



Aquaforest

Autobahn DX 6.0

Release Notes

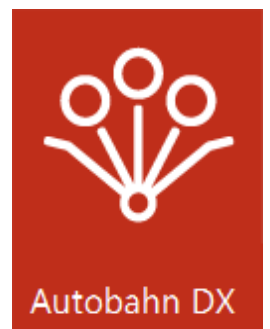


Aquaforest

Autobahn DX 6.0

Release Notes

Autobahn DX **Release Notes**



Version 6.0
February 2025

Content

1	UPGRADING	8
1.1	Upgrade overview	8
1.2	License Keys	8
1	VERSION 6.0.2502.14	9
1.1	New Branding	9
1.2	Improvements	9
1.2.1	Improved Service Monitor tab	9
1.2.2	Any File to Office email attachments do not overwrite existing files	9
1.2.3	Add black and white to image conversion steps	9
1.2.4	Improved Help Tab	9
1	VERSION 6.0.2411.11	10
1.1	New Steps	10
1.2	Improvements	10
1.2.1	Add option to force DPI value when converting image to pdf	10
1.2.2	KVP Expected Keys now gives all data points	10
1.2.3	Improve license naming when first opening ADX and inputting new key	10
1.3	Bug Fixes	11
1.3.1	GdPicture HTML/EML/MSG conversion issue	11
1.3.2	Incorrect default value for GdPicture conversion settings	11
1.3.3	Page Range does not work for Smart Redaction	11
1.3.4	Changing the order of steps does not get saved automatically	11
2	VERSION 6.0.2404.10	11
2.1	Improvements	11
2.1.1	Update documentation on Kingfisher Jobs	11
2.1.2	Designer tab can be slow to load	11
2.1.3	Skip previously OCR'd pages in GdPicture OCR step	12
2.1.4	Start time of each job process is now shown in logging	12
2.1.5	Custom step now provided with the error folder as an extra argument	12
2.1.6	Custom script step now warns users if custom script file field left blank	12
2.1.7	Remove reliance on license.txt	12
2.1.8	Update PDF To Text step to new GdPicture class	12
2.1.9	Add option to replace invalid characters found in barcodes	12
2.1.10	Update terms in the High Availability step	13
2.2	Bug Fixes	13
2.2.1	When "Remove Hidden Text" is set to "True", pages with "visible text" are re-OCR'd	13
2.2.2	Temporary files not deleted when converting email files	13
2.2.3	SharePoint step displaying incorrect license information	13
2.2.4	Stamp PDF step not returning error messages	13
2.2.5	Send Documents step renaming prevents moving original files after processing	13

3	VERSION 6.0.2311.16	15
3.1	Bug Fixes	15
3.1.1	Command line returns before command has completed.	15
3.1.2	Autobahndx.exe Exit Code returns early before files are in output location	15
4	VERSION 6.0.2311.03	15
4.1	New Steps	15
4.1.1	Key Value Pair Extraction	15
4.1.2	Pattern Redaction	15
4.1.3	Pattern Highlight	15
4.1.4	Split PDF (GdPicture)	15
4.1.5	Split By Barcode	15
4.2	Improvements	16
4.2.1	JSON and XML files can now be converted to PDF	16
4.2.2	Updated license key messages to better reflect license status	16
4.2.3	Add SSN and Postal Address support to Smart Redaction	16
4.2.4	Source / Destination folder warning popup for SP and Azure Connectors	16
4.2.5	Users can now control threads used in GdPicture OCR	16
4.2.6	Implement fileProcessOrder for merge steps	16
4.3	Bug Fixes	17
4.3.1	Fields in the designer not being recognized as changed	17
4.3.2	Jobs stuck in 'Finishing job execution' after file error	17
4.3.3	PDF File fails Compression with PDFInvalidContent status	17
4.3.4	Generic Error when running Email steps in ADX	17
4.3.5	SharePoint Download fails to download nested files from SPO	17
4.3.6	OCR Removing visible text during processing	17
5	VERSION 6.0.2304.10	18
5.1	Improvements	18
5.1.1	Improve properties in PDF to Image steps	18
5.1.2	Improve error messages for invalid page ranges	18
5.1.3	Add 'Order by Int' option to Combine PDFs step	18
5.1.4	Improvement for PDF to TIFF (GdPicture) to align with other PDF to Image steps	18
5.1.5	General Improvements and Maintenance to Autobahn code	18
5.2	Bug Fixes	18
5.2.1	Extensions filter in Azure download step not working as expected	18
5.2.2	Azure download step throws error for some special chars	19
5.2.3	Periods removed from file name of Cloud steps	19
5.2.4	Folder created as file when downloaded from Azure	19
5.2.5	Standard OCR DPI property only available when 'Convert To TIFF' set to 'Yes'	19
5.2.6	Create XML Property File does not escape special XML characters	19
6	VERSION 6.0.2302.08	20
6.1	Improvements	20
6.1.1	Add option to Overwrite signed files with same name	20
6.1.2	Add option to create the signed file path folder if it doesn't exist	20
6.1.3	Review the 'Used by another process...' job error's severity	20
7	VERSION 6.0.2301.18	21

7.1 New Steps	21
7.2 Improvements	21
7.2.1 Allow users to disable SSL in email steps	21
7.2.2 Add step to detect Digital Signatures	21
7.2.3 Normalize the OCR Dictionary properties	21
7.2.4 DPI is not a changeable property for GdPicture PDF to Image steps	21
7.2.5 Add a Search bar for Step Properties	21
7.3 Bug Fixes	22
7.3.1 Inconsistency in naming of PDF to Image step property value	22
7.3.2 %JOBSTATUS% does not return job status in email alerts	22
8 VERSION 6.0.2211.18	23
8.1 Upgrading	23
8.1.1 License Key	23
8.1.2 .Net Framework	23
8.2 New Steps	23
8.2.1 GdPicture input file types	24
8.3 Improvements	25
8.3.1 The job upgrader tool could be automatic	25
8.3.2 The BCL Service being turned off gave unhelpful errors	25
8.3.3 Custom script examples are not easy to use	25
8.3.4 Custom script documentation needs updating	25
8.3.5 When first opening ADX with no license key set, the license prompt did not appear at the front	25
8.4 Bug Fixes	26
8.4.1 Warnings sometimes displayed when switching between jobs	26
8.4.2 Non-Latin characters are not rendered correctly in logging	26
8.4.3 Single clicking a job did not properly select the job	26
8.4.4 Job Name label not updating correctly	26
8.4.5 Convert PDF To TIFF step incorrectly giving exit code 0 with specific error	26
8.4.6 SharePoint Downloads with path length over 260 causes error	26
8.4.7 SharePoint steps missing a 'Continue on Error' property	26
8.4.8 Pause function not working as intended	27
8.4.9 SharePoint Upload - first file in batch fails when folder threshold reached.	27
8.4.10 Deleting steps does not update the 'Step Properties'	27
8.4.11 Canceling creating a new job causes strange behavior	27
8.4.12 PDF/A Option in Standard OCR steps not working	27
8.4.13 Output File Name %DIRNAME not working in 'Merge Image to Searchable PDF (Extended)'	27
8.4.14 SharePoint Upload - first file in batch fails when threshold reached.	27
8.4.15 SharePoint - many file entries are created when accessing SPO sites	27
9 VERSION 5.5.2111.16	29
9.1 Enhancements	29
9.1.1 Set the emails sent out from Autobahn back to their original filename	29
9.2 Bug Fixes	29
9.2.1 Files with # in name not uploading to SharePoint correctly	29
9.2.2 Emails over 3MB failing to send	29
9.2.3 Any File to Searchable PDF (Extended) does not produce a CSV log	29
9.2.4 Pages processed in CSV not correct when processing in multicore	29
9.2.5 Console logging during Email steps do not output to the logs	29
9.2.6 DynamicRasterizer and pdfbox version mismatch	29

