## Contents

1 Introduction ......................................................................................... 4
  1.1 Description .................................................................................. 4
  1.2 How It Works ............................................................................... 4
  1.3 Processing Documents .................................................................. 5
  1.4 Integration Options ...................................................................... 5
  1.5 Features ....................................................................................... 6
  1.6 Operating Systems ....................................................................... 7
  1.7 Compatible Microsoft Office Versions .......................................... 7

2 Quick Start Guide .................................................................................. 8
  2.1 Installation ................................................................................... 8
  2.2 Convert your first file using the preinstalled configuration ............. 8
  2.3 Configure for your scenario .......................................................... 8
  2.4 Integrate into your system ............................................................. 8

3 License Management ............................................................................. 10

4 User’s Guide .......................................................................................... 11
  4.1 Service Configuration .................................................................... 11
  4.2 Profile Configuration .................................................................... 13
  4.3 Connection Configuration ............................................................ 16
  4.4 Office Configuration ..................................................................... 19
    4.4.1 Office User Configuration .................................................... 19
      Let the Conversion Service create a new local user ..................... 19
      Use an existing domain user ..................................................... 19
    4.4.2 Microsoft Office License Considerations ............................. 19
  4.5 Health & Activity .......................................................................... 20
  4.5.1 Jobs, Tasks and Parallel Execution .......................................... 20
  4.6 Statistics ...................................................................................... 22
  4.7 Clients .......................................................................................... 25
    4.7.1 Client Installation ............................................................... 25
    4.7.2 Shell Client pdfclient .......................................................... 25
    4.7.3 GUI Client ............................................................................ 26
      Settings Window ......................................................................... 27
      Pre-Configuration ...................................................................... 28
      Set default service endpoint .................................................... 28
      Add workflows to blacklist ...................................................... 28
    4.7.4 Office Add-In ....................................................................... 28
  4.8 REST API ..................................................................................... 28
  4.8.1 API Usage ............................................................................... 29
  4.9 Plugins .......................................................................................... 31

5 Workflows ............................................................................................. 32
  5.1 Archive PDF/A-2 .......................................................................... 32
    5.1.1 Supported File Formats ....................................................... 32
    5.1.2 Process ............................................................................... 32
    5.1.3 Configuration ................................................................. 33
  5.2 Archive PDF/A-3 .......................................................................... 35
    5.2.1 Supported File Formats ....................................................... 35

© PDF Tools AG – Premium PDF Technology
4-Heights® Conversion Service, August 4, 2022 | 1/65
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2.4</td>
<td>Execute Command</td>
</tr>
<tr>
<td>6.2.5</td>
<td>Output Mailbox (IMAP)</td>
</tr>
<tr>
<td>6.2.6</td>
<td>Output Mailbox (Exchange Online)</td>
</tr>
<tr>
<td>6.2.7</td>
<td>Send Email (SMTP)</td>
</tr>
<tr>
<td>6.2.8</td>
<td>Send Email (Exchange Online)</td>
</tr>
<tr>
<td>7</td>
<td>Troubleshooting</td>
</tr>
<tr>
<td>7.1</td>
<td>Error Codes</td>
</tr>
<tr>
<td>7.2</td>
<td>Office Conversion Problems</td>
</tr>
<tr>
<td>7.2.1</td>
<td>Configuration Problems</td>
</tr>
<tr>
<td>7.2.2</td>
<td>Specific errors related to Word, Excel and PowerPoint documents</td>
</tr>
<tr>
<td>7.3</td>
<td>Service Log</td>
</tr>
<tr>
<td>7.4</td>
<td>Submitting a Support Request</td>
</tr>
<tr>
<td>8</td>
<td>Version History</td>
</tr>
<tr>
<td>8.1</td>
<td>Version 3</td>
</tr>
<tr>
<td>8.2</td>
<td>Version 2</td>
</tr>
<tr>
<td>8.3</td>
<td>Version 1</td>
</tr>
<tr>
<td>9</td>
<td>Licensing, Copyright, and Contact</td>
</tr>
</tbody>
</table>
1 Introduction

1.1 Description

The 4-Heights® Conversion Service is an all-round carefree package for automating your document processes. It takes input documents of different formats from various sources and processes them according to your use case to produce high-quality PDF/A documents.

1.2 How It Works

The 4-Heights® Conversion Service can be thought of as a factory with a production line (workflow) for each product it offers. The products it manufactures are documents prepared for a specific use case. For example, the workflow Archive PDF/A-2 is engineered specifically for preparing documents for archiving.
A production line consists of a series of processing steps (e.g. validate, OCR, convert, merge, sign, …), each of which moves the raw materials (input documents) closer to being a finished product (output document).

Some steps are optional and most of them have options (e.g. paper size, image resolution, …). These can be adjusted to your individual needs by configuring one or more setups (profiles) for each production line. A profile defines which processing steps to activate and what options to use.

1.3 Processing Documents

If you want the factory to manufacture a product (output document), you send them an order (job) consisting of the raw materials (input documents), the production line (workflow) and setup (profile) to use and where to deliver the finished product (output path).

Input

Similarly, to prepare documents for a specific use case, you send a job consisting of the following input to the 4-Heights® Conversion Service:

Input Documents The documents to be processed.

Workflow Name Which workflow to use, i.e. the use case to prepare the documents for.

Profile Name Which profile to use, i.e. how the documents should be processed.

Output Path Where to store the resulting document.

Output

Once conversion is completed, the 4-Heights® Conversion Service returns:

Output Document If the conversion was successful, the resulting document can be found in the specified output directory.

Report Conversion events, warnings and errors. Useful for analysis.

1.4 Integration Options

The communication with the factory can happen through different channels. Maybe you send the order by mail. Or you bring everything to the factory in person. Similarly, the 4-Heights® Conversion Service offers different interfaces to communicate through.

The 4-Heights® Conversion Service is a Windows service designed to be installed on a server and integrated into your system using one of the following interfaces:
**Watched Folder**  Automatically convert documents copied to a pre-configured input folder to a pre-configured output folder. Recommended for integrating with file share architectures. Suited for interactive and manually triggered processing.

**Shell Client**  Easy to use shell tool. Recommended for integrating using scripts. Suited for automated and manually triggered processing.

**REST API**  Easy to use webservice interface. Recommended for integrating into your existing application (Enterprise Application Integration). Suited for automated and manually triggered processing from within your application.

**GUI Client**  Easy to use graphical tool. Recommended for testing and familiarizing yourself with the 4-Heights® Conversion Service. Suited for manually triggered processing only.

### 1.5 Features

- Robust client/server architecture
- Powerful REST API (HTTP or HTTPS)
- High throughput job processing
- Specialized workflows
  - Archive PDF/A-2
  - Archive PDF/A-3
  - Archive PDF/A-1
  - Conversion
- Invoice
- Dossier
- Multiple configuration profiles for each workflow
- Convenient GUI configurator
- Watched folder
- Statistics of activity and job processing
- High availability and stability
- Logging and monitoring

1.6 Operating Systems

The 4-Heights® Conversion Service is available for the following operating systems:

- Windows Server 2016, 2019, 2022 | x64

1.7 Compatible Microsoft Office Versions

As an optionally available feature Word, Excel and PowerPoint documents can be converted to PDF with the 4-Heights® Conversion Service. The conversion of Word, Excel and PowerPoint documents requires a Microsoft Office installation on the same server where the Conversion Service is running. Compatible Microsoft Office versions are:

- Office 2016 (64 Bit)
- Office 2019 (64 Bit)
- Office 2021 (64 Bit)
- Office 365 (64 Bit)

US English is the recommended and supported language setting.
2 Quick Start Guide

The setup is simple: you just install the 4-Heights® Conversion Service on your server and you are all set for your first conversion.

2.1 Installation

1. Download and install the required .NET runtimes:
   1. .NET Framework 4.7
   2. Microsoft .NET Desktop Runtime 6.0 (x64)
   3. Microsoft ASP.NET Core Runtime 6.0 (x64)
2. Download and execute the MSI installation package and follow the instructions on screen. This will install the following components:
   - Conversion Service
   - Configurator GUI
   - Shell and GUI Clients
   - License Manager
   - Watched Folder
3. Add your license in the license manager as described in chapter License Management.

**Note:** To convert Word, Excel and PowerPoint documents, an Office User must be configured, as described in Office Configuration. Also, Office conversion must be activated in the profile, e.g. using the Configurator GUI as described in Profile Configuration.

2.2 Convert your first file using the preinstalled configuration

1. Copy a document, e.g. input.pdf, to the server you just installed the service on.
2. Open a command prompt and type

   ```
   pdfclient -v input.pdf output.pdf
   ```

3. The client will show the progress and when it is finished you will find the newly generated output.pdf in the current working directory.

2.3 Configure for your scenario

- Configure your service as described in Service Configuration.
- Adapt the default profile or create new ones as described in Profile Configuration.

2.4 Integrate into your system

Choose your integration:

- Shell Client pdfclient
- Watched Folder
REST API
3 License Management

The 4-Heights® Conversion Service requires a valid license in order to run correctly. More information about license management is available in the license key technote.
4 User’s Guide

This section gives an overview of the usage and configuration of the 4-Heights® Conversion Service features and interfaces.

4.1 Service Configuration

The service can be configured using the 4-Heights® Conversion Service Configurator GUI, which is added to the Windows start menu during installation.

**Service configuration tab of the 4-Heights® Conversion Service Configurator**

- **Status**
  - Shows the status of the 4-Heights® Conversion Service and provides the controls to start, stop and restart the Windows service.
  - For the 4-Heights® Conversion Service to be “running”, the Windows service must be started and its configuration must be valid.
  - If the 4-Heights® Conversion Service is not operable even though the Windows service has been started, check the log file for more information on the cause. Common problems are a missing license or low disk space.
Note that whenever the configuration is changed, the Windows service must be restarted for the changes to take effect. When the service is stopped, all current jobs are canceled and deleted.

The Windows service is installed with service name ConversionService and display name “4-Heights® Conversion Service”. So its status can also be changed using Windows’ “Services” desktop app or the sc command line program.

**Service Host Address**  The address of the service host of the form

http[s]://<hostname>:<port>/conversion/v1.0/rest.

Default: http://localhost:13033/conversion/v1.0/rest

It is recommended to not change the address, unless the default port 13033 is already in use on the local machine. If the port is changed and firewall rules have been added, they must be changed accordingly. The suffix conversion/v1.0/rest is not configurable.

The configured service host address is also the endpoint URL used by the Shell Client pdfclient or REST API.

**HTTPS**  By default the service endpoint uses HTTP. Activating HTTPS will disable support for HTTP. This is to prevent clients from accidentally sending sensitive information over HTTP.

HTTPS can be activated by setting the scheme of the service host address to https. In addition, the following is required:

1. A valid host certificate. The certificate must be provided as PKCS#12 file (.pfx or .p12) that includes the certificate’s private key and issuer certificates. If the private key is password protected, the password must be configured.
2. The <hostname> of the service host address must be set to the name of the host certificate. Otherwise clients, such as the Shell Client pdfclient, will refuse to connect. This hinders correct operation of many components, including the Configurator GUI or the Connection Configuration.
3. The host certificate must be trusted.

**Windows Defender Rules**  Exclude the service processes and working directories from monitoring by the Windows Defender Antivirus. It is highly recommended to enable this setting, because the service might otherwise be slowed down significantly.

**Open Port**  The port configured in the Service Host Address is by default available from the local machine only. If the 4-Heights® Conversion Service should be accessible from other machines, e.g. using the Shell Client pdfclient or REST API, the port must be opened in the Windows Firewall.

This switch opens the port for connections from all other computers. If more specific rules are required, they must be created manually (see chapter Security Considerations).