10	VERSION 5.5.2109.30	31
10.1	Enhancements	31
10.1.1	Add a tool to update Autobahn 5.01 jobdef files	31
10.1.2	Add a new compression step in Autobahn DX	31
10.1.3	Include total page count per job as additional logging	31
10.1.4	Add PowerShell Scripts capability	31
10.1.5	Allow Autobahn email steps to send files in bulk with 'Send Email' step	31
10.2	Bug Fixes	32
10.2.1	PDF page rasterization causes Runtime error	32
10.2.2	PDF files containing both visible and hidden text in a page not processed correctly	32
10.2.3	Files not moved to error folder if not able to copy to output	32
10.2.4	Exception thrown when uploading large document to SharePoint On-Prem	32
10.2.5	The '&' symbol in a SharePoint address causes the job to fail	32
10.2.6	Some PDF steps gave error messages when processing '.PDF' files	32
11	VERSION 5.5.2105.28	33
11.1	Bug Fixes	33
11.1.1	Convert PDF to PDF/A step grayed out with Standard license.	33
11.1.2	%DIRNAME gives the file name and not the directory name for extended jobs	33
12	VERSION 5.5.2104.30	34
12.1	Upgrading	34
12.2	Enhancements	34
12.2.1	Give ability to use templates for metadata settings in "Set PDF Properties" step	35
12.2.2	Added Encrypted PDF Handling options	35
12.2.3	Allow users to sort input files by date	35
12.2.4	Default input delay set to 5 seconds	35
12.2.5	Support for conversion of .eml files	35
12.2.6	Improve license key error messages for Cloud steps	35
12.2.7	Update Email libraries	35
12.2.8	Add Modern Authentication to SharePoint libraries	36
12.2.9	Renamed 'Extract PDF Image via' option	37
12.2.10	Select up to 8 languages in Extended OCR	37
12.2.11	Post Job Completion Alerts	37
12.3	Bug Fixes	37
12.3.1	Subfolders not created in output	37
12.3.2	Barcode zones not being used	37
12.3.3	Merge PDF properties controlling processing encrypted documents	37
12.3.4	Email attachments containing the ';' character in their name causes error	37
13	VERSION 5.01 200316	39
13.1	Bug Fixes	39
13.1.1	Office documents always output as Portrait	39
13.1.2	Cloud OCR option 'Extract PDF Images Via' default value invalid	39
13.1.3	'Keep original image' setting available when processing in Native	39
13.1.4	Binarization Mode property not keeping value	39
13.1.5	Contents of 'temp' folder not being properly removed	39
13.1.6	'Run Continuously' checkbox not keeping value	39
13.1.7	Option "Keep original image" does not save changes and remains blank	39
13.1.8	Image to Searchable PDF (standard) with text output fails to save	40
13.1.9	OCR Any file to PDF not keeping overwrite setting in GUI	40

13.1.10	Job analyzer not deleting temporary files it creates	40
13.1.11	Job API displays inconsistent behavior with concurrent jobs	40
13.1.12	'Include unprocessed PDFs Only' generates a PDFBox error	40
13.1.13	Hangs when processing certain types of PDFs in Native mode	40
13.2	Enhancements	40
13.2.1	SharePoint Step does not support SP2013 OR SP2010	40
13.2.2	Job summary tool uses job temp for its file processing	40
13.2.3	Improved CPU Core throttling	42
13.2.4	Added 'forcecores' to allow setting cores manually	42
13.2.5	Moved 'Save Options' in Module & Options tab	42
13.2.6	Merge Image to Searchable PDF (extended)	42
14	VERSION 5.0.190905	43
14.1	Bug Fixes	43
14.1.1	Removed WIF 3.5 as a Prerequisite	43
14.1.2	Fixed the PDF to Searchable PDF (Cloud OCR) Step properties.	43
15	VERSION 5.0.190805	43
15.1	Bug Fixes	43
15.1.1	SharePoint connector license check failure	43
16	VERSION 5.0.190715	43
16.1	Bug Fixes	43
16.1.1	Any File to PDF fails when generating PDF/A files from text documents	43
16.1.2	Batch Size gets dropped after the first iteration of job from the service	43
17	VERSION 5.0.190605	43
17.1	Bug Fixes	43
17.1.1	CSV Log Files for Paths with Comma(s)	43
17.1.2	Merging PDF Files with Acroforms	43
18	VERSION 5.0.190430	45
18.1	Upgrading from earlier Versions	45
18.1.1	Preserving Existing Job Definitions when Upgrading	45
18.1.2	License Key	45
18.1.3	Removed Files	45
18.2	Enhancements in v5.0	45
18.2.1	Pause Job	45
18.2.2	New Job Steps	45
18.2.2.1	Cloud OCR	45
18.2.2.2	Stamp PDF Files	47
18.2.2.3	Any File to Searchable PDF (Extended)	47
18.2.2.4	Azure Storage Download	47
18.2.2.5	Azure Storage Upload	47
18.2.2.6	Distributed Polling	47
18.2.3	Job API Changes	48
18.2.3.1	Remote API Enhancement	48
18.2.3.2	GetLastRunDate	48
18.2.4	New Alerts Method	48

1.	48
18.2.5 OCR Updates	49
18.2.5.1 Extended Engine	49
18.2.5.1.1 Default Values	49
18.2.5.1.2 New High-Quality OCR engine	49
18.2.5.1.3 Recognition of images scanned with dithering	50
18.2.5.1.4 Automatic language detection of a single-language page	50
18.2.5.1.5 Punch-hole removal	52
18.2.5.1.6 Retain pre-processing settings	52
18.2.5.1.7 Advanced pre-processing settings	52
18.2.5.1.8 New languages available with High-Quality OCR engine	52
18.2.5.1.9 Performance improved for page orientation detection on Korean documents	53
18.2.5.1.10 Memory consumption reduced for document conversion	53
18.2.5.1.11 Turn off PDF/A validation	53
18.2.5.1.12 Default Values	53
18.2.6 Step Types that have changed name	54
18.2.7 Delete Empty Input Folders	54
18.2.8 CPU Core Licensing and Job Control	55
18.2.9 Autobahn DX Directory Changes	55
18.3 Bug Fixes	56
18.3.1 [SDK-120] Graphics state	56
18.4 Known Issues	56
18.4.1 Recognition of accented characters with High-Quality OCR engine (Extended OCR module)	56

1 Upgrading

1.1 Upgrade overview

Two versions of Autobahn cannot be installed at the same time.

Upgrading Autobahn DX will involve:

- Backing up your job definition files
- Uninstalling your current installation.
- Downloading the latest version from our website.
- Installing the new version. This process should only take a few minutes.

When the Autobahn DX UI starts, it checks the version number of each of the step types and will make the required updates to any step. A status bar will show the progress of this process. This only happens the first time the job definition is checked by Autobahn DX. Once completed, your previous job list will appear in the new version as it was previously.

Updating from older versions (before 5.0) may require either manual opening, editing, and saving of the job or recreating the job from a new one.

On the install server, Autobahn requires .Net Framework 4.7.2. The Autobahn installer will check this requirement before installation.

1.2 License Keys

The license key will remain from your previous version. For upgrades between minor releases, the license key should still be valid for the newly installed version. If you are on a new major version (5.0, 5.5, 6.0), then the license key will no longer be valid. In this case, you should contact support for a new license key.

1 Version 6.0.2502.14

1.1 New Branding

Ref: ADX-823

Our company has rebranded and aligned our vision for the future. With these changes, we have changed our naming and branding of products to have a cohesive suite of offerings. That means that Autobahn DX is getting a new coat of paint, transforming into the Nutrient Document Automation Server (DAS).

This current release has minimal changes, only graphical differences, but future versions will be more impactful. We will give guidance for the upgrade to the future version to make the transition to DAS as smooth as possible.

1.2 Improvements

1.2.1 Improved Service Monitor tab

Ref: ADX-839

In previous versions, the image that shows the status of the service will only change when the service is manually turned on/off. Now it will update every time the tab is opened, and also when the new **Refresh** button is pressed. This allows the monitor to reflect any changes to the services outside of DAS.

1.2.2 Any File to Office email attachments do not overwrite existing files

Ref: ADX-846

Anyfile to Office allows users to also process the attachments of email files that are processed by the step. However, we were not allowing overwriting for the output attachments only. We have changed this so that attachments now have the same behavior as other output files, overwriting files already present with the same name.

1.2.3 Add black and white to image conversion steps

Ref: ADX-848

This feature was requested by a customer. They wanted to convert a pdf document to an image file. However they also wanted the output to be bitonal. We have now added a new setting to achieve this.

1.2.4 Improved Help Tab

Ref: ADX-856

Certain functions and links in the Help tab needed updating after the rebranding, and so these changes have been made to clean up this tab. Documentation and support issues now link to our new website.

1 Version 6.0.2411.11

1.1 New Steps

New steps have been added to this Autobahn release, based on customer requests. The list of new steps are as follows:

Convert PDF To Office

Converts and input PDF file to the selected Office type. Options include docx, pptx, xlsx, svg

Convert Any File To Office

Converts many different input file types to the selected Office type. Email files can also have their attachments converted. Output file options include docx, pptx, xlsx, svg

Get Document Information

Creates a report on the searchability of a pdf document. The report includes information about page count, searchable page count, and image page count. Output options include txt, csv, json, xml

Pattern Enumeration

Select a pattern and/or list of terms. A csv file report is created that list the occurrences of each term in the input pdf.