**Temporary Files**  The directory where temporary files are written to. To optimize performance, this directory should be on a local drive with fast read and write access.

**Proxy**  Configuration of the proxy url to be used for all http and https communication to external hosts. The option can either be left empty for no proxy or must be a string with the following syntax:

http[s]://[<user>:<password>@]<host>[:<port>]

Where:

- <user>:<password> (optional): Credentials for connection to proxy (basic authorization).
- <host>: Hostname of proxy.
- <port>: Port for connection to proxy.

**Office User**  Configure the dedicated user account required for Word, Excel and PowerPoint to PDF conversion. Click on Configure and follow the on-screen instructions, further details are explained in Office Configuration.
**Log Directory**  The directory where all log files are written to. This includes the log files of the service (see Service Log), the connectors and the configurator.

**Statistics**  Configure whether the service stores service, job and task information to a database file to analyze the conversion history with the Statistics tab. The location of the database is configured in the Database Directory setting.

**Database Directory**  The directory where the database is stored when the setting Statistics is enabled. This setting is only configurable when the setting Statistics is turned on.

**Max Age**  The configured maximum age of the database entries. Information of the conversion history will be deleted after reaching the maximum age. This setting is only configurable when the setting Statistics is turned on.

**Max File Size**  The configured maximum file size of the database. Older entries of the conversion history will be deleted if the maximum database file size is exceeded. This setting is only configurable when the setting Statistics is turned on.

### 4.2 Profile Configuration

Profiles are managed using the 4-Heights® Conversion Service Configurator GUI, which is added to the Windows start menu during installation.

![Workflows & Profiles tab of the 4-Heights® Conversion Service Configurator](image)

**Workflows**  Workflows can be activated or deactivated by clicking the toggle button right next to the workflow name. Your changes are kept in memory, until you click Save & Restart Service in the notification.
Profiles  A profile defines all options and optional processing steps for a given workflow, see Workflows. Each workflow is delivered with a default profile called “default”. To add and configure a new profile click the Add Profile button. To edit an existing profile simply click on it. To copy an existing profile click the Copy icon.

Import Window of 4-Heights®
Conversion Service Configurator

Import Profiles

Deselect All
☑ Archive PDF/A-2
☑ default
☑ Profile2

![Duplicate profiles selected. What should happen during import?](image) Replace originals with imported profiles
☑ Keep originals and imported profiles

[Deselect duplicates] [Cancel] [Import]

Import Profiles  Click Import Profiles in the top-right menu to import profiles from a profile configuration file. If a profile with the same name is selected in the same workflow as an already existing profile, a dialog with the following options appears:

- **Deselect duplicates**: Select this option if you only want to add, but not replace any existing profiles.
- **Replace originals with imported profiles**: Select this option if you want to replace the existing profiles.
- **Keep originals and imported profiles**: Select this option if you want to keep both, the existing and the imported profiles.

When importing profiles of a workflow, the following workflow properties are also imported:

- **Workflow state**: Whether the workflow is activated or deactivated.
Export Profiles  Click Export Profiles in the top-right menu to export one or more profiles to a profile configuration file.

Configure a new profile with the 4-Heights® Conversion Service Configurator

Name (required)  Enter the desired name for the profile.

Default  Check this box to set the current profile as the default profile for the given workflow.
**Processing Steps**  Enable/disable optional processing steps. Corresponding options will be displayed accordingly.

**Options**  Set options for individual processing steps. A summary for each option is documented with tooltips. For a detailed explanation of the options, click on the title of the corresponding section to open the documentation panel.

Any configuration change requires a restart of the 4-Heights® Conversion Service in order for it to take effect. When you add or edit profiles, the Configurator detects changes and displays an unobtrusive notification. For your convenience, your changes are kept in memory, until you click **Save & Restart Service** in the notification.

### 4.3 Connection Configuration

Easily integrate the 4-Heights® Conversion Service into the environment by connecting input sources (e.g. a watched folder) to output destinations (e.g. an output folder) and configuring how documents should be processed along the way. To achieve this configure a connection using the 4-Heights® Conversion Service Configurator, which is added to the Windows start menu during installation.

**Integration tab of the 4-Heights® Conversion Service Configurator**

- **Connection**  A connection defines how the 4-Heights® Conversion Service is integrated into an existing system. Each connection requires input and output connectors. To add and configure a new connection click the **Add Connection** button. To edit or delete an existing connection, click on the respective icon. To copy an existing connection click the **Copy** icon.
Configure a new connection with the 4-Heights® Conversion Service Configurator

**Name (required)**  Enter a connection name.

**Description (optional)**  Enter a connection description.

**Input(s) (required)**  Configure an input connector by clicking **Add**.

**Processing (required)**  Select the workflow and profile for the conversion.

**Output(s) (required)**  Configure an output connector by clicking **Add**.
Add Connector  Click **Add** in the desired section to add an input or output connector. For a detailed explanation of the connector, click on the title to open the documentation panel. Select the desired connector and click **Next** to confirm the selection.

Configure Connector Window of 4-Heights® Conversion Service Configurator
Configure Connector  Set options for the selected connector. A summary for each option is documented with tooltips. For a detailed explanation of the connector, click on the title to open the documentation panel. See Connectors / Integration for more information on the individual connectors.

Any configuration change requires a restart of the 4-Heights® Conversion Service in order for it to take effect. When you add or edit connections, the Configurator detects changes and displays an unobtrusive notification. For your convenience, your changes are kept in memory, until you click Save & Restart Service.

4.4 Office Configuration

The 4-Heights® Conversion Service uses Microsoft Office Applications to convert Word, Excel and PowerPoint documents to PDF. A dedicated user account is necessary to run and interact with Microsoft Office Applications. We refer to this user as the Office User.

4.4.1 Office User Configuration

1. Open the Configurator
2. Navigate to the Service configuration tab of the 4-Heights® Conversion Service Configurator
3. Click on Office User and configure one of the following two options:
   - Let the Conversion Service create a new local user
   - Use an existing domain user

Let the Conversion Service create a new local user

In this case the Conversion Service creates and configures a local user account ConvsrvOfficeUser automatically.

- The account does not have any group memberships.
- A complex random password with 36 characters is created and encrypted using the Windows Data Protection API. The password does not expire.
- The "Log on as a batch job" Security Policy Setting is granted to this user.
- The Registry and DCOM Permissions of Word, Excel and PowerPoint are adjusted.

Use an existing domain user

In this case you need to provide an existing domain user account. The account is not managed by the configurator. The following properties and settings are set automatically.

- The password is encrypted using the Windows Data Protection API before it is stored to a configuration file.
- The "Log on as a batch job" Security Policy Setting is granted to this user.
- The Registry and DCOM Permissions of Word, Excel and PowerPoint are adjusted.

4.4.2 Microsoft Office License Considerations

The 4-Heights® Conversion Service uses Microsoft Office applications to process jobs containing Word, Excel and PowerPoint documents. For compliance with the licensing guidelines from Microsoft, it is required that every client sending jobs with Word, Excel or PowerPoint documents to the Conversion Service is licensed for Microsoft Office.

---

1 If you update the password you need to reconfigure the Office User
2 For further details we refer to the Documentation from Microsoft, last accessed November 2020
4.5 Health & Activity

The state and recent activity of the service is monitored by the Health & Activity tab of the 4-Heights® Conversion Service Configurator GUI, which is added to the Windows start menu during installation. The purpose is to provide an overview whether the service has been doing well so far. It consists of the following parts:

- The number jobs the service currently processes
- Time passed since when the service was started
- Validity of the license
- Pie chart about the results of jobs processed on the current day
- Time chart tracking service's load and performance

4.5.1 Jobs, Tasks and Parallel Execution

A job is the unit that is sent to and processed by the Conversion Service. A job consists of input documents, a workflow name, a profile name and an output path. Depending on the job content a set of processing steps is required to complete its execution (such as destructuring, conversion, merging, etc). A task is the execution unit of such a processing step. Tasks are generally executed in parallel and tasks are held in a queue in case the service is charged to capacity. The parallel processing capacity of the service depends on the license and the underlying hardware system. The license and hardware therefore limit the number of tasks that can be processed in parallel by the service. We refer to this number as the maximum number of concurrent tasks (maxConcurrentTasks).
**Health & Activity tab of the 4-Heights® Conversion Service Configurator**

**Status**  Summarizes the status of the 4-Heights® Conversion Service by displaying the count of currently processed jobs, status of the license, service uptime.

**Processed Jobs**  The pie chart summarizes and monitors jobs processed during the current day in the following categories:

- **Jobs successful**: Jobs converted successfully.
- **Jobs with warnings**: Jobs converted with warnings.
- **Jobs failed**: Jobs that failed with errors related to the input document like for example Corrupt, UnsupportedFormat or Generic.
- **Jobs with errors**: Jobs that failed due to a Timeout, Configuration or Internal error.
- **Jobs canceled**: Jobs that were canceled by the user.

**Activity**  The time chart is an indicator for the service’s load and performance. It displays the following two curves which are limited by the parallel processing capacity of the service, see explanation in Jobs, Tasks and Parallel Execution:

- **Utilization** (green): Plots the percentage of processing capability used. Note that the total possible processing capability of the service depends on your license and your system.
- **Number of Tasks in Queue** (yellow): Shows the number of tasks are waiting in the queue for being processed.
On top of the diagram, on the very right, there are four numbers, each next to a symbol. These values are in the color of the curve they are associated with. The top green value is the maximum, the bottom green value is the average of the utilization. The top yellow value is the maximum number of tasks in queue for the selected time span, the bottom yellow value is the average time a task spends waiting in the queue before being processed.

Next to the title 'Activity', the following time spans can be configured for the chart:
- Last ten minutes
- Last hour
- Last two hours
- Last four hours

When you change the selected item in the legend at the bottom of the chart, the curve it represents is brought to the foreground in addition with its associated grid for easier reading. Note that for the y-Axis, the range of values is labeled at the right hand side.

### 4.6 Statistics

Information about the past activity of the service is reported by the Statistics tab of the the 4-Heights® Conversion Service Configurator GUI, which is added to the Windows start menu during installation. With this tab the conversion history of the service can be analyzed. The Statistics tab consists of the following parts:

- **Filter section**: A section containing a set of filters to control the data displayed in the tab.
- **Simple Statistics**: An overview displaying the performance figure of the service.
- **Processed Jobs**: Pie chart about the results of the processed jobs.
- **Activity**: Time chart indicating service's load and performance for a selected group of jobs

In addition, the data displayed in the tab can be exported for further analysis.
Statistics tab of the 4-Heights® Conversion Service Configurator

Filter section  In the top row of the tab, below its title, the following filters can be configured to be applied to the data displayed in the tab:

**Time Unit & Period**  Determines the period of time when the jobs were processed. On the left hand side, the length of the period can be configured by selecting one of the following time units:
- Day
- Week
- Month

On the right hand side, the point of time where the period occurred can be selected either by clicking on the calendar icon and choosing it from the calendar dropdown which has appeared or by clicking the arrows.

**Workflows**  Filter data to show usage, load and job results of certain workflows.

**Query Time**  The time point when the query for the displayed data was made. The query time is only relevant if the current day is part of the selected time period. The query time can be updated with the refresh icon in this case.
Export Data as CSV  Click Export Data as CSV in the top-right menu to export the data selected by the filters from the database. When the export has finished, a ZIP archive containing the following CSV files has been saved:

- **service.csv** contains the following columns:
  - `started`: Service start time.
  - `stopped`: Service stop time.
  - `maxConcurrentTasks`: Concurrency supported by the license, i.e. the maximum number of tasks which can be executed in parallel.