1.2 Improvements

1.2.1 Add option to force DPI value when converting image to pdf

Ref: ADX-774

Some tiff files have incorrect DPI, which causes problems when converting to pdf. The dimensions of the output pdf file will be warped based on the DPI. To counter this, we have added the option to set ConvertTiffForceDPI. This lets the user select a sensible DPI, so the output will have reasonable dimensions.

1.2.2 KVP Expected Keys now gives all data points

Ref: ADX-773

Previously, when retrieving an expected key from the Key Value Pair Extraction step, only the key and value would be retrieved. It has now been improved so that the other data points can also be returned alongside the KVP output.

1.2.3 Improve license naming when first opening ADX and inputting new key

Ref: ADX-760

When opening Autobahn DX for the first time, you must input your license key in the prompt. However, the wording on this prompt was unclear and confusing. It has now been updated to address areas of ambiguity.

1.3 Bug Fixes

1.3.1 GdPicture HTML/EML/MSG conversion issue

Ref: ADX-820 & ADX-818

An update to Google Chrome and Microsoft edge caused an issue with conversions from html, eml, and msg files when using GdPicture steps. GdPicture was updated to fix the issue for Google Chrome, and has dropped support for Edge. A new config setting called **WebBrowserPath** has been added. This can be pointed towards a substitute for Portable Chromium, such as Blink Binaries.

1.3.2 Incorrect default value for GdPicture conversion settings

Ref: ADX-811 & ADX-808

In GdPicture conversion steps, the default value for both **TrackOfficeDocumentRevisions** and **SpreadsheetRenderOnlyPrintArea** are intended to be No. However, it was currently set to Yes by default. This has now been corrected for any future GdPicture conversion steps that are newly created.

1.3.3 Page Range does not work for Smart Redaction

Ref: ADX-793

The Smart Redaction step has a Page Range property that will determine on which pages the redaction will be applied. However, the page range was not being applied properly, and so all pages were being redacted. This has now been fixed, and only the given pages will redact.

1.3.4 Changing the order of steps does not get saved automatically

Ref: ADX-768

If an existing job has multiple steps, and the steps are moved around, this was not recognised as a change in the job. Because of this, leaving the Designer tab would not trigger a prompt to save changes. This has now been fixed and moving steps is seen as a change to the job.

2 Version 6.0.2404.10

2.1 Improvements

2.1.1 Update documentation on Kingfisher Jobs

Ref: ADX-746

When using Kingfisher jobs, a user requires a separate instance of Kingfisher with a job set up in that program for Autobahn to call. It is important that this job is set to single core, so this point has now been emphasized in the Reference Guide for users.

2.1.2 Designer tab can be slow to load

Ref: ADX-738

When switching to to Designer tab, Autobahn has to load the data from the selected job. The time taken to load this data has now been improved to be much faster, making the user experience when switching between tabs better.

2.1.3 Skip previously OCR'd pages in GdPicture OCR step

Ref: ADX-737

Autobahn provided the user with a start time for the entire job, and an end time once all jobs were complete. However, some customers requested more information about when each individual process starts. We have now added this as the **Job Start time** in the logs.

2.1.4 Start time of each job process is now shown in logging

Ref: ADX-736

Currently, the GdPicture OCR step will OCR every page that it processes. In some cases, customers prefer that previous OCR/hidden text layers are skipped instead. We have now added this choice as the options **OCR Searchable Pages** and **OCR Hidden Text**. Setting will prevent text from being overwritten with a new OCR text layer on previously OCR'd pages.

2.1.5 Custom step now provided with the error folder as an extra argument

Ref: ADX-733

When using a custom step, Autobahn provides the input and output file locations to the script file. We have now included the error folder as an extra flag send to the script to give users more options when dealing with files that do not process in the custom script as expected.

2.1.6 Custom script step now warns users if custom script file field left blank

Ref: ADX-732

The custom script step needs the name of the script file that it will run. If this field is left empty, a helpful message box will now appear to warn the user that there needs to be a value for the step to function.

2.1.7 Remove reliance on license.txt

Ref: ADX-729

In previous version of Autobahn, the license.txt was checked to see if jobs needed upgrading. It would also be the version number displayed in the Help tab. This has now been changed to use the version number in the program itself. The license.txt will still be maintained for legacy.

2.1.8 Update PDF To Text step to new GdPicture class

Ref: ADX-728

The underlying GdPicture library is always improving, and a specific class for text extraction was added that fit the requirements for the current PDF To Text step perfectly, while improving on performance and adding new settings and features. This change will help keep the step performing at the best level for future GdPicture upgrades as the library is constantly improving.

2.1.9 Add option to replace invalid characters found in barcodes

Ref: ADX-727

When using the Split By Barcode step, it is very common to use the barcode value in the output filename. This can cause problems in the case that the barcode contains characters that are invalid for filenames. In this situation, we now find and replace these characters with an underscore by default, but this can be customised in the step by changing the "Replace Invalid Characters With" field.

2.1.10 Update terms in the High Availability step

Ref: ADX-388

The high availability step is design to utilize 2 separate instances of Autobahn running on separate hosts. The first host would process the job, and if it ever ran into issues and stopped processing, the second host would take over. The terms referring to the two hosts has been updated to Controller for the processing host, and Replica for the host that monitors and takes over from a Controller that stops processing.

2.2 Bug Fixes

2.2.1 When "Remove Hidden Text" is set to "True", pages with "visible text" are re-OCR'd

Ref: SDK-210

An issue was introduced with the previous SDK update that caused pages to be re-OCR'd when they should have been skipped. Remove hidden text should not have affected visible text, but these page were treated as needing OCR again. This has now been fixed and visible text is no longer re-OCR'd.

2.2.2 Temporary files not deleted when converting email files

Ref: ADX-752

There was an uncommon issue when converting email files to PDF when they contained attachments. A temporary files would be created when processing the attachments, but this file would not be cleaned up after the conversion. It is now handled as expected.

2.2.3 SharePoint step displaying incorrect license information

Ref: ADX-751

The SharePoint Download and Upload steps display the expirtation date of the current Autobahn license. This information was incorrect in previous versions, but will now display the correct date.

2.2.4 Stamp PDF step not returning error messages

Ref: ADX-741

There was an issue with the Stamp PDF step where error messages were not returning if the job failed due to invalid job settings. The validation for the settings has now improved and will return relevant error messages if needed. Also, Page Range will now accept * as a valid input, to bring this in line with other steps.

2.2.5 Send Documents step renaming prevents moving original files after processing

Ref: ADX-740

The Autobahn Send Documents email step has a feature that allows the file to be sent using it's original filename. This renaming caused the file to not move out of the input folder if the option was also set.

This has now been fixed, and you can also use **%EMAILNAME%** in the **Rename Input File** settings to set the input filename to it's original name before moving.

3 Version 6.0.2311.16

3.1 Bug Fixes

3.1.1 Command line returns before command has completed.

Ref: ADX-722

Command line returns before command has completed. This has been fixed in the latest build.

3.1.2 Autobahndx.exe Exit Code returns early before files are in output location

Ref: ADX-720

Sometimes, the job will return an exit code saying that the job has succeeded, but the output has not been copied to the output location. This has now been fixed and exit codes will return after the step is completed.

4 Version 6.0.2311.03

4.1 New Steps

Five new steps have been added to Autobahn DX in this release, which all utilize the GdPicture toolkit's capabilities, including the new Key-Value Pairs processing engine.

4.1.1 Key Value Pair Extraction

Extracts key-value pairs from PDFs alongside information about their location, data type etc. The step outputs the information in the choice of JSON, CSV, and XML output. The user can also specify Expected Keys, and the closest match for each Expected Key will be output in a separate file.

4.1.2 Pattern Redaction

Redacts text in the input PDF based on a regex expression or terms listed in a text file. The user can also customize the color of the redaction zone.

4.1.3 Pattern Highlight

Adds a highlight to text in the input PDF based on a regex expression or terms listed in a text file. The user can also customize the color and transparency of the highlight.

4.1.4 Split PDF (GdPicture)

Split a PDF based on the criteria specified. The available options are to split by page ranges, bookmarks, or split into single pages.

4.1.5 Split By Barcode

Split a PDF based on the barcodes found. The output files can be named based on the value of each barcode found. After specifying the barcode types, page ranges and regular expressions can further specify the barcodes to split the document on.

4.2 Improvements

4.2.1 JSON and XML files can now be converted to PDF

Ref: ADX-673

The 'Convert Any File To PDF (GdPicture)' step previously could not process JSON or XML files as they were not supported. Now, they are treated as text files when converted to PDF, so they will now process successfully.

4.2.2 Updated license key messages to better reflect license status

Ref: ADX-661 and ADX-625

In various areas, generic messages were returned that did not fully reflect the characteristics of the license being used. The messages will now depend on the license, and so will give a clearer and more individual message in these cases.

4.2.3 Add SSN and Postal Address support to Smart Redaction

Ref: ADX-599

Improvements to the GdPicture engine have increased the possibilities of the Smart Redaction step, adding two more redaction categories. You can now choose to redact both Social Security Numbers and Postal Addresses from input PDFs.

4.2.4 Source / Destination folder warning popup for SP and Azure Connectors

Ref: ADX-584

Connector steps in Autobahn are unique in the fact that they ignore either the input or output folder of the job. This is because they instead connect to email, SharePoint, or Azure to send/retrieve their files. When adding an email job, a pop-up would mention that the unused input/output folder would be a dummy folder. To avoid confusion, this warning pop-up message has now been extended to both the SharePoint and the Azure steps.

4.2.5 Users can now control threads used in GdPicture OCR

Ref: ADX-565

The GdPicture OCR step is unique in that it uses multiple concurrent threads to process multiple pages simultaneously, drastically increasing speed. However, this also uses more CPU to run these threads. We have added the option to limit the threads so as not to use all the CPU power. If needed we would recommend between 2-8 threads.

4.2.6 Implement fileProcessOrder for merge steps

Ref: ADX-563

We are adding features to the GdPicture merge steps that will bring it in line with the old merge step. This feature allows some control over the order of the merging of files. Currently, all files are processed alphabetically. This option allows files to be merged in a numerical order instead.

4.3 Bug Fixes

4.3.1 Fields in the designer not being recognized as changed

Ref: ADX-624

There were a few fields in the designer that would not register the job as changed if they were edited. This would mean that the job would not ask you to save changes when changing tabs without saving. This has now been fixed, and the job will correctly recognize that a job has been changed.

4.3.2 Jobs stuck in 'Finishing job execution' after file error

Ref: ADX-591

In very specific circumstances, Autobahn had an intermittent issue when a file had failed during processing. The issue would leave the job in the running state, and so it would be unable to run again until the service was restarted. This issue was investigated and has now been fixed in the latest version.

4.3.3 PDF File fails Compression with PDFInvalidContent status

Ref: ADX-587

This issue was specific to files with a very specific structure. When 'Remove Formfields' and 'Remove JavaScript' is enabled in Compression step, the file fails to process. This was due to leftover data when the formfields were removed, and has now been fixed.

4.3.4 Generic Error when running Email steps in ADX

Ref: ADX-585

Running Autobahn email steps with Modern Authentication was giving a 'Generic Error' for specific setups. The cause of this issue has been resolved and email step should now run as normal.

4.3.5 SharePoint Download fails to download nested files from SPO

Ref: ADX-580

The SharePoint Download step had an issue that prevented it from successfully downloading files stored in subfolders of the target SharePoint location. Only files at the top level would successfully download. This issue has been resolved, and all folders will now download successfully.

4.3.6 OCR Removing visible text during processing

Ref: ADX-578 (SDK-199)

For very specific files, visible text was being removed in addition to hidden text during the OCR process when in native mode, damaging the OCR'd file. This issue was found in the Aquaforest SDK, but a fix was developed and has been deployed to all products using this SDK for OCR processing.

5 Version 6.0.2304.10

5.1 Improvements

5.1.1 Improve properties in PDF to Image steps

Ref: ADX-540

The PDF to TIFF/PNG/JPEG steps contain many properties for user customization. We have added a default value of 0 to the following properties: Brightness, Contrast, Saturation, Gamma, Crop Left, Crop Top. We have also renamed the cropping properties to better reflect their purpose. Crop Height and Crop Width determine the output area of the crop, and so have been renamed to Crop Area Height and Crop Area Width. Crop Left and Crop Top determine the location of the crop area from the top left corner, and so have been renamed to Crop Location Left and Crop Location Top.

5.1.2 Improve error messages for invalid page ranges

Ref: ADX-491

In the PDF to TIFF/PNG/JPEG steps, if a document was not in range for the given page range, the error message would simply state that the file failed to save. This has been improved so that the user will be told that the file does not fall within the specified page range, giving a much clearer reason for failure.

5.1.3 Add 'Order by Int' option to Combine PDFs step

Ref: ADX-490

A common issue that users come across with sorting is that, by default, '*file11*' comes before '*file2*' alphabetically. This can be frustrating when trying to merge files in a numerical order. The new **Combine PDFs** GdPicture step now has the option to sort files by number. This sorts files with the same prefix by a number comparison of their numbers, putting '*file2*' before '*file11*'.

5.1.4 Improvement for PDF to TIFF (GdPicture) to align with other PDF to Image steps

Ref: ADX-451

Our previous implementation of the 'PDF to TIFF (GdPicture)' step did not allow many options for customization to the user. The step has now been rewritten to be more in line with the other PDF to Image steps and now has access to many more properties that users can change depending on their use case.

5.1.5 General Improvements and Maintenance to Autobahn code

Ref: ADX-444

As a part of keeping our products up to date, we review existing code and aim to make improvements when necessary. This change focused on improving the steptype checks at the start of each decision tree, making them slightly faster and easier to update when new steps are added.

5.2 Bug Fixes

5.2.1 Extensions filter in Azure download step not working as expected

Ref: ADX-544

The extensions property filters out any files that do not have the given extension so that they are not downloaded. The list of extensions to be picked up are added in as a string separated by commas. e.g. 'docx,xlsx,pptx'. However, files with extensions that were a part of any of the full extensions in the list

would also be picked up. In the example, **doc xls ppt** files would all be picked up incorrectly. This has now been fixed and only exact matches are picked up.

5.2.2 Azure download step throws error for some special chars

Ref: ADX-539

Azure Blob storage allows for the following special characters to be included in file names: `" * : < > ? \ |`. However, these symbols are illegal to use in the Windows file manager. To allow users to download these files, we replace these symbols with a replacement symbol, which can be set as the new **Replace Invalid Characters With** property. The default value is an underscore.

5.2.3 Periods removed from file name of Cloud steps

Ref: ADX-533

The Cloud steps had an issue where filenames containing multiple periods would have the last period removed along with the subsequent text. For example *'file1.2.3.pdf'* would be output as *'file1.2.pdf'*. The final text was being treated as a file extension and was replaced with *'pdf'*. This issue has now been fixed for all future versions.

5.2.4 Folder created as file when downloaded from Azure

Ref: ADX-500

A new issue came from a new setting in Azure Blob Storage. The 'HNS Enabled' setting was causing folders to be downloaded as if they were files when set to true. This was due to a change in behavior of methods in the code. This has now been fixed and only files will be downloaded, regardless of the blob settings.

5.2.5 Standard OCR DPI property only available when 'Convert To TIFF' set to 'Yes'

Ref: ADX-477

There was an unintended restriction set on the DPI property in the 'PDF to Searchable PDF (Standard)' step. It was only selectable if the 'Convert to TIFF' property was set to 'Yes'. This was not intended as they are independent properties, and so this issue has now been fixed and DPI can always be changed.

5.2.6 Create XML Property File does not escape special XML characters

Ref: ADX-386

There was an issue with the Create XML Property File step where metadata fields were not being escaped. This was an issue if the metadata fields of the input pdf contained any of the following five characters: `< > & ' "`. These would cause the output XML to be invalid. This has now been addressed and the characters are properly escaped before XML creation.

6 Version 6.0.2302.08

6.1 Improvements

6.1.1 Add option to Overwrite signed files with same name

Ref: ADX-448

The new 'Detect Signatures' step copy/moves digitally signed files to a user defined folder to preserve the signatures. This step would fail if a file with the same name already existed in the defined folder. The user now has the option to enable overwrite these files.

6.1.2 Add option to create the signed file path folder if it doesn't exist

Ref: ADX-447

With the new 'Detect Signatures' step, users could copy/move signed files to a specific folder to prevent the signing from being broken. Previously, processing the file would fail if the folder did not exist. There is now the option to create the folder automatically if it does not exist.