- **jobs.csv**
  - `id`: Row number.
  - `jobId`: Unique identifier name of the job.
  - `worklow`: Name of the workflow the job was processed with.
  - `profile`: Name of the profile the job was processed with.
  - `started`: Point in time when the service started to process the job.
  - `finished`: Point in time when the service completed the job.
  - `error`: Name of the error if the job failed.
  - `warnings`: List of the names of the warnings that occurred when processing the job.
  - `outputFilesCount`: Total number of output files resulting from the conversion.
  - `overallOutputSize`: Total number of bytes summed over all output files resulting from the conversion.
  - `overallPages`: Total count of pages summed over all output files resulting from the conversion.

- **tasks.csv**
  - `id`: Row number.
  - `job`: Row number of the job the task is associated with.
  - `queued`: Point in time when the task began waiting in the queue for being processed.
  - `started`: Point in time when the task moved out of the queue and the service started to process it.
  - `stopped`: Point in time when the service was finished with processing the task.
  - `component`: Names the type of processing step applied through this task.

**Simple Statistics**  Displays some characteristic numbers about the jobs considered by the current filter configuration and the service:

- **Service Uptime**: The total amount of time the service was running.
- **Processed Jobs**: The total number of jobs that have been processed.
- **Processed Pages**: The total sum of pages of documents converted to PDF.
- **Processing Time**: The total sum of job processing time. Note that this number can be higher than Service Uptime since multiple tasks can be processed in parallel.
- **Output File Size**: The total sum of output file size of the processed documents.

**Processed Jobs**  The pie chart which is also described in Section 4.5 summarizes the jobs considered by the current filter configuration in the following categories:

- **Jobs successful**: Jobs converted successfully.
- **Jobs with warnings**: Jobs converted with warnings.
- **Jobs failed**: Jobs that failed with errors related to the input document like for example Corrupt, UnsupportedFormat or Generic.
- **Jobs with errors**: Jobs that failed due to a Timeout, Configuration or Internal error.
- **Jobs canceled**: Jobs that where canceled by the user.

**Activity**  The time chart which is described in Section 4.5 as well is an indicator for the service's load and performance. The filters apply to this diagram as well. It displays the following two curves:

- **Utilization** (green): Plots the percentage of processing capability used. Note that the total possible processing capability of the service depends on your license and your system.

---

\[1\] Refer to Section 4.5.1 for more information.
- **Number of Tasks in Queue** (yellow): Shows the number of tasks are waiting in the queue for being processed.

On top of the diagram, on the very right, there are four numbers, each next to a symbol. These values are in the color of the curve they are associated with. The top green value is the maximum, the bottom green value is the average of the utilization. The top yellow value is the maximum number of tasks in queue for the selected period, the bottom yellow value is the average time a task spends waiting in the queue before being processed.

When you change the selected item in the legend at the bottom of the chart, the curve it represents is brought to the foreground in addition with its associated grid for easier reading. Note that for the y-Axis, the range of values is labeled at the right hand side.

## 4.7 Clients

### 4.7.1 Client Installation

The ConversionService-Client.msi offers a standalone installation for the client applications. For automated installations, the Property `SERVICE_URL` can be used to preconfigure the service url of the 4-Heights® Conversion Service.

```bash
msiexec /i ConversionService.msi SERVICE_URL=http://localhost:13033/conversion/v1.0/rest
```

The installation of the Office Add-In can be disabled in the installer.

### 4.7.2 Shell Client pdfclient

The shell client application called pdfclient interacts with the 4-Heights® Conversion Service via the [REST API](https://example.com/restapi).

The application creates jobs based on its input parameters. The jobs are then uploaded in parallel to the 4-Heights® Conversion Service and the shell client monitors the completion, before the output files are downloaded. In verbose mode (option `-v`), a detailed report including all performed actions on the documents is written to the console. In case warnings or errors occured during processing, they are reported as well.

A selection of examples is explained in detail below.

### Examples

#### Show Usage

```bash
pdfclient
```

By typing `pdfclient` without parameters on the command line, the usage and examples are shown:

#### Basic

```bash
pdfclient -v -url http://<hostname>:<port>/conversion/v1.0/rest file.pdf
-pw <password> encrypted.pdf output.pdf
```

---

4 The pdfclient is added to the PATH environment setting during installation.
Sends the files `file.pdf` and `encrypted.pdf` to the 4-Heights® Conversion Service through the REST API endpoint url `http://<hostname>:<port>/conversion/v1.0/rest` (see Service Host Address) and saves the result as `output.pdf`. Option `-v` turns on verbose mode, option `-url` sets the service endpoint url and option `-pw` supplies the password for `encrypted.pdf`. If the service endpoint url (`-url`) is not set, "localhost" and the default port are used.

**Workflow and Profile**

```
pdfclient -w "Archive PDF/A-2" -p "myCustomProfile" input.pdf output.pdf
```

Use options `-w` and `-p` to select a specific workflow and profile for the job. In this example the workflow is set to "Archive PDF/A-2" and the profile is set to "myCustomProfile". If the options are not specified the service's configured default workflow and profile are used.

**Wildcards**

```
pdfclient -r *.ext C:\path\to\output.pdf
```

Sends all files with extension `.ext` from the current working directory and all subdirectories to the 4-Heights® Conversion Service and saves the result as a single output document `C:\path\to\output.pdf`.

```
pdfclient -s *.ext C:\path\to\output
```

Process the files with extension `.ext` as separate jobs and save the results in `C:\path\to\output`. The output directory must already exist. Existing output files are overwritten by default. Output options such as creating unique filenames instead of overwriting existing files are explained in the usage of the tool.

**Batch Processing**

```
pdfclient -r -s C:\path\to\input C:\path\to\output
```

Sends the files inside `C:\path\to\input` including all subdirectories to the 4-Heights® Conversion Service, and saves the results in `C:\path\to\output` with the same file structure as in the input directory. Existing output files are overwritten. Output options such as creating unique filenames instead of overwriting existing files are explained in the usage of the tool.

**4.7.3 GUI Client**

The GUI Client is an easy to use graphical tool intended for testing and familiarizing yourself with the 4-Heights® Conversion Service. It interacts with the 4-Heights® Conversion Service via REST API.
**4-Heights® Conversion Service GUI Client**

**Input**  Drag & drop the documents to be processed to the input area. To remove individual documents from the list select them and press the **delete** or **back** key. To remove all documents from the list click on **Clear List**.

**Conversion Options**  Select the workflow and the profile to be used for the conversion. Click the refresh icon to reload the workflows and profiles. The **merge** option controls the processing mode. If it is switched on, all input documents are gathered into one job, resulting in one merged output document. If it is switched off, every document gets processed in a separate job, resulting in one output document per input document.

**Output**  Choose where to store the resulting document(s). If **merge** is switched on, choose the filename of the resulting document.

**Settings Window**

The service endpoint URL can be changed in the settings window of the 4-Heights® Conversion Service GUI Client. The default value is `http://localhost:13033/conversion/v1.0/rest`.

![Settings window of 4-Heights® Conversion Service GUI Client](image)
Pre-Configuration

Set default service endpoint

To set a default service endpoint URL, add it to the serviceEndpoint in the project's appsettings.json file located in the bin directory of the installation path.

```
{
    "service": {
        "serviceEndpoint": "http://localhost:13033/conversion/v1.0/rest"
    }
}
```

Add workflows to blacklist

Individual workflows can be blacklisted in the project's appsettings.json file in the directory bin of the installation directory, effectively making them invisible in the UI. To blacklist a specific workflow, add it to the blacklistWorkflows in the aforementioned file.

```
{
    "client": {
        "blacklistWorkflows": [
            "Dossier",
            "Invoice"
        ]
    }
}
```

4.7.4 Office Add-In

The Office Add-Ins are a simple way for converting documents directly from an Office Application. A 4-Heights® Conversion Service button is added to the toolbar, through which the conversion can be started immediately.

If not disabled in the ConversionService-Client.msi, the add-ins are installed for the following Office products:
- Microsoft Word
- Microsoft Outlook
- Microsoft Excel
- Microsoft PowerPoint

4.8 REST API

The 4-Heights® Conversion Service offers a REST API that allows to schedule jobs and get service status information. The REST API is also used by the other clients, e.g. the Shell Client pdfclient or Watched Folder.

The service endpoint URL of the REST API is defined by the Service Host Address of the service configuration. The default value is http://localhost:13033/conversion/v1.0/rest.
Note: For easy integration with existing systems, a simplified REST interface is provided by the input connectors REST Input (Plain HTTP) and REST Input (JSON).

Security Considerations

The API is by default available on the local machine only. If it should be accessible remotely, the computer's firewall must be configured accordingly, e.g. using the Configurator GUI. When opening the port in the firewall it is recommended to add a rule that is as strict as possible, i.e. to not allow connections from untrusted computers.

Note that the REST API is designed for use in a protected intranet only. It offers no user authentication nor other security measures, e.g. against denial-of-service attacks. If this is required, a web application firewall (WAF) is recommended.

The 4-Heights® Conversion Service does support Cross-Origin Requests (CORS). This is required when sending requests using JavaScript from a browser. By default, requests from all origins are allowed. If necessary, cross-origin requests can be restricted to certain origins using the service's configuration.

4.8.1 API Usage

The API is described in detail by the OpenAPI document `doc/openapi.yaml` inside the installation directory. The YAML document can be viewed in an OpenAPI editor or used to generate client stub code of any programming language.

The API supports XML and JSON in the bodies of requests and responses. So it is recommended to set the headers `Content-Type` and `Accept` to the preferred type.

In case of an error, the API returns a suitable HTTP status code as well as a problem details object (RFC 7807). This object contains more information on the type and cause of the error. Notably the object's property `detail` contains a human-readable explanation that is helpful to troubleshoot the issue. Therefore, it is recommended to parse and use the problem details object (Content-Type `application/problem+json` and `application/problem+xml`) whenever the returned HTTP status code indicates an error.

Job Processing

For a general overview of how jobs are processed by the 4-Heights® Conversion Service, see chapter Processing Documents or the documentation of a specific workflow, such as the Archive PDF/A-2.

To schedule a job and retrieve its result, the following simplified sequence described in Job Processing Sequence can be used.
There is no limit of the number of jobs that can be started concurrently. The service will process the jobs in the order they were created, using the highest concurrency allowed by the system’s CPU and the license. Nonetheless, it is recommended to not start much more jobs than the service can process. For example, on a machine with 8 CPU cores and a license for 8 cores, not much more than 8 jobs should be started.

Service Information

These methods return service status information.

The `getServiceStatus` method can be used to retrieve general status information. This is suitable to verify if the service is running, e.g. for health check monitoring. In addition, information on the service’s load and general job count information is returned.

The `listJobs` method returns a list of all jobs and their status. This is useful to see what tasks are executing.

The `listWorkflows` method returns a list of all workflows and their profiles.

---

5 Note that this is a very simplified example. To determine the maximum concurrency, the whole system and configuration must be taken into consideration. Dependent systems, such as an OCR service or Office conversion might further limit the maximum concurrency.
4.9 Plugins

Plugins are non-standard components used for extending the 4-Heights® Conversion Service with custom workflows. They are managed in the plugins tab of the 4-Heights® Conversion Service Configurator GUI. This tab is only displayed if the corresponding license feature is active. All installed plugins are listed here.

To install a new plugin, click **Install Plugin** and choose the ZIP file containing it. To update or delete a plugin, click on the respective icon. These actions require a service restart to take effect.

Each plugin version is designed for a specific version of the 4-Heights® Conversion Service. So when upgrading the 4-Heights® Conversion Service, all plugins must be upgraded as well. Unless the plugins are updated, the service is not operational.

Plugins tab of the 4-Heights® Conversion Service Configurator
5 Workflows

5.1 Archive PDF/A-2

This workflow is engineered specifically for preparing documents for archiving.

5.1.1 Supported File Formats

This workflow supports all file formats listed in Supported File Formats. The conversion of most file formats is enabled by default in the Convert Mode Configuration for Child Documents (Attachments).

5.1.2 Process

All input documents of a job are processed as follows:

[Diagram of the process flow]
Analyze Filetype  In a first step, the documents are analyzed. If the filetype of any document is unsupported, conversion is aborted. Otherwise, they are sent to the next processing step depending on their filetype.

Validate & Repair (Quality Assurance)  To ensure document quality, PDF and PDF/A documents are validated. If a corruption is detected, the 4-Heights® Conversion Service attempts to repair it.

Convert to PDF  Non-PDF documents (e.g. images, Office documents, …) are converted to PDF if their format is supported by the 4-Heights® Conversion Service.

Note: the conversion of Office documents requires an additional step which can only be enabled with an appropriate license.

OCR  In order to make the resulting document searchable, OCR is performed on documents that require it. The recognized text is stored directly in the PDF.

Note: this is an optional step and can only be enabled with an appropriate license.

Stamping  See Stamping.

Convert to PDF/A  PDF documents that are not already PDF/A-2 conforming are converted to a high-quality PDF/A-2.

Merge / Collect  The converted documents of a job are merged or collected into one document, depending on the profile setting.

Optimize  The resulting document is optimized for archiving. This includes several optimizations: redundant and unnecessary data for archiving is removed, images are compressed intelligently and fonts are merged and subset.

Note: this is an optional step and can only be enabled with an appropriate license.

Sign  The resulting document is digitally signed using the signature settings in the selected profile.

Note: this is an optional step and can only be enabled with an appropriate license.

5.1.3 Configuration

The workflow’s profile offers a fine-grained configuration of how files are converted. All of the processing steps described above can be enabled and configured in the profile configuration. Furthermore, the following configuration options are available:

Convert Mode Configuration for Child Documents (Attachments)  The convert mode defines, which documents are converted (configuration "Convert") and which are skipped (removed) from the result. When removing documents, a warning (configuration “Skip with Warning”) or an informational message (configuration “Skip”) is generated.

The convert mode can be specified based on the type of the child document, its filename, or the type of its parent document.

For example, by default Office files are converted to PDF/A, executables are removed, and other non-convertible documents are removed with a warning.

Collect Mode Configuration  The collect mode configuration defines how a converted document and its child documents are combined.

The collect mode can be specified for each document type individually. Please refer to the documentation panel of the 4-Heights® Conversion Service Configurator for a detailed description of the available collect modes.
For example, emails can be converted by creating a PDF collection (Portfolio) of its body and attachments. When converting Word documents, their embedded files can be attached to the converted PDF.

**Merge vs. Attach** There are two categories of collect modes. Either the pages of multiple PDF documents can be merged into a single document, or child documents can be attached (embedded) into a PDF document. A PDF Collection (Portfolio) is a special case of the latter, where the parent document contains no pages, but shows a convenient table of the attached documents for easy navigation.

The advantage of the **Merge** collect modes is, that this creates simple files that can be processed and viewed by all PDF applications. The disadvantage is that only PDF files can be merged. Furthermore, not all information can be preserved when merging PDF files. For example, document metadata, signatures, and certain interactive form fields cannot be merged and must be removed. Also, logical structure (tagging) information might be less meaningful after merging.

Recommended collect mode configuration for the “Merge” use case:

<table>
<thead>
<tr>
<th>Type</th>
<th>Collect Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Job</td>
<td>Merge</td>
</tr>
<tr>
<td>Word</td>
<td>Merge or Attach</td>
</tr>
<tr>
<td>Excel</td>
<td>Merge or Attach</td>
</tr>
<tr>
<td>PowerPoint</td>
<td>Merge or Attach</td>
</tr>
<tr>
<td>Email</td>
<td>Merge or Attach</td>
</tr>
<tr>
<td>Archive (ZIP)</td>
<td>Merge or Attach</td>
</tr>
<tr>
<td>PDF</td>
<td>Preserve Structure</td>
</tr>
</tbody>
</table>

The advantages of the **Attach** collect modes is, that all information of the input files and the files’ structure can be preserved. Especially PDF Collections (Portfolios) provide a convenient way to show and navigate the attached documents.

Recommended collect mode configuration for the “Attach” use case:

<table>
<thead>
<tr>
<th>Type</th>
<th>Collect Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Job</td>
<td>Collection or Single Document</td>
</tr>
<tr>
<td>Word</td>
<td>Attach</td>
</tr>
<tr>
<td>Excel</td>
<td>Attach</td>
</tr>
<tr>
<td>PowerPoint</td>
<td>Attach</td>
</tr>
<tr>
<td>Email</td>
<td>Attach</td>
</tr>
<tr>
<td>Archive (ZIP)</td>
<td>Collection or Single Document</td>
</tr>
<tr>
<td>PDF</td>
<td>Preserve Structure①</td>
</tr>
</tbody>
</table>

① If necessary, the collect mode “Flatten” can be used to flatten the structure of PDF documents. Up to version 3.1 of the 4-Heights® Conversion Service this has been the default behavior.
The **Child Error Handling** configuration defines how errors during the conversion of child documents are handled. In case of an error, the child document can either be skipped (removed) from the result and a warning generated (configuration “Skip with Warning”). Alternatively, the conversion of the parent document can be aborted with an error (configuration “Strict”).

### 5.2 Archive PDF/A-3

This workflow is engineered specifically for preparing documents for archiving. All features and processing steps of the workflow Archive PDF/A-2 are also available in Archive PDF/A-3. However, the PDF/A-3 format allows additional features regarding embedded files.

#### 5.2.1 Supported File Formats

This workflow supports all file formats listed in Supported File Formats. The conversion of most file formats is enabled by default in the Convert Mode Configuration for Child Documents (Attachments).

#### 5.2.2 Features

In addition to the features of the Archive PDF/A-2 workflow, this workflow also supports the following features:

**Convert to PDF/A-3** The workflow’s profile offers a fine-grained configuration of how files are converted.

**Convert Mode Configuration for Child Documents (Attachments)** This extends the Convert Mode Configuration for Child Documents (Attachments) of the Archive PDF/A-2 workflow with PDF/A-3 features. Specifically the PDF/A-3 standard allows to embed child documents “As Is”, i.e. without converting them to PDF/A.

For example, by default Office files are converted to PDF/A-3, images are used as-is, and executables are removed.

**Collect Mode Configuration** This is the same as the Collect Mode Configuration of the Archive PDF/A-2 workflow. However, in the workflow Archive PDF/A-3 it is common to choose the convert mode “As Is” for some child document types. Therefore the use of the “Merge” collect modes is strongly discouraged, because they are limited to PDF documents (see Merge vs. Attach). Instead, the collect modes “Collection or Single Document” and “Attach” are recommended. Please refer to the documentation panel of the 4-Heights® Conversion Service Configurator for a detailed description of the available collect modes.

**Attach Source Document** The source document (original document) can be attached. The configuration allows for each file type to decide if the source document should be attached or not. By default, the source documents for Office files are attached. Note that this may increase the file size of the result substantially.

**Attach Conversion Report** All events of a conversion can be written to a report file and attached to the result document.

### 5.3 Archive PDF/A-1

This workflow is engineered specifically for preparing documents for archiving, specifically in case the conformances PDF/A-2 or PDF/A-3 are not allowed. The features and processing steps of this workflow are similar the workflow Archive PDF/A-2.

This workflow is disabled by default because we generally recommend to use Archive PDF/A-2 for archiving. Since PDF/A-2 is based on a newer version of the PDF standard, more PDF features are allowed. These features are nowadays commonly used and include transparency, layers, embedded files, and a less restrictive internal file structure.
For these reasons, fewer conversion errors and a better conversion quality can be achieved. More information on PDF/A-2 can be found here:


5.3.1 Supported File Formats

This workflow supports all file formats listed in Supported File Formats. The conversion of most file formats is enabled by default in the profile’s “Convert Mode Configuration for Child Documents (Attachments)”.

5.3.2 Features and Limitations

Compared to the Archive PDF/A-2 workflow, this workflow has the following features and limitations:

Convert to PDF/A-1

Collect Mode Configuration The PDF/A-1 standard does not allow embedded files. Therefore, the only collect mode configuration available is “Merge” (see Merge vs. Attach).

Document Content Rasterization Certain graphical features of PDF documents are not allowed in PDF/A-1, for example transparency. The removal of these features can lead to severe visual differences that may render the page's content unreadable.

The Windows version of the product can preserve the visual appearance of such documents by converting the pages' content to images (content rasterization). When doing so, a “Content Rasterized” warning is generated. If possible, content rasterization preserves the pages' extractable text, links, outlines and viewer preferences.

The Docker version cannot rasterize pages. Therefore, the conversion of such files will lead to a warning “Transparency Removed” or “Visual Differences” depending on whether the visual differences are caused by the removal of transparency.

PDF/A-1 Specific Conversion Warnings The following warnings occur more frequently when converting to PDF/A-1: “Annotation Removed”, “Content Rasterized”, “Layers Removed”, “Transparency Removed”, “Visual Differences”. So the profile's conversion settings for these warnings is particularly important.

Sign Document signing is currently not implemented.

5.4 Conversion

This workflow is engineered specifically for the conversion of documents to PDF. In contrast to the workflows Archive PDF/A-2 and Archive PDF/A-3 files are only converted to PDF (and not PDF/A), the file format is not validated and the output documents can not be signed. OCR is available as an optional processing step.

5.4.1 Supported File Formats

This workflow supports all file formats listed in Supported File Formats.

---

7 Only on Windows, where document content rasterization is available.
8 Only in Docker image, where document content rasterization is not available.
5.4.2 Features

The Conversion workflow offers the following additional features compared to the workflows Archive PDF/A-2 and Archive PDF/A-3:

Optimize for Speed or Size The workflow's profile offers an option to optimize for processing time (speed) or for minimal document file size.

Convert Mode Configuration for Child Documents (Attachments) Certain child documents can be skipped (removed) during conversion to PDF, such as attachments of emails or PDF documents. The convert mode can be specified based on the type of the child document, its filename, or the type of its parent document.

For example, by default executables attached to an email are removed. If desired, rules can be added to attach files that can not be converted (e.g. PDF documents containing unrendered XFA, HTML documents) in their original source format to the resulting output document.

Collect Mode Configuration The collect mode configuration defines how a converted document and its child documents are combined. The collect mode can be configured for each document type and also defines how errors are handled.

For example, emails can be converted by creating a PDF collection (Portfolio) of its body and attachments. Or when converting Word documents, all embedded files can be merged to the converted PDF.

5.5 Invoice

This workflow is engineered specifically for preparing invoices. This workflow converts a main document to PDF/A-3 and attaches one or more files to the result. For example, to a PDF invoice an Excel table containing additional data can be attached.

5.5.1 Supported File Formats

This workflow only supports PDF and MS Word as input format. (See Supported File Formats)

5.5.2 Features

Convert to PDF/A-3 This workflow has been optimized for the conversion of transactional documents. Supported are conversions of Word and PDF documents to PDF/A-3.

Attach Documents In order to attach additional documents to a main document, all must be added to a single job. Which documents to attach and additional properties can be set using the following document options. These values are written into the result document and are visible to a user in the list of attached documents.

DOCROLE (required) Defines the document’s role. Supported values are:
- main (default): The main document. In each job, there must be one main document.
- attached: A document that is attached to the main document. There may be multiple attached documents in a job.