6.1.3 Review the 'Used by another process...' job error's severity

Ref: ADX-323

Autobahn has the option to move or delete input files after processing. However, if the input file is forced to stay open by another process, Autobahn will be unable to move the file. Previously in this case, the job would error out and would be unable to restart without manual intervention. A new config option 'ErrorIfDeleteInputFails' has now been added. If set to true, the individual file will fail but the job will continue running for other files.

7 Version 6.0.2301.18

7.1 New Steps

We have continued to improve the variety of steps Autobahn can offer by utilizing the functions available in the GdPicture toolkit. We have added a step that allows for automatic Smart Redaction, and a step that with Detect Signatures on pdf files.

Smart Redaction

The Smart Redaction step has many options for redacting a variety of sensitive data, including credit card numbers, email addresses, and phone numbers. If a section of text matches the criteria, the visible text will be covered and searchable text removed.

Detect Signatures

The Detect Signatures step will analyze pdf files and will act if it detects a signature in a file. It has the option to copy or move the signed file from the processing location to a custom location. If the file was copied, then there is an option to attach the signed copy to the processing file. The signature in the attachment will remain intact even when the base file is processed.

7.2 Improvements

7.2.1 Allow users to disable SSL in email steps

Ref: ADX-429

SSL is a security feature that helps establish secure connections when sending emails. However, sometimes emails are only sent internally, so requiring an SSL certificate is may not be required. Though we recommend keeping SSL enabled in most cases, it can now be disabled by setting the config option "secureoption" to **None**. It is set to **Auto** by default.

7.2.2 Add step to detect Digital Signatures

Ref: ADX-409

PDF documents containing digital signatures are becoming more common. A signed file will lose its signature when processed by jobs in Autobahn. The new Detect Signatures step can be used at the start of a job to filter these files. They can then be copied/moved to a new location before processing.

7.2.3 Normalize the OCR Dictionary properties

Ref: ADX-402

In the previous release of Autobahn, the GdPicture OCR step had two properties that related to the OCR language option. It used a drop-down box for common languages, and a text box to enter any additional languages. The drop-down box was redundant as you could add as many languages as you wanted into the text box below. We decided to remove the drop-down box for clarity and conciseness.

7.2.4 DPI is not a changeable property for GdPicture PDF to Image steps

Ref: ADX-401

The new GdPicture PDF to Image steps (PDF to JPEG and PDF to PNG) did not include an option to set the DPI. This option was set to 300 by default, which is a higher quality that some people need. It is now a changeable property in the step's settings, allowing greater flexibility and control of the output.

7.2.5 Add a Search bar for Step Properties

Ref: ADX-385

Some of the available steps in Autobahn have many properties, particularly the 'Any File to PDF' steps which have settings for multiple file types. We have now added a search bar above the properties which can be used to jump to a specific property without scrolling.

7.3 Bug Fixes

7.3.1 Inconsistency in naming of PDF to Image step property value

Ref: ADX-411

The 'PDF to JPEG' and 'PDF to PNG' steps both had a small visual issue where the **Remove Lines** property had a default value of **No**, but the options only contained **Horizontal**, **Vertical**, and **None**. The values of **None** has now been change to **No** for consistency.

7.3.2 %JOBSTATUS% does not return job status in email alerts

Ref: ADX-408

Autobahn jobs can be set to send out email alerts after jobs are finished. These emails contain information specific to the job such as the name, source, and target. However, the status of the job was not being correctly replaced with the specific job value. Instead, it would always return as 'Status'. This has now been fixed and the emails will give a proper status. Test email alerts now give a 'Testing' status.

8 Version 6.0.2211.18

8.1 Upgrading

8.1.1 License Key

This is a major release of Autobahn and will require a new license key to use. If you have current Support and Maintenance Cover (SMC) for a perpetual license or a current subscription license, please contact support@aquaforest.com to request a new key. From 31 March 2022 Autobahn DX is only available as a subscription product. Existing permanent Autobahn DX 5.5 licenses will remain valid (and function as permanent licenses). Additional SMC for existing permanent licenses will continue to be available for version 6.0.

8.1.2 .Net Framework

This release of Autobahn requires a .Net Framework of 4.7.2 or higher. The Autobahn installer will check this requirement before installation.

8.2 New Steps

The main improvement of this Autobahn release is the inclusion of 11 new steps. These steps all use the GdPicture toolkit, which includes a wide variety of functions and will be constantly improved and updated. A brief description of the new job steps is below:

- **Validate PDFA**
Checks PDFs against a PDFA version and gives an error for the file if it does not conform.
- **Linearize PDF**
Optimizes PDFs for web-viewing, rendering the document one page at a time.
- **Convert Any File To PDF (GdPicture)**
Able to convert a large variety of file types to PDF. This step does not require an Office installation to process Office files.
- **Combine Any File To PDF**
Converts a folder of files into PDF and then merges them, so create a single output PDF. This step uses GdPicture and does not require an Office installation to process Office files.
- **Combine PDFs**
Merges a folder of PDF files to create a single output PDF.
- **PDF To JPEG**
Converts an input PDF to a JPEG file using the GdPicture toolkit.
- **PDF To PNG**
Converts an input PDF to a PNG file using the GdPicture toolkit.
- **PDF To TIFF (GdPicture)**
Converts an input PDF to a TIFF file using the GdPicture toolkit.
- **PDF To Text**
Extracts the searchable text from the pages of a PDF file and creates an output text file.

- **PDF To Searchable PDF (GdPicture)**
Carries out Optical Character Recognition on the input PDF using the GdPicture toolkit, creating an invisible searchable text layer over the document.
- **Create Pdf Portfolio**
Combines a folder of files into an integrated PDF unit. There are a wide range of file types that can be used to create the PDF Portfolio.

8.2.1 GdPicture input file types

The following file types can be used with the Convert Any File to PDF and Combine Any File to PDF steps.

Description	Suffix
Windows bitmap format	BMP
Microsoft Word (.doc) binary file format	DOC
Microsoft Word OpenXML	DOCX
Microsoft Word Macro-Enabled OpenXML format	DOCM
Enhanced Windows Meta-format	EMF
Graphics Interchange Format	GIF
HTML format	HTML
Icon and cursor format (single or multi page)	ICO
Joint Photographic Expert Group	JPEG
Portable Gray-map File	PGM
Portable Network Graphics Format	PNG
Portable Pix-map File	PPM
Microsoft PowerPoint Presentation format	PPTX
Microsoft PowerPoint Macro-Enabled Presentation format	PPTM
Rich Text File Format	RTF
Tagged Image Format	TIFF
Plain text file	TXT
Standard Windows Meta-format	WMF
Microsoft Excel (.xls) binary file format	XLS
Microsoft Excel Spreadsheet format	XLSX
Electronic Mail format	EML
Outlook Item File Format	MSG
Scalable Vector Graphics File	SVG
Device Independent Bitmap format	DIB
24-bit compressed JPEG Graphic format	JPE
MIME HTML format	MHTML
OpenDocument Text file format	ODT
Portable Bitmap Image file format	PBM
Picture Exchange image file format	PCX
TARGA raster graphics format	TGA

8.3 Improvements

8.3.1 The job upgrader tool could be automatic

Ref: ADX-390

A previous issue in Autobahn was upgrading jobs between versions. This was solved with an upgrading tool created for version 5.5. However, the visibility for the tool was low and it was a very manual process. All jobdef files need upgrading if they are not the correct version, so it made sense to make this process automatic. When opening Autobahn, the jobdef folder will now be checked for jobs that need upgrading. The process will be shown in a brief pop-up, but this will be very quick if you only have a few jobs that need to be upgraded.

8.3.2 The BCL Service being turned off gave unhelpful errors

Ref: ADX-353

The Any File To PDF steps that use BCL previously had no check to see if the BCL service was running. This would be an issue if the service was not running, as the job fails, and the error returned would not explain the cause of the issue. The service is now checked to see if it is running, and an informative error message is returned if it is not running.

8.3.3 Custom script examples are not easy to use

Ref: ADX-352

The Custom Script Step allowed users to run their own scripts as part of an Autobahn job. However, the custom folder we provide had example scripts that were old and not easy to use. We recently created new scripts with simpler functions. The examples were added to the custom folder, and there are examples for these new scripts in the Reference Guide.

8.3.4 Custom script documentation needs updating

Ref: ADX-347

The Custom Script section of the Autobahn Reference Guide had become outdated, and the information was both confusing and had links to sites that no longer existed. The documentation has now been updated to provide examples for example scripts with more straightforward functions. Any information that was no longer beneficial has been removed.

8.3.5 When first opening ADX with no license key set, the license prompt did not appear at the front

Ref: ADX-338

When first opening Autobahn, a prompt will appear asking for a license key. However, this prompt did not have priority and would often appear behind other windows. This has now been improved so that the prompt will be brought to the front of all the current windows, so users can more easily see it.

8.4 Bug Fixes

8.4.1 Warnings sometimes displayed when switching between jobs

Ref: ADX-362

If the settings of a job required the job to use work folders, a warning will be displayed to alert the user. This alert was incorrectly popping up when switching to a job that does not require work folders, as the check was being carried out while loading the settings for the new job. This has been fixed and will no longer occur.

8.4.2 Non-Latin characters are not rendered correctly in logging

Ref: ADX-360

Files that contained non-Latin characters were not displayed as expected in the logs, instead appearing as question marks. This was due to the character encoding, and this has now been updated so that most common characters will now correctly display in the logs.

8.4.3 Single clicking a job did not properly select the job

Ref: ADX-354 and ADX-345

In the Autobahn Job Manager tab, single clicking a job would highlight that job, but would not select it. This would cause confusion when then clicking the Design tab, as it would take the user to the previously selected job. This issue also had potential to create blank jobs if creating a new job was canceled before saving it. In this version, the jobs now are selected after a single click, as expected. New jobs that are canceled are also cleared from being selected, so blank jobs are no longer an issue.

8.4.4 Job Name label not updating correctly

Ref: ADX-351

The Job Name label in the top right-hand corner of Autobahn keeps track of the currently selected job. However, in certain situations, the label was not updating correctly. When creating new jobs or renaming old jobs, the label would display the previous job name instead, causing confusion. This has now been fixed and the label updates as expected.

8.4.5 Convert PDF To TIFF step incorrectly giving exit code 0 with specific error

Ref: ADX-342

A file with a broken page was creating a memory issue with the 'Convert PDF To TIFF' step and was not producing an output tiff. However, the step was reporting back that the conversion was a success. This has now been changed so that the output file is checked after processing. If the output file does not exist, an error is reported correctly.

8.4.6 SharePoint Downloads with path length over 260 causes error

Ref: ADX-335

When running the 'SharePoint download' step in Autobahn, if a download would create a file with an output path above 260 chars, the download fails with the message: "The specified path, file name, or both are too long. The fully qualified file name must be less than 260 characters, and the directory name must be less than 248 characters." This error has now been addressed and the restriction should no longer occur.

8.4.7 SharePoint steps missing a 'Continue on Error' property

Ref: ADX-334

For most Autobahn steps, the Continue on Error property exists to allow the job to continue processing files if a single file fails, which is generally desirable behavior. However, this property was missing in the SharePoint steps, meaning that if a single file failed to transfer to/from SharePoint, the job would error. This useful property has now been added to align with other steps.

8.4.8 Pause function not working as intended

Ref: ADX-333

The option to pause a job was intended to allow a job to be stopped temporarily, and it would start from the last file when restarted. However, the pause would instead start the job from the beginning. This was due to a file incorrectly recording the progress. This has now been fixed and restarting a job after a pause will start from the last file processed previously.

8.4.9 SharePoint Upload - first file in batch fails when folder threshold reached.

Ref: ADX-331

This issue is related to ADX-320. When uploading a batch of files to SharePoint Online, if too many folders already exist on SharePoint, the upload operation would be prohibited.

8.4.10 Deleting steps does not update the 'Step Properties'

Ref: ADX-327

Whenever a user would delete a job step in the 'Designer' tab, Autobahn would automatically select the next step (if any), but the step properties of the deleted jobstep would still show. The designer now refreshes properly so that the newly selected job's properties are displayed.

8.4.11 Canceling creating a new job causes strange behavior

Ref: ADX-326

If a user canceled creating a new job, and then immediately selected an existing job, Autobahn would mistake the job ID with the canceled job, loading a job that doesn't exist and causing a UI error. This has now been fixed and the correct job ID will be read, and the job loaded successfully.

8.4.12 PDF/A Option in Standard OCR steps not working

Ref: ADX-324

The PDF/A conversion for Standard OCR steps had an issue and was being skipped during the processing. This has now been fixed, and files should now be output as a PDF/A file if this property is set.

8.4.13 Output File Name %DIRNAME not working in 'Merge Image to Searchable PDF (Extended)'

Ref: ADX-321

Users can use the %DIRNAME template to name the output file from a merge step after the folder. Unfortunately, there was a bug with this specific merge step that replaced the template with the folder name of the parent folder instead. This has now been fixed and %DIRNAME is now replaced by the correct folder name.

8.4.14 SharePoint Upload - first file in batch fails when threshold reached.

Ref: ADX-320

When uploading a batch of files to SharePoint Online, if too many files already exist on SharePoint, the first file in the batch would fail to upload. All other files would upload successfully. This issue has been addressed and all files should successfully upload.

8.4.15 SharePoint - many file entries are created when accessing SPO sites

Ref: ADX-317

Autobahn was creating many certificate key files when accessing SPO sites, found in the *C:\ProgramData\Microsoft\Crypto\RSA\MachineKeys* folder. Some files are critical, but many others could be removed after the accessing is complete. These files are now removed automatically at the end of a SharePoint job step when they are no longer needed.

9 Version 5.5.2111.16

9.1 Enhancements

9.1.1 Set the emails sent out from Autobahn back to their original filename

Ref: ADX-315

Previously all files emailed by the Autobahn 'Send Documents' step were received in the format "name@timestamp@email@.ext". This format is important as it holds information about the email that the file needs to be sent to. However, the file can now be renamed after this information is extracted by setting the new "Use Original Filename" property to true. The received file will have the timestamp, email and any extra @ signs removed.

9.2 Bug Fixes

9.2.1 Files with # in name not uploading to SharePoint correctly

Ref: ADX-319

When using the SharePoint Upload step in Autobahn, any file that contained a # symbol in its name will lose all text after this symbol, including the extension, when it is uploaded to SharePoint. This has been fixed in the latest version and will no longer be an issue with files named this way.

9.2.2 Emails over 3MB failing to send

Ref: ADX-318

Emails that had an attachment size greater than 3MB were getting a general exception error, which would cause the file to not be sent successfully. The issue causing this was fixed and the email attachment limit size now depends more on the email provider. Any file over 50MB will fail in Autobahn with a "file too large" error, but this limit is higher than most providers allow, so please be aware of these limits for your own provider.

9.2.3 Any File to Searchable PDF (Extended) does not produce a CSV log

Ref: ADX-314

It was discovered that for Any File to PDF (Extended) job steps specifically, the CSV log was not being written to when adding the run data. This has been fixed for this version of Autobahn, so it will produce a CSV log like every other OCR step.

9.2.4 Pages processed in CSV not correct when processing in multicore

Ref: ADX-313

The new feature of counting the pages processed for each file and adding this information to the CSV was giving incorrect data when in multicore. This issue was fixed in the latest version, so Autobahn will now correctly output the pages processed in the CSV logs as long as the config option "verboseLogs" is enabled.

9.2.5 Console logging during Email steps do not output to the logs

Ref: ADX-312

The way that the email steps in Autobahn were previously called caused some unintended behavior, including the console output not being saved to the log files. After some reworking, these email steps are now called the same way as every other step. The logging is now saved to the log file as expected, and this has also fixed all the other unintended behavior.

9.2.6 DynamicRasterizer and pdfbox version mismatch

Ref: ADX-311

In earlier versions of the Autobahn release that included PDF Compression, some internal components were mismatched, causing certain jobs to fail. This has been fixed in the current version, and we encourage anyone experiencing these issues to upgrade to the latest version.

10 Version 5.5.2109.30

10.1 Enhancements

10.1.1 Add a tool to update Autobahn 5.01 jobdef files

Ref: ADX-309

When moving jobs from Autobahn 5.01 to 5.5, jobdefs relating to jobs that added new fields or changed field IDs will have missing information. This can cause issues, especially with the new field "Embed Font Subset" that is present in a few common steps.

An 'Upgrade Jobs' tool has been added to the 'Help' tab that will update 5.01 jobs in the 'jobdef' folder when run.

10.1.2 Add a new compression step in Autobahn DX

Ref: ADX-306

We have added the new step 'Modern Compress PDF'. This step has many great options for compression of both image and searchable PDFs. It is a big upgrade on our old compression step, which is now deprecated and will be removed in a later release.