AFRELATIONSHIP (optional) Defines the relationship of the attached document to the main document. Supported values are:
- Unspecified (default): Shall be used when the relationship is not known or cannot be described using one of the other values.
- Source: Shall be used if the attached document is the original source material for the main document.
- **Data**: Shall be used if the attached document represents information used to derive a visual presentation - such as for a table or a graph.
- **Alternative**: Shall be used if the attached document is an alternative representation of content, for example audio.
- **Supplement**: Shall be used if the attached document represents a supplemental representation of the original source or data that may be more easily consumable (e.g., A MathML version of an equation).

**AF.MODDATE** (optional) Defines the date and time when the attached document was last modified. 
Example: "2009-05-19T09:15:22.000000+02:00"

**AF.DESCRIPTION** (optional) Descriptive text associated with the attached document.

**Example:**
Example using the [Shell Client pdfclient](#) to attach `supplement.xls` to the main document `invoice.docx`.Note that the option `-do <name> <value>` applies a document option to the subsequent input file.

```
C:\Temp> pdfclient -v -w Invoice invoice.docx ^
   -do DOC.ROLE attached -do AF.RELATIONSHIP Supplement supplement.xls ^
   out.pdf
Creating job (id job123_4q5mprmmrz0)
Adding file "invoice.docx" (id 0kk1ddv0uzg)
Adding file "supplement.xls" (id 2fg4ufg3qmz)
Output: "invoice.pdf"
- Info: Converted Word document 'invoice.docx' to PDF.
- Info: Converted 'invoice.pdf' to PDF/A-3u.
- Info: Attached file 'supplement.xls' to 'invoice.pdf'.
```

### 5.6 Dossier

This workflow is specifically designed to compile multiple PDF documents into a single dossier.

All PDF document supplied to the job are merged into a single PDF in the order that they were added to the job.

The workflow supports the following optional features:

- Title page with title, subtitle, and a custom background.
- Extensively customizable table of contents.
- Stamps to create e.g. watermarks, header or footer lines.
- Bookmarks to structure the dossier.

#### 5.6.1 Supported File Formats

This workflow only supports PDF files. (See [Supported File Formats](#))

#### 5.6.2 Features

**Title Page**

This workflow can be configured to create a title page for the dossier.

The title page contains the title and subtitle of the dossier. The font, size, placement and alignment of the text is fully configurable.
Additionally, a PDF document can be configured to be used as a custom background of the title page. The title page can be customized for each job using the following options:

**DOSSIER-TITLE**  (optional) The title of the dossier used on the title page.
   This option overrides the value specified in the profile configuration. The option should only be set at job level and is ignored at document level.

**DOSSIER-SUBTITLE**  (optional) The subtitle of the dossier used on the title page.
   This option overrides the value specified in the profile configuration. The option should only be set at job level and is ignored at document level.

**Structuring the Dossier**

The dossier can be structured by creating a table of contents and a document outline (bookmarks).

By default, the title of each document is taken from its metadata. If the title from the metadata is not suitable, it can be overridden by using the following option:

**DOCUMENT-TITLE**  (optional) The title of the document used for the table of contents, document outline, and stamp placeholder.
   This option can be overridden by the more specific options DOCUMENT-OUTLINE-TITLE, DOCUMENT-TOC-TITLE and DOCUMENT-STAMP-TITLE.
   If this option is omitted and none of the more specific options is set, one of the following values is used instead (in order of precedence):
   - The document title from the document metadata.
   - The filename of the input document (without the extension).
   This option should only be set at document level and cannot be configured in the profile configuration.

**Document Outline**

The document outline, also called “bookmarks”, provides a way to quickly jump to a specific part of the dossier.

The "Dossier" workflow uses the existing outline tree of the input documents. Optionally, a parent bookmark can be created for each individual document.

The document outline can be customized for each job using the following options:

**TOC-OUTLINE-TITLE**  (optional) The name of bookmark created for the table of contents (if enabled).
   If this option is omitted, one of the following values is used instead (in order of precedence):
   - The value of the option TOC-TITLE.
   - The title of the table of contents configured in the profile.
   This option should only be set at job level and cannot be configured in the profile configuration.

The document outline can be further configured for each individual input document using the following options:

**ADD-DOCUMENT-OUTLINE**  (optional) Defines whether a bookmark is added to the document outline for the input document. This option can be set at the job level where it applies to all documents, or at the document level where it only applies to the specific document.
   **true** (default) A new bookmark is added to the document outline for the input document. Existing bookmarks of the input document are preserved and added as children to the newly created bookmark.
false  The existing bookmarks of the input document are added directly to the top level of the document outline.

This is useful e.g. for the following use cases:

- Special input documents that don't represent a chapter in the output document, e.g. a cover sheet or table of contents.
- Documents that already contain all necessary bookmarks

**DOCUMENT-OUTLINE-TITLE** (optional) The name of the newly created bookmark (if enabled).

If this option is omitted, one of the following values is used instead (in order of precedence):

- The value of the option DOCUMENT-TITLE.
- The document title from the document metadata.
- The filename of the input document (without the extension).

This option should only be set at document level and cannot be configured in the profile configuration.

### Table of Contents

The automatically generated table of contents provides a great overview over the dossier.

The table of contents supports multiple individually configurable nesting levels by using the existing outline tree of the input documents. Optionally, an parent entry can be generated for each input document.

The table of contents can be customized for each job using the following options:

**TOC-TITLE** (optional) The title of the table of contents.

This option overrides the value specified in the profile configuration. The option should only be set at job level and is ignored at document level.

The table of contents can be further configured for each individual input document using the following options:

**ADD-DOCUMENT-TOC-ITEM** (optional) Defines whether an entry in the table of content is added for the input document. This option can be set at the job level where it applies to all documents, or at the document level where it only applies to the specific document.

true  (default) An entry in the table of contents is added for the input document. Existing bookmarks of the input document are added as sub-entries.

false  The existing bookmarks of the input document are added directly to the top level of the table of contents.

This is useful e.g. for the following use cases:

- Special input documents that don't represent a chapter in the output document, e.g. a cover sheet or table of contents.
- Documents that already contain all necessary bookmarks

**DOCUMENT-TOC-TITLE** (optional) The name of the entry for the document in the table of content (if enabled).

If this option is omitted, one of the following values is used instead (in order of precedence):

- The value of the option DOCUMENT-TITLE.
- The document title from the document metadata.
- The filename of the input document (without the extension).

This option should only be set at document level and cannot be configured in the profile configuration.
**Stamps**

A general overview about stamps can be found in [Stamping](#).

Stamps can be further customized for the “Dossier” workflow by using the following document options:

**DOCUMENT-Stamp-Title** (optional) The title of the document used for stamping, i.e. for the `[input:DOCUMENT-TITLE]` placeholder (if used).

If this option is omitted, one of the following values is used as placeholder value instead (in order of precedence):
- The value of the option DOCUMENT-TITLE.
- The document title from the document metadata.
- The filename of the input document (without the extension).

This option should only be set at document level and cannot be configured in the profile configuration.

**CUSTOM.<OPTION-NAME>** Options whose name starts with “CUSTOM.” define custom placeholder variables that can be used with the placeholder format `[custom:<OPTION-NAME>]`.

Example:
- The stamp text in the profile is configured as: “Reviewed by [custom:REVIEWER].”
- The document option CUSTOM.REVIEWER is set to “John Doe”.
- The resulting stamp text is: “Reviewed by John Doe.”

**Example:**

Example using the [Shell Client pdfclient](#) to create a dossier from a cover page and two chapters. For the second chapter, the name of the bookmark is explicitly overridden.

```bash
C:\Temp>pdfclient -v -w Dossier -do ADD-DOCUMENT-OUTLINE false ^
Creating job (id job1_Sg5Fkycbbhe)
Adding file ".\chapter-1.pdf" (id jgvmhpyqah0)
Adding file ".\chapter-2.pdf" (id 4ogify1xd31)
Output: "out.pdf"
- Info: Assembled PDF files 'chapter-1.pdf' and 'chapter-2.pdf' into dossier 'out.pdf'.
```

**5.7 Processing Steps**

**5.7.1 Office Conversion**

The conversion of Word, Excel and PowerPoint documents to PDF requires an installation of Microsoft Office (see [Compatible Microsoft Office Versions](#)) on the same server where the Conversion Service is installed. The Conversion Service runs and interacts with Microsoft Office Applications with a dedicated user account. We refer to this user account as the Office User. To convert Word, Excel and PowerPoint documents with the 4-Heights® Conversion Service you need to configure the Office User. Please follow the steps described in [Office Configuration](#).

**Processing**

Prior to the PDF conversion, embedded documents from Office documents in Open XML format (e.g. DOCX, XLSX, PPTX) are extracted in a destructuring step. Thereafter the documents are converted to PDF by using Microsoft Word, Excel and PowerPoint. Quality is the reason why we use Office Applications for this processing step, because the layout and resemblance of the resulting PDF and the original document is high. This is especially true, if the document...
was created with the same version of Microsoft Office, which is installed on the server with the Conversion Service. If a document contains macro instructions, these are disabled before opening the document because of security concerns. During conversion CPU and memory usage of Word, Excel and PowerPoint processes are monitored by the Conversion Service.

User parameters (e.g. Ignore Embedded Files, Layout settings such as Fit Page Width, etc) to control the Office to PDF conversion are not yet available. With regard to performance, the parallel execution of Word and Excel Applications is currently limited to four and to one for PowerPoint. Furthermore, the extraction of embedded files from documents in classic Office formats (e.g. DOC, XLS, PPT) is not supported. The extraction of embedded files from password protected Office documents is only supported for Word documents in Open XML format (e.g. DOCX). Finally, parallel operation of the 4-Heights® Conversion Service and the 3-Heights® Document Converter on the same server is not recommended because the Office configurations of the two services interfere with each other, see Office Conversion Problems.

5.7.2 Stamping

Stamping allows adding small bits of content to the existing content of the pages of a PDF.

**Text Stamp**

Text stamps can be used to apply a single line of arbitrary text, e.g. to create header or footer lines or to apply a transparent watermark.

The stamps provide many configuration options such as font, size, color, opacity and positioning.

**Placeholders**

The text of the stamp can be customized using placeholder variables. The supported placeholder names are documented separately in the configurator for each workflow.

The following placeholders are supported by all workflows:

[custom:OPTION-NAME]  

Placeholder variables whose name start with "custom:" are replaced with the value of the corresponding document option CUSTOM.‹OPTION-NAME›.

Example:

- The stamp text in the profile is configured as: “Reviewed by [custom:REVIEWER]”.
- The document option CUSTOM.REVIEWER is set to “John Doe”.
- The resulting stamp text is: “Reviewed by John Doe”.

5.7.3 Metadata

The metadata of the resulting PDF can be customized in all workflows.

**Note:** All metadata is applied to the main result document only and not to embedded files or any other document contained therein. All runtime options to further customize the metadata must be set at job level, not at document level.

**Standard Properties**

The four standard PDF metadata properties that can be customized are “Author”, “Title”, “Subject” and “Keywords”.
A fixed value can be configured in the profile configuration for each of the properties. If no value is configured, the original value is preserved (if available).

The values can also be provided dynamically using the following job options:

**META.AUTHOR**  The author of the document

**META.TITLE**  The title of the document

**META.SUBJECT**  The subject of the document

**META.KEYWORDS**  Keywords that apply to the document

**Extended Properties**

Extended metadata is defined by the XMP standard and the properties are grouped into schemas.

The Conversion Service provides the possibility to set values from the following schemas:

**Dublin Core Schema**

The Dublin Core schema provides a set of commonly used properties.

The values can be configured statically for each profile in the configurator or dynamically using the following job options:

**META.EXT.DC.CONTRIBUTOR**  Contributors to the resource (other than the authors)

**Note:** Although the schema technically supports multiple values, only a single value can be set here.

**META.EXT.DC.COVERAGE**  The extent or scope of the resource.

**META.EXT.DC.IDENTIFIER**  Unique identifier of the resource.

**META.EXT.DC.RIGHTS**  Informal rights statement.

**META.EXT.DC.SOURCE**  Unique identifier of the work from which this resource was derived.

**META.EXT.DC.TYPE**  A document type; for example, “novel”, “poem” or “working paper”.

**Note:** Although the schema technically supports multiple values, only a single value can be set here.

**XMP Basic Schema**

The XMP Basic Schema contains properties that provide basic descriptive information.

The values can be configured statically for each profile in the configurator or dynamically using the following job options:

**META.EXT.XMP.NICKNAME**  A short informal name for the resource.

**META.EXT.XMP.LABEL**  A word or short phrase that identifies a document as a member of a user-defined collection. Used to organize documents in a file browser.

**Note:** This property is not available in PDF/A-1
META.EXT.XMP.RATING A number that indicates a document’s status relative to other documents. Used to organize documents in a file browser. Values are user-defined within an application-defined range.

**Note:** This property is not available in PDF/A-1

**XMP Rights Management Schema**

This schema includes properties related to rights management.

The values can be configured statically for each profile in the configurator or dynamically using the following job options:

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>META.EXT.XMP-RIGHTS.CERTIFICATE</td>
<td>URL of an online rights management certificate.</td>
</tr>
<tr>
<td>META.EXT.XMP-RIGHTS.MARKED</td>
<td>Indicates that this is a rights-managed resource.</td>
</tr>
<tr>
<td>META.EXT.XMP-RIGHTS.OWNER</td>
<td>The legal owner of a resource.</td>
</tr>
</tbody>
</table>

**Note:** Although the schema technically supports multiple values, only a single value can be set here.

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>META.EXT.XMP-RIGHTS.USAGE-TERMS</td>
<td>Text instructions on how a resource can be legally used.</td>
</tr>
<tr>
<td>META.EXT.XMP-RIGHTS.WEB-STATEMENT</td>
<td>The location (URL) of a web page describing the owner and/or rights statement for this resource.</td>
</tr>
</tbody>
</table>

**Custom Extension Schemas**

For metadata properties that are not covered by the predefined schemas, a custom schema can be defined.

The schema definition must be provided statically in the profile configuration.

**Note:** XMP extension schemas are expected to be stable over time. Changes to a schema definition should only add new properties and never change the meaning or type of existing properties. If incompatible changes are necessary, a new schema should be created instead.

The actual property values can be configured statically or provided dynamically using a placeholder and the following custom job option:

**CUSTOM.⟨OPTION-NAME⟩** The value for the placeholder [custom:⟨OPTION-NAME⟩,'Default value'].

- If the default value is omitted ([custom:⟨OPTION-NAME⟩]), the job option is required and an error is signaled if it is missing.
- If the default value is left empty ([custom:⟨OPTION-NAME⟩,' ']), no value is set if the job option is not provided.

The single quotes ('') around the default value are part of the syntax and must not be omitted.

**Advanced**

For advanced users, the metadata can be customized further using the following job options:

**META.XMP** A complete XMP packet that replaces the metadata of the input document.
This option can be used to set all kind of metadata, including Standard Properties, Extended Properties and Custom Extension Schemas.

Any properties set using the standard mechanisms above are applied to this packet and any values set using the standard mechanisms take precedence over the values provided by this XMP packet.

**Note:** When creating a PDF/A document, the XMP metadata must contain a full schema definition for all custom extension schemas. Such a definition can be provided
- directly in the XMP packet.
- using Custom Extension Schemas configuration. In this case the "Force Definition" setting in the schema should be enabled.

### 5.8 Supported File Formats

Each workflow supports a specific subset of the following file formats, which is documented in the description of the workflow itself.

- **Document formats**
  - PDF 1.x, 2.0
  - PDF/A-1, PDF/A-2, PDF/A-3
- **Other formats**
  - Email: EML, MSG (without encryption)
  - Excel: XLS, XLT, XLSX, XLSM, XLTX, XLTM, XML (SpreadsheetML 2003)
  - PowerPoint: PPT, PPS, PPTX, PPTM, PPSX, PPSM
  - OpenOffice⁹: ODT, ODS, ODP
  - CSV
  - HTML, HTM (prepared for archiving¹⁰)
  - TXT
  - XML¹¹
  - ZIP (without password protection)
- **Image formats**
  - JPEG, JPEG2000
  - TIFF
  - BMP
  - GIF
  - JBIG2
  - PNG
  - HEIC
  - HEIF

Where applicable, the 4-Heights® Conversion Service adheres to the following standards:

- ISO 32000-1 (PDF 1.7)
- ISO 32000-2 (PDF 2.0)
- ISO 19005-1 (PDF/A-1)

---

⁹ PDF conversion of OpenDocument Format depends on the rendering in Microsoft Word, Excel or PowerPoint, in particular visual differences may occur with tables and tabs. The visual differences caused by the rendering of shapes are usually not acceptable.

¹⁰ HTML documents need to be self-contained (layout information and images are either inline or available on the web) and suited for portrait page layout. Javascript content is disabled during processing.

¹¹ Layout information and images need to be available on the web.
- ISO 19005-2 (PDF/A-2)
- ISO 19005-3 (PDF/A-3)
6 Connectors / Integration

The 4-Heights® Conversion Service can be integrated into an existing system, by combining suitable input and output connectors.

6.1 Input Connectors

Input connectors can receive or fetch conversion jobs from various sources.

6.1.1 Watched Folder

The "Watched Folder" connector converts all files that are placed into a configurable input directory. After the conversion is finished, the files are deleted.

Note: This connector does not work well with Windows NFS shares. If the watched folder is shared to be accessible by other computers using the NFS protocol, the service will start processing files before they are copied completely. This is due to a limitation in the Windows Server NFS implementation.

Possible solutions:
- Use the SMB protocol instead of NFS
- Use a two-stage process to place the file in the watched folder:
  1. Copy the file to a temporary directory on the NFS share.
  2. From there, move the file to the watched folder.

The following additional configuration options are available:
- Optionally, files in subdirectories are also converted.

6.1.2 REST Input (Plain HTTP)

The "REST input" connector converts all files that are sent to a URL which is customizable by setting the connection id during configuration.

Request

The connector can handle the following two request formats:
- Multiple files can be sent as multipart/form-data body. All files are merged into the same job.
- A single file can be sent as "raw" request body. The query parameter "filename" can be set in the URL to specify the name of the file, otherwise a default name is used.

All URL query parameters are available as variables that can be used to configure certain output connectors, e.g. as placeholders in URL setting of the REST Output connector.
Response

The connector can be configured to provide one of the following three different response types:

- The call returns with a HTTP status "202 Accepted" immediately. The result is handled exclusively by the configured output connectors.
- The call blocks until the job has finished and the first result file is returned as response body.
- The call blocks until the job has finished and all result files are returned as multipart/form-data response.

6.1.3 REST Input (JSON)

REST Input connector using JSON as a structured format for request and response.

Request

This connector allows to specify a job in a structured format (JSON). The files to be converted can either be referenced by URL or included directly in the job description.

```json
{
   "name": "Job name",
   "options": [
      { "name": "<option name>", "value": "<option value>" } ]
},
"variables": {
   "<variable name>": "<variable value>"
},
"data": [
   // First file referenced by URL
   { "url": "http://example.com/path/to/file.jpg" },
   // Second file inline
   { "fileName": "file.jpg", "content": "<base64 encoded file>" },
   // Third file with options
   { "fileName": "file2.jpg", "content": "<base64 encoded file>",
     "options": [ { "name": "DOC.PASSWORD", "value": "mypassword" } ] } ]
}
```

Options and Variables

Options are used to customize the processing of the documents in the workflow itself. The available options depend on the selected workflow.

Variables however are used to pass through information to the output connector. They can be used in certain configuration values of some output connectors in the form of [input: <variable name>]. For example as argument in the Execute Command connector or as placeholders in the URL setting of the REST Output connector.

Response

The connector can be configured to provide one of the following two different response types:

- The call returns with a HTTP status "202 Accepted" immediately. The result is handled exclusively by the configured output connectors.
- The call blocks until the job has finished and all result files are returned as a structured JSON response.
6.1.4 Watched Mailbox (IMAP)

This connector converts all emails that are placed into a configurable mailbox on an IMAP server. After the conversion is finished, the mails are deleted.

The following additional configuration options are available:

- Option to only process the attachments instead of the entire mail.

Variables

The following variables are set by this input connector:

**From**  The sender of the email, including the display name if available.

Examples:
- John Doe <john.doe@example.com>
- john.doe@example.com

**Subject**  The subject of the email.

6.1.5 Watched Mailbox (Exchange Online)

This connector converts all emails that are placed into a configurable mailbox on an Microsoft Exchange Online server. After the conversion is finished, the mails are deleted.

When selecting a mailbox, the connector asks for access to the emails on the specified account. The user must consent to this, otherwise the connector will not work.

The following additional configuration options are available:

- Option to only process the attachments instead of the entire mail.

Variables

The following variables are set by this input connector:

**From**  The sender of the email, including the display name if available.

Examples:
- John Doe <john.doe@example.com>
- john.doe@example.com

**Subject**  The subject of the email.
6.2 Output Connectors

Output connectors feed the output back into an existing system.

6.2.1 Output Folder

The “Output Folder” output connector copies files into a configurable output directory.

The following additional configuration options are available:

- The connector can either store the unchanged input files, or the converted result files.
- Optionally, the folder structure of the input folder can be replicated inside the output folder.
- In case of a filename conflict with an existing file, the file can either be renamed or overwrite the existing file.

6.2.2 Create Text File

This connector creates a text file with additional information in a configurable output directory.

The content of the file is customizable and may contain placeholders:

- [output:ERROR]: The error message.
- [output:WARNINGS]: The warnings that happened during processing. Each warning on a separate line.
- [input:NAME]: The name of the input job.
- [input:<variable name>]: If the configured input connector supports variables, those can also be used as placeholders.

The following additional configuration options are available:

- Optionally, the folder structure of the input folder can be replicated inside the output folder.
- In case of a filename conflict with an existing file, the file can either be renamed or overwrite the existing file.

6.2.3 REST Output

The “REST Output” output connector posts files as a multipart/form-data request to a configurable output URL.

The URL may contain placeholder variables coming from the input connector, e.g:

- http://example.com/path/on/server/?id=[input:id]
- [input:url]
- http://example.com/[input:serverpath]

The following additional configuration options are available:

- The connector can either store the unchanged input files, or the converted result files.

6.2.4 Execute Command

The “Execute Command” output connector executes a configurable command with arguments on files. In the arguments string either placeholder [output:FILES] can be used to pass a list of all files quoted and separated by a white space or [output:FILE_‹i›] to pass the i-th file. If the configured input connector supports variables, those can also be used as [input:<variable name>].

The following additional configuration options are available:

- The connector can either operate on the unchanged input files, or the converted result files.
6.2.5 Output Mailbox (IMAP)

This connector copies files to a configurable mailbox on an IMAP server.

In the case that the files are the original files coming from a "Watched Mailbox" input connector, the original email is copied as is. In all other cases, a new empty email is created with the files as attachment.

6.2.6 Output Mailbox (Exchange Online)

This connector copies files to a configurable mailbox on a Microsoft Exchange Online server.

In the case that the files are the original files coming from a "Watched Mailbox" input connector, the original email is copied as is. In all other cases, a new empty email is created with the files as attachment.