10.1.3 Include total page count per job as additional logging

Ref: ADX-305

A new config option has been added to Autobahn. With 'verboseLogs' set to **true**, the total processed page count will be added to the output log as extra information. There are a few conditions for this to work:

- This option is for single step jobs, as pages may be processed multiple times in multistep jobs
- This information will not work with every steptype, as not every step processes in pages
- The log file path should contain %TIMESTAMP% otherwise data from previous runs may affect the output

10.1.4 Add PowerShell Scripts capability

Ref: ADX-303

Previously, the 'Custom Script Step' was unable to run PowerShell scripts directly. This has been addressed in the latest release, and there should be no issues attempting to run PowerShell script as part of a job in Autobahn.

10.1.5 Allow Autobahn email steps to send files in bulk with 'Send Email' step

Ref: ADX-299

In the previous version, the Autobahn 'Send Email' step would send each file in a separate email. For jobs that included splitting or big volumes, this created a lot of emails. We have now added the option to have multiple attachments in one email. This can be controlled with limits to the total file size or file number attached to the emails that Autobahn sends.

10.2 Bug Fixes

10.2.1 PDF page rasterization causes Runtime error

Ref: SDK-180

In rare cases with specific documents, an exception was being thrown intermittently during rasterization of an Extended OCR step. This has been fixed in the latest Aquaforest SDK version, and incorporated into Autobahn.

10.2.2 PDF files containing both visible and hidden text in a page not processed correctly

Ref: SDK-177

The combination of processing documents that contain both visible and hidden text using settings that set 'RemoveExistingPDFText = true' and 'PdfToImageIncludeText = false' caused the resulting files to be damaged. This rare issue has been fixed completely and should no longer cause any problems in output documents.

10.2.3 Files not moved to error folder if not able to copy to output

Ref: ADX-307

In the rare case that the output folder became unavailable during processing in Autobahn, an error would occur when attempting to copy the completed files over. However, there was no process of sending the original files to the error folder.

Autobahn now attempts to send the contents of the first work folder (or source folder) to the error folder. If this fails, it will disable the deletion of files if set.

10.2.4 Exception thrown when uploading large document to SharePoint On-Prem

Ref: ADX-302, PS-223

An issue was found in the SharePoint Upload step when users used Modern Authentication to upload a very large document, causing the upload to fail. This has been fixed in the latest version and should no longer occur.

10.2.5 The '&' symbol in a SharePoint address causes the job to fail

Ref: ADX-301

SharePoint now allows & symbols to be used in its naming system. However, an error in Autobahn caused jobs to fail when they contained these addresses as it was unable to find the location. This is no longer a problem in the latest version, any location containing a '&' symbol will now be found correctly.

10.2.6 Some PDF steps gave error messages when processing '.PDF' files

Ref: ADX-298

Previously, some Autobahn PDF steps were incorrectly identifying '.PDF' files. The steps would only accept files with a lowercase extension. This has been fixed in the latest version, and the steps are no longer case-sensitive when verifying the file extension.

11 Version 5.5.2105.28

11.1 Bug Fixes

11.1.1 Convert PDF to PDF/A step grayed out with Standard license.

Ref: ADX-297

The Convert PDF To PDF/A step was newly added in Version 5.5. However, it was only available for users with extended licenses. This was not intended, and it is now available for both standard and extended licenses in the latest version.

11.1.2 %DIRNAME gives the file name and not the directory name for extended jobs

Ref: ADX-296

When using the %DIRNAME template for the output file name for extended jobs, the template would be replaced with the filename instead of the directory name in the output file. This has now been fixed in all future versions going forward.

12 Version 5.5.2104.30

12.1 Upgrading

License Key – This is a major release of Autobahn, and will require a new license key to use. Please contact support@aquaforest.com to request a new key.

.Net Framework – This release of Autobahn requires a .Net Framework of 4.7.2 or higher. The Autobahn installer will check this requirement before installation.

Cloud OCR Upgrade – The Cloud OCR steps have been split into Google and Microsoft counterparts. Upgrading jobs with this step will require remaking the job. This affects the following steps:

Image to Searchable PDF (Cloud), PDF to Searchable PDF (Cloud)

Note - Some jobs have been improved and have changed from previous versions. When upgrading, property values may be missing from these steps. We recommend viewing the step in the designer and making sure all the properties are correctly filled out. If you still have issues, deleting and re-adding the step will update the properties to their default values. This affects the following steps:

Read Mailbox, Send Documents, SharePoint Download, SharePoint Upload, Barcode TIFF/PDF

12.2 Enhancements

- Autobahn DX 5.5 contains new OCR engines, including for Extended OCR, which will help improve the accuracy of the output from all OCR steps
- Added “PDF Recognition to JSON” step
 - This step will automatically extract important data from PDF files in the form of Key/Value pairs
 - No need for training or specifying extraction zones
 - Example document types that work well with this step
 - Invoices, shipping documents, etc.
- Added “PDF to PDF/A” step
- Added new improved barcode library with enhanced accuracy when reading certain barcode types e.g. QR codes
- Email and SharePoint libraries updated
 - Now supports OAuth2 Authentication

12.2.1 Give ability to use templates for metadata settings in "Set PDF Properties" step

Ref: ADX-284

Please note that these config options have now been replaced by the new Detect Signatures step. See ADX-468 for more details on this change.

Previously, users could only use static data for metadata input for the "Set PDF Properties" step. Users are now able to use %FILENAME% or %DIRNAME% in the Author, Title, Subject, Creator and Keywords field, and they will be replaced by the input file name or directory name for the output PDF.

12.2.2 Added Encrypted PDF Handling options

Ref: ADX-273

Two config options have been added to Autobahn. The first option, 'securedPdfHandling', allows you to 'pass' through secured files, or you can 'move' or 'copy' these files to another location. The second option, 'securePdfOutputLocation', defines the location to move/copy the secured files to.

12.2.3 Allow users to sort input files by date

Ref: ADX-270

The ability to sort the input file order can now be changed with the new option 'File Order' in the 'Properties' tab of the Designer. There are UTC and local time variants of the date options, totaling nine options: **Alphabetically**, **Created Date (Ascending)**, **Created Date (Descending)**, **Modified Date (Ascending)**, and **Modified Date (Descending)**. Note: this setting does not work for "Merge Image to PDF..." steps, the merge and OCR must be done in two separate job steps.

12.2.4 Default input delay set to 5 seconds

Ref: ADX-269

A common issue that users came across when using Autobahn as part of a workflow is that the input file would not be fully uploaded from the last process when it is picked up by Autobahn. This would cause the incomplete file to error when processed. This extra delay should give the last process time to finish properly uploading the file.

12.2.5 Support for conversion of .eml files

Ref: ADX-253

The latest version of BCL, the third-party library we use for conversions, now supports the conversion of .eml files. We updated the library in Autobahn DX version 5.5 and added the filetype to the configuration, so the .eml filetype should now successfully convert.

12.2.6 Improve license key error messages for Cloud steps

Ref: ADX-242

With the implementation of the new Cloud OCR library, we have also updated the error messages returned to users when there are errors with the Google/Microsoft license key input. This should make it easier for the user to diagnose the problem with the job.

12.2.7 Update Email libraries

Ref: ADX-226

The email library has been updated with a new version that address many of the issues with the previous library. It also includes support for Modern Authentication, which can be used in both email alerts and the individual email job steps.

12.2.8 Add Modern Authentication to SharePoint libraries

Ref: ADX-225

The update to the SharePoint library includes support for Modern Authentication, which has been a feature that many users have requested. The jobs have been updated so users can customize their authentication type to one that suits their needs.

12.2.9 Renamed 'Extract PDF Image via' option

Ref: ADX-219

This small update to the naming of a job option was made to provide more clarity on the way that the PDF Image would be extracted. The option is now 'Convert to TIFF', which will convert a PDF File to a TIFF file before extracting the image. The option is false by default.

12.2.10 Select up to 8 languages in Extended OCR

Ref: ADX-214

The Extended OCR jobs have been updated to allow users to input up to 8 languages, given that they are from the same character set. This change will help users that process multilanguage documents.

12.2.11 Post Job Completion Alerts

Ref: ADX-209

The feature to send an email when a job completes has been in previous versions of Autobahn, but it has now recently been updated with the new email library changes, including the addition of Modern Authentication. On the 'Modules and Options' tab, users are able to select their preferred authentication, and these details will be used for each job that has alerts set in the 'alerts' tab.

12.3 Bug Fixes

- The deployment executables, including TIFF Junction, can no longer be called directly
- The support tool has been removed from the 'Help' tab
- "PDF To Searchable PDF (Standard)" can no longer process TIFF files

12.3.1 Subfolders not created in output

Ref: ADX-294

In an early version of 5.5, Autobahn was not creating any subfolders in the output location. This meant that any folder tree that was processed would be output directly to the source file. This bug has now been fixed for future releases of Autobahn.