6.2.7 Send Email (SMTP)

This connector sends an email to a configurable recipient.

Email properties such as TO, CC, BCC, subject and body text can be configured freely and may contain variable placeholders to include dynamic data, e.g:

- Information about the processing such as error message, warnings or events.
- Variables set in the request of the REST Input (JSON) connector.
- Query parameters from the REST Input (Plain HTTP) connector.

The email can be configured to contain either the results or the original files as attachments.

6.2.8 Send Email (Exchange Online)

This connector sends an email to a configurable recipient by using a Microsoft Exchange Online account.

Email properties such as TO, CC, BCC, subject and body text can be configured freely and may contain variable placeholders to include dynamic data, e.g:

- Information about the processing such as error message, warnings or events.
- Variables set in the request of the REST Input (JSON) connector.
- Query parameters from the REST Input (Plain HTTP) connector.

The email can be configured to contain either the results or the original files as attachments.
7 Troubleshooting

7.1 Error Codes

If an error occurs during processing in a workflow, an error code with an explanatory message is returned.

The following are common error codes:

**Internal**  The 4-Heights® Conversion Service is not operational because of an internal issue, e.g. an incomplete installation.

A detailed description of the problem is written to the service log file with severity Error. The service administrator should be notified of the problem and submit a support request (see Submitting a Support Request).

Because the issue is not related to the client nor the request, no specific error message is returned. This behavior can be changed in the service configuration. However, since the error is related to the service's configuration and the detailed description is designed to help the administrator resolve the problem quickly, the message might reveal internal configuration settings that one might not want to disclose to clients. Therefore, this is generally only recommended during installation and testing of the 4-Heights® Conversion Service.

**Configuration**  The 4-Heights® Conversion Service is not operational because of a configuration issue, e.g. an invalid or an incompatiable setting.

A detailed description of the problem is written to the service log file with severity Error. The service administrator should resolve the problem with the help of the Configurator.

**Generic**  A generic error occurred.

**UnsupportedFormat**  The format of the input data is not supported.

**UnsupportedFeature**  An unsupported feature was requested. This might be a feature of the input data or one requested using options.

**Option**  An error occurred that is related to job or document options passed by the client. For example:

- A required option is missing.
- An option has an invalid or unsupported value.

**Canceled**  The job has been canceled by the user.

**Timeout**  The job has been canceled because of a processing timeout.

**Password**  The data cannot be processed because of a missing or invalid password. Retry the conversion specifying the missing password using the document option DOC.PASSWORD. This option can be added multiple times to try several passwords for a document or to specify passwords for multiple files, e.g. attachments or embedded files.

**Conformance**  There is a problem with the conformance of the input data. For example:

- The input data's conformance is not supported.
- The input data cannot be converted to meet the conformance required by the workflow and profile.

**Corrupt**  Data cannot be processed because it is corrupt.
7.2 Office Conversion Problems

7.2.1 Configuration Problems

Some of the Warnings and Errors that can be encountered inside the Configurator during the Office Configuration are detailed hereafter and suggestions to fix the issues are explained.

Solution The conversion of Word, Excel and PowerPoint documents to PDF requires an installation of Microsoft Office on the same server where the Conversion Service is installed. Please install a compatible Microsoft Office version (see Compatible Microsoft Office Versions).

Description A parallel operation of the 4-Heights® Conversion Service and the 3-Heights® Document Converter on the same server is not recommended because the Office configurations of the two services interfere and break each other.

Solution This error usually occurs when the domain of the domain user is not recognized. Double check the spelling and please specify the domain by writing the username in SAM or UPN format, e.g. YourDomainYourDomainUser or YourDomainUser@YourDomain.com, respectively.
**Solution**  This error usually signals a connection problem between the server and the domain controller. Double check that the server where the Conversion Service is installed is a member of the domain you are trying to access.

**Solution**  The message *Office configuration problem* can occur, when a setting related to the Office Configuration was changed after the Office User was successfully configured. Try to configure the Office User again, if the problem persists or another error occurs please follow the steps described in [Submitting a Support Request](#).

### 7.2.2 Specific errors related to Word, Excel and PowerPoint documents

**Generic:** *This error can occur if the document is corrupt or a popup blocks the conversion process.*  This error occurs when no progress in the conversion of the Office document was observed for a certain amount of time. If the error was triggered by a popup that is blocking the conversion process, the problem can sometimes be resolved by opening and editing the document manually.

**Generic:** *The application WINWORD exceeded the memory limit of 2048 MB.*  This error indicates that during the conversion of a Microsoft Word document (e.g. DOCX), the application process (Microsoft Word) required more than 2048 MB of memory during conversion. This limit is a precaution set by the Conversion Service to support the overall stability of the conversion process. For further assistance, please follow the steps in [Submitting a Support Request](#) and include the input file.

### 7.3 Service Log

The 4-Heights® Conversion Service writes verbose information to a log file. This allows the administrator to monitor the service’s operation and also troubleshoot problems after an event.
The location of the standard log files can be configured using the Service Configuration. The name of the log files is ConversionService-Service.log. After each day, the old log file is compressed and moved to the file ConversionService-Service-yyyy-mm-dd.log.zip. A maximum of 90 archived log files are kept.

Log Entry Properties

In addition to standard properties, log entries contain the following:

**Level**  The severity of the message. The following severities are common:

- **Fatal**  Severe error. See Error below.
- **Error**  Error that prevents the service from operating, e.g. because of an incomplete installation or invalid configuration. Clients sending processing requests to the service will get an error code Internal or Configuration.

  It is recommended, that the service administrator is notified whenever a message of severity Error or higher occurs. This can be achieved by monitoring the standard service log file or creating a Custom Log.

- **Warn**  Errors that are not critical and do not prevent the service from operating. Even though no immediate action by the service administrator is required, it is advisable to review warnings periodically and decide, if an action is required.

- **Info**  Informational events are useful to monitor the service's operation.

- **Debug**  Debug and tracing messages are strictly for development purposes and analysis by the PDF Tools AG support team. During productive use of the service, messages of this level should be disabled for performance reasons.

- **Trace**  See Debug above.

**Process Name**  The name or type of the process in which the log event occurred.

**Message**  The log message.

**Exception**  Critical log messages often have an exception associated that contains more detailed information on the message's cause.

**Job ID**  Using the job id all messages of a particular processing job can be filtered for analysis.

**Task ID**  This is an internal ID used by the service that is useful for analysis by PDF Tools AG support team.

**Remote IP**  The IP address of the remote host (client). This is meaningful only for messages that are associated with a request at the REST API.

Reading the Log File

The standard log file is in CSV format, for which many different tools are available. Many of them offer a tabular view of CSV files, highlighting by log message severity, and filtering.

Custom Log

For logging the 4-Heights® Conversion Service uses NLog, a very flexible logging platform. This allows to create additional log outputs, e.g. to write to a database or the Windows Event Log.
**Layout Renderers**

The following additional layout renderers are available in the 4-Heights® Conversion Service:

- `${gdc:processName}` The Process Name.
- `${mdlc:jobId}` The Job ID.
- `${mdlc:taskId}` The Task ID.
- `${mdlc:remote-ip}` The Remote IP.

**Examples**

To add a new log output a file NLog.config can be created in the directory bin of the installation directory. Also see the documentation of NLog for more information.

```
Note: A new logging configuration can be tested by disabling the 4-Heights® Conversion Service's license in the license manager. When the service is restarted and a job scheduled, an error is logged.
More information on logging configuration issues can be obtained by activating internal logging:

```xml
<nlog internalLogLevel="Info"
    internalLogFile="c:\path\to\nloglog.txt"
    ...
```

**Example:** A NLog.config that generates a log file C:\path\to\mylog.log.

```xml
<?xml version="1.0" encoding="utf-8"?>
<nlog xmlns="http://www.nlog-project.org/schemas/NLog.xsd"
     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
    <targets>
        <target name="mylog" xsi:type="File" fileName="C:\path\to\mylog.log">
            <layout xsi:type="CsvLayout" delimiter="Comma" withHeader="true">
                <column name="time" layout="${longdate}" />
                <column name="level" layout="${level:upperCase=true}" />
                <column name="processName" layout="${gdc:processName}" />
                <column name="message" layout="${replace-newlines:$message}" />
                <column name="exception" layout="${replace-newlines:${exception:format=ShortType,Message
                   :innerFormat=ShortType,Message:maxInnerExceptionLevel=1}}" />
                <column name="jobId" layout="${mdlc:job-id}" />
                <column name="taskId" layout="${mdlc:task-id}" />
                <column name="remoteIp" layout="${mdlc:remote-ip}" />
            </layout>
        </target>
    </targets>
    <rules>
        <logger name="*" minlevel="Info" writeTo="mylog" />
    </rules>
</nlog>
```
Example: A NLog.config that writes critical messages to the Windows Event Log.

```xml
<?xml version="1.0" encoding="utf-8"?>
<nlog xmlns="http://www.nlog-project.org/schemas/NLog.xsd"
     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <extensions>
    <add assembly="NLog.WindowsEventLog"/>
  </extensions>
  <targets>
    <target name="eventlog" xsi:type="EventLog"
           source="ConversionService"
           layout="${message}\n${gdc:processName}\n${gdc:job-id}\n${gdc:task-id}\n${exception:format=ToString}/">
    </target>
  </targets>
  <rules>
    <logger name="*" minlevel="Error" writeTo="eventlog" /></rules>
</nlog>
```

Example: A NLog.config that sends critical messages by email.

```xml
<?xml version="1.0" encoding="utf-8"?>
<nlog xmlns="http://www.nlog-project.org/schemas/NLog.xsd"
     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <extensions>
    <add assembly="NLog.MailKit"/>
  </extensions>
  <targets>
    <target name="maillog" xsi:type="Mail"
           subject="4-Heights® Conversion Service Error: ${message}" 
           body="${message}\n${gdc:processName}\n${gdc:job-id}\n${gdc:task-id}\n${exception:format=ToString}/">
      to="admin@example.com"
      from="system@example.com"
      smtpUserName="user"
      smtpPassword="password"
      smtpServer="smtp.example.com"
      smtpPort="587"
      smtpAuthentication="Basic"
      secureSocketOption="StartTls"
    </target>
  </targets>
  <rules>
    <logger name="*" minlevel="Error" writeTo="maillog" /></rules>
</nlog>
```
7.4 Submitting a Support Request

For submitting a support request, use the support request form on www.pdf-tools.com. Depending on the type of issue, different data should be provided in addition to the problem description.

General Issues

Use the 4-Heights® Conversion Service Configurator GUI to gather the necessary files for a support request. In the Support Request tab, click on Create Zip and attach the resulting ZIP file to your support request.

The resulting file will contain configuration and log files. However, no sensitive information is included, notably no input files, nor any passwords from the configuration or client options are included.

Issue related to a particular job

Please submit the following data:

- Data described above in chapter General Issues.
- Input files and options used for the conversion.
- The jobId of the conversion.
8 Version History

8.1 Version 3

Changes in Version 3

REST Interface

- **New** query parameter "url" on "addData" POST request to load a file directly from a webserver. [v3.4]
- **New** method "storeJobResultData" to send the result file directly to a webserver via HTTP PUT/POST request. [v3.4]
- **New** property "Status" in result of method "getJobResult" to get Success/Warning/Error status of job result instead of using legacy "Success" property. [v3.9]

Installation and Deployment

- **New** Windows Server 2022 is supported. [v3.3]
- **New** Office 2021 (64 Bit) is supported. [v3.3]
- **New** dependency on Microsoft .NET Desktop Runtime 6.0 (Windows, x64) and Microsoft ASP.NET Core Runtime 6.0 (Windows, x64). [v3.7]

Workflows

- **New** workflow "Archive PDF/A-1". [v3.3]

Workflow "Archive PDF/A-2"

- **New** convert mode configuration for child documents (attachments) for detailed configuration of which documents to convert. [v3.2]
- **New** support to convert HTML, CSV and XML child documents (attachments). Note that this is disabled by default. [v3.2]
- **New** collect mode configuration. This offers a detailed configuration of how to convert documents that have child documents (attachments), specifically how to combine the converted documents and how to handle errors when processing child documents. This supersedes the configuration options "Document Collection Mode" and "Child Error Handling". [v3.2]
- **New** stamping functionality to add single-line text content to the converted documents. [v3.4]

Workflow "Archive PDF/A-3"

- **New** convert mode settings "Skip" and "Skip with Warning", which allows to choose whether an informational event or a warning is generated when skipping (removing) a child document. The new default is "Skip with Warning". When updating an installation, the profiles' behavior is not changed. [v3.2]
- **New** stamping functionality to add single-line text content to the converted documents. [v3.4]
Workflow "Conversion"

- New convert mode settings "Skip" and "Skip with Warning", which allows to choose whether an informational event or a warning is generated when skipping (removing) a child document. The new default is "Skip with Warning". When updating an installation, the profiles' behavior is not changed. [v3.2]
- New stamping functionality to add single-line text content to the converted documents. [v3.4]

Workflow "Dossier"

- New profile setting to control the horizontal alignment (Left, Center, Right) of the title of the table of contents. [v3.2]
- Changed the document outline to contain a bookmark for the table of contents. [v3.2]
- New job option TOC.OUTLINE.TITLE to set the title of the bookmark for the table of content. [v3.2]
- New profile setting for automatic numbering of the entries in the table of contents. [v3.2]
- New profile setting to apply stamps to the pages of the table of contents. [v3.2]
- New optional preprocessing step to convert input documents to PDF. [v3.2]
- New support for custom placeholder variables in Text Stamps. [v3.4]

Common functionality

- New conversion of CSV, HTML (prepared for archiving) and XML documents. [v3.1]
- New conversion of WordprocessingML 2003 (.xml) and SpreadsheetML 2003 (.xml) documents via Microsoft Word and Microsoft Excel, respectively. [v3.3]
- Changed newly created bookmarks to be closed instead of open by default. [v3.3]
- New conversion of signed Emails (s/mime). [v3.4]
- New conversion of OpenDocument Text (.odt), Spreadsheet (.ods) and Presentation (.odp) via Microsoft Word, Excel and PowerPoint, respectively. [v3.4]
- New profile settings and options in all workflows to add metadata to the documents. [v3.5]
- Improved performance especially for small files. [v3.7]
- Improved robustness in low memory conditions on Windows. [v3.8]

Configurator

- Changed export and import of profiles to also include the workflow's activation state. [v3.3]
- Changed name of "Profiles" tab to "Workflows & Profiles". [v3.1]
- New possibility to activate and deactivate workflows in the "Workflows & Profiles" tab. [v3.1]

GUI Client

- Improved UI design to provide a clear overview of the job progress and results. [v3.1]
- New Office Add-Ins for Word, Outlook, Excel and PowerPoint. [v3.3]

Integration/Connections

- New input connector "Watched Mailbox (IMAP)" to automatically convert emails or files sent to a configurable mailbox on an IMAP server. [v3.2]
- New input connector "Watched Mailbox (Exchange Online)" to automatically convert emails or files sent to a configurable mailbox on a Microsoft Exchange Online server. [v3.5]
- New output connector "Output Mailbox (IMAP)" to copy output files to a mailbox without actually sending an email. [v3.4]
New output connector "Output Mailbox (Exchange Online)" to copy output files to a mailbox without actually sending an email. [v3.5]

New output connector "Send Email (SMTP)" to send an email to a configurable or dynamic email address, containing the results or original files [v3.4] and/or additional information (error message, warnings) [v3.6].

New output connector "Send Email (Exchange Online)" to send an email to a configurable or dynamic email address, containing the results or original files [v3.4] and/or additional information (error message, warnings) [v3.6].

New output connector "Create Text File" to create a text file with additional information (error message, warnings) in an output folder. [v3.6]

Connector "REST Input (JSON)"

- New key "options" in JSON format to pass job-level and document-level options to the workflow. [v3.1]
- New key "variables" in JSON format to pass information directly to the output connectors. [v3.1]
- Changed default request body size limit from unlimited to 100MB. [v3.1]

Connector "REST Output"

- New support for placeholder variables in URL setting. [v3.1]

8.2 Version 2

Changes in Version 2

REST Interface

- New support for Cross-Origin Requests (CORS). By default, requests from all origins are allowed. [v2.3]
- Changed value of property type of problem details object (RFC 7807) to http://www.pdf-tools.com/service/rest/errors/‹type›. [v2.4]
- New job result error codes Configuration and Timeout. [v2.6]
- New support for HTTPS service endpoint. [v2.9]

Installation and Deployment

- New Docker image "pdftoolsag/conversion-service" of the 4-Heights® Conversion Service. [v2.2]
- Changed license key format to 4H-V2-XXXXX-XXXXX-XXXXX-XXXXX-XXXXX-XXXXX-XXXXX. License keys with the format 1-XXXXX-XXXXX-XXXXX-XXXXX-XXXXX-XXXXX-XXXXX-XXXXX are no longer supported. [v2.2]
- New dependency on Microsoft .NET Desktop Runtime 5.0 (Windows, x64) and Microsoft ASP.NET Core Runtime 5.0 (Windows, x64). [v2.9]
- New client installer (msi). [v2.10]

Workflows

- New workflow "Archive PDF/A-3" similar to "Archive PDF/A-2" but with support for non-PDF/A attachments. [v2.5]
- New workflow "Invoice" to create PDF/A-3 conformant invoices with attachments. [v2.5]
- **New** workflow "Dossier" for merging PDF documents into a dossier while applying headers, footers or watermarks [v2.8] and support for a title page and table of contents. [v2.11]
- **New** workflow "Conversion" provides document conversions from multiple file formats to PDF with an emphasis on performance and document size. [v2.8]

**Workflow "Archive PDF/A-2"

- **New** profile option "Flatten Annotations" in section "Document Optimization". [v2.3]
- **New** profile option "Child Error Handling". [v2.7]
- **New** profile option for emails "Use Subject As Filename". [v2.8]

**Common functionality

- **New** conversion of HEIC and HEIF images. [v2.11]
- **Improved** email conversion: Attachments of emails where parts of the attachment are removed during the conversion process (e.g. a not convertible part of a ZIP file) are now listed as "Partially removed" inside the email header. [v2.7]
- **New** profile setting "Task Timeout" for all workflows, to avoid the situation where a single long-running processing step can lead to timeouts in other unrelated jobs.
- **Improved** and harmonized messages of warnings and events.

**Office Documents

- **New** conversion of RTF documents. [v2.2]
- **Improved** conversion of Word, Excel and PowerPoint: These types of documents are now converted directly within the 4-Heights® Conversion Service through Microsoft Office Applications. A parallel installation of the 3-Heights® Document Converter is no longer required and no longer supported for this purpose. [v2.1]
- **Improved** conversion of DOCX documents: Certain corrupt documents are now repaired automatically. [v2.6]
- **Improved** conversion of password protected DOCX documents: Embedded files from these documents can now be processed. [v2.6]

**Digital Signatures

- **New** signature types "Windows - PAdES-B-B" and "PKCS#11 - PAdES-B-B" to create basic signatures that require no TSA nor revocation information. [v2.5]
- **Improved** processing of signed PDF documents: If signatures need to be removed, their visual appearance is preserved. [v2.10]

**OCR

- **Improved** OCR recognition task scheduling: Each page is now processed in a separate task, which significantly reduces the amount of time a worker is blocked by OCR tasks and thus improves the latency of other high priority tasks. [v2.4]
- **Improved** OCR engine configuration: Specific configuration for each engine type. [v2.4]
- **New** support for load-balancing to multiple instances of 3-Heights® OCR Service. [v2.6]
PDF to PDF/A Conversion

- **Improved** conversion event messages to be more specific. [v2.11]
- **Improved** detection of optimal target conformance level where the highest level is chosen that can be achieved with a perfect conversion. [v2.11]
- **New** warning types “Annotation Removed” and “Multimedia Removed” for events that were previously reported as “Action Removed”. [v2.11]

Configurator

- **Improved** UI design of the profiles tab to provide a clear overview of profiles and their related workflow. [v2.7]
- **New** tab “Integration” for configuring connections to allow easy integration of the 4-Heights® Conversion Service into existing systems. [v2.4]
- **New** tab “Health & Activity” for monitoring the state and recent activity of the service. [v2.10]
- **New** tab “Statistics” to report and analyze the service’s conversion history. [v2.11]
- **New** possibility to copy profiles and connections. [v2.11]
- **Improved** UI validation of service, profiles and connection configuration to provide greater clarity of errors and warnings. [v2.12]

Integration/Connections

- **New** service “4-Heights® Conversion Service Connections” that allows combining input sources with output for easy integration into existing systems. [v2.4]
- **Removed** service “4-Heights® Conversion Service Folders” (superseded by the new service “4-Heights® Conversion Service Connections”) [v2.4]

Input Connectors

- **New** REST Input connector for plain HTTP requests. [v2.9]
- **New** REST Input connector for structured JSON requests. [v2.10]

Output Connectors

- **New** REST Output connector. [v2.9]
- **New** Command Execute Output Connector. [v2.9]

Various

- **Changed** trademark ™ to registered trademark ®. [v2.10]

8.3 Version 1

Changes in Version 1

- **Changed** minimum required version of 3-Heights® Document Converter Enterprise Edition to version 6.5.
- **Changed** port to connect to 3-Heights® Document Converter Enterprise Edition to 7983.
- **New** requirement of Microsoft Windows Desktop Runtime 3.1 (Windows, x64) and Microsoft ASP.NET Core Runtime 3.1 (Windows, x64). This is in addition to the .NET Framework 4.7.
- Changes to the REST API
  - Changed to return RFC 7807 Problem object instead of proprietary error object.
  - Changed JSON serialization of enum values from integer to string.
  - Changed XML objects to use no namespace in accordance with openapi.yaml.
  - New codes in the service status response.
- New GUI Client for manual processing.
- New view in the 4-Heights® Conversion Service Configurator GUI for submitting a support request.
- New profile option for job priority in workflow "Archive PDF/A2".
- New profile configuration (Digital Signature) available in the configurator GUI.
- New attachment information is shown as part of the email header.
- New possibility to import/export profile configurations.
- New license dependent view in the 4-Heights® Conversion Service Configurator GUI for installing plugins.
- New document content overflow into the margin when converting emails is signaled with ContentOverflowWarning.
- New service setting for proxy configuration.
- New documentation panel in the 4-Heights® Conversion Service Configurator GUI.
9 Licensing, Copyright, and Contact

PDF Tools AG is a world leader in PDF (Portable Document Format) software, delivering reliable PDF products to international customers in all market segments.

PDF Tools AG provides server-based software products designed specifically for developers, integrators, consultants, customizing specialists and IT-departments. Thousands of companies worldwide use our products directly and hundreds of thousands of users benefit from the technology indirectly via a global network of OEM partners. The tools can be easily embedded into application programs and are available for a multitude of operating system platforms.

Licensing and Copyright  The 4-Heights® Conversion Service is copyrighted. This user’s manual is also copyright protected; it may be copied and given away provided that it remains unchanged including the copyright notice.

Contact
PDF Tools AG
Brown-Boveri-Strasse 5
8050 Zürich
Switzerland
http://www.pdf-tools.com
pdfsales@pdf-tools.com