12.3.2 Barcode zones not being used

Ref: ADX-275

In the previous barcode step, the zonal information was not correctly sent to the barcode executable, so the whole page was always used as the zone. This has now been fixed with the new barcode step, so declaring zones will ignore the other areas of the page when searching for barcodes.

12.3.3 Merge PDF properties controlling processing encrypted documents

Ref: ADX-274

When processing encrypted files in the Merge PDF step, the job was expected to fail. However, some setting would allow the step to succeed. The step has been improved so it can now process encrypted files. Due to the new encrypted file handling implemented in version 5.5, encrypted files are handled by Autobahn before the merge step.

12.3.4 Email attachments containing the ';' character in their name causes error

Ref: ADX-272

In the rare case that an email attachment's name contained the ';' character, the file would fail to download. This was a limitation of the previous email library, and is no longer an issue with the new email library.

13 Version 5.01 200316

13.1 Bug Fixes

13.1.1 Office documents always output as Portrait

Ref: ADX-265

When processing Office documents, the output would always be portrait, regardless of the input orientation or the "Paper Orientation" setting in Autobahn. This has been fixed so that both orientations can now be produced as output.

13.1.2 Cloud OCR option 'Extract PDF Images Via' default value invalid

Ref: ADX-258

The PDF to Searchable PDF (Cloud OCR) step had an invalid default value of 'No' for this setting. This has now been changed to the valid value of 'Native'.

13.1.3 'Keep original image' setting available when processing in Native

Ref: ADX-257

In the PDF to Searchable PDF (Extended) step, the setting of 'Keep original image' will not take effect if processing with the 'Native' method, but the setting could still be changed. This could be confusing to a user, so this option is now grayed out unless processing using the 'Convert to Tiff' method.

13.1.4 Binarization Mode property not keeping value

Ref: ADX-256

When saving a PDF to Searchable PDF (Extended) step, the Binarization Mode property would store invalid values, due to the property referencing values of another property instead. This has been fixed in the latest version of Autobahn, so the property saves the correct value.

13.1.5 Contents of 'temp' folder not being properly removed

Ref: ADX-255

The main 'temp' folder holds job definitions for quick jobs, but extra definitions were generated every time Autobahn was reopened and these were not removed. This has been fixed so only necessary definitions are created, and these definitions are removed after use.

13.1.6 'Run Continuously' checkbox not keeping value

Ref: ADX-252

In the previous release of Autobahn, the "Run Continuously" checkbox would not keep its value when switching between jobs. This did not affect the schedule of jobs where the value was switched. This UI bug has been fixed in the new release.

13.1.7 Option "Keep original image" does not save changes and remains blank

Ref: ADX-251

When editing an 'Any File to Searchable PDF (extended)' job step, the value for 'Keep original image' will reset its value to false and appear blank in the UI. This has been fixed in the latest version of Autobahn.

13.1.8 Image to Searchable PDF (standard) with text output fails to save

Ref: ADX-246

When processing Image files with the step Image to Searchable PDF (standard), using "OCR to TextFile" and "Output File = Plain Text (no PDF)" the file will be processed, but the file will fail to save. This has now been fixed.

Convert to TIFF returning blank PDF forms

Ref: ADX-243

When converting PDF Forms to TIFF files, output files would lose their data, or return blank. This was an issue in a third-party component, and has been fixed in their latest build, which has been included in this build of Autobahn.

13.1.9 OCR Any file to PDF not keeping overwrite setting in GUI

Ref: ADX-240

The previous release of Autobahn DX 5.0 had a GUI bug that did not retain the overwrite setting in the Any file to PDF OCR job. This has been fixed.

13.1.10 Job analyzer not deleting temporary files it creates

Ref: ADX-238

In the previous release of Autobahn DX 5.0, the Job analyzer was not deleting the files it generated in a temp location. This no longer occurs, and the location of the generated files has been changed (see ADX-239).

13.1.11 Job API displays inconsistent behavior with concurrent jobs

Ref: ADX-233

In the previous release of Autobahn DX 5.0, the Job API encountered a problem if two or more jobs attempted to start at the same time and would instead show one job starting twice. We have fixed this issue.

13.1.12 'Include unprocessed PDFs Only' generates a PDFBox error

Ref: ADX-229

In the previous release of Autobahn DX 5.0, including this filter setting in a job would cause the error to occur. This was due to a PDFBox version mismatch. This issue has been resolved.

13.1.13 Hangs when processing certain types of PDFs in Native mode

Ref: SDK-135

In the previous release of Autobahn DX 5.0, PDFs that contained recursive code were causing hangs when processed in native mode. This was fixed in the latest PDFBox version, which has been implemented in this build.

13.2 Enhancements

13.2.1 SharePoint Step does not support SP2013 OR SP2010

Ref: ADX-261

The SharePoint Upload and Download steps were previously unsupported in SP2013 and SP2010. These steps have now been improved and can function with these versions of SharePoint.

13.2.2 Job summary tool uses job temp for its file processing

Ref: ADX-239

In the previous release of Autobahn DX 5.0, the job summary tool used the User Temp location to store files. These files are now stored in the job temp location.

13.2.3 Improved CPU Core throttling

Ref: ADX-234

Changes have been made in the latest version to improve CPU Core throttling.

13.2.4 Added 'forcecores' to allow setting cores manually

There now exists the config option 'forcecores', allowing users to tell Autobahn how many cores their machine has directly. The user will still be restricted by the cores allowed on their license and will need to restart the service for the config option to take effect.

13.2.5 Moved 'Save Options' in Module & Options tab

Ref: ADX-231

'Save Options' has been moved closer to the settings that it relates to, and the text has been changed to 'Update' instead.

13.2.6 Merge Image to Searchable PDF (extended)

This OCR step has now been updated to support PNG and BMP files.

14 Version 5.0.190905

14.1 Bug Fixes

14.1.1 Removed WIF 3.5 as a Prerequisite

Ref: ADX-227

The previous release of Autobahn DX 5.0 failed during installation on a Windows Server 2019 system, this was because of the WIF 3.5 prerequisite. This has been replaced.

14.1.2 Fixed the PDF to Searchable PDF (Cloud OCR) Step properties.

Ref: ADX-228

In the previous release of Autobahn DX 5.0, the Convert to TIFF option in was not getting passed to the OCR engine when using the new Cloud OCR step.

15 Version 5.0.190805

15.1 Bug Fixes

15.1.1 SharePoint connector license check failure

Ref: ADX-222

The previous release of Autobahn DX 5.0 failed executing the SharePoint steps, this was because of a bug in the license key validation

16 Version 5.0.190715

16.1 Bug Fixes

16.1.1 Any File to PDF fails when generating PDF/A files from text documents

Ref: ADX-218

The previous release of Autobahn DX 5.0 failed when generating PDF/A files from text documents when using **GenericExtension**, **AutoExtension** and **AutoExtensionEx**. Updating to the latest version of **BCL** fixes this issue.

16.1.2 Batch Size gets dropped after the first iteration of job from the service

Ref: ADX-220

The previous release of Autobahn DX 5.0 was setting the job filter limit to zero after the job executes for the first time under the service.

17 Version 5.0.190605

17.1 Bug Fixes

17.1.1 CSV Log Files for Paths with Comma(s)

In the previous version of Autobahn DX, the presence of commas (,) in file paths adds unwanted columns in the CSV log file. We have fixed this issue.

17.1.2 Merging PDF Files with Acroforms

In the previous version of Autobahn DX, there was a bug that caused merging of PDF files with Acroforms to fail. We have fixed this issue.

18 Version 5.0.190430

18.1 Upgrading from earlier Versions

- This release requires version 4.5.2 of the .NET framework. The setup will check whether they are installed on your system and if not, will take you to the appropriate Microsoft site to download and install.
- To upgrade from earlier versions, request a new license key from Aquaforest: sales@aquaforest.com.
- Upgrade blog: <http://www.aquaforest.com/wp/index.php/upgrading-autobahn-dx-server/>

18.1.1 Preserving Existing Job Definitions when Upgrading

When Upgrading to a new version of Autobahn DX, your old jobs will not have all the new step properties added. To rectify this issue, open all your old jobs from the **Job Manager** and save them.

18.1.2 License Key

Autobahn DX 5.0 uses different license keys from the previous versions of Autobahn DX. You will need to request a new license key from Aquaforest: sales@aquaforest.com.

18.1.3 Removed Files

Autobahn DX no longer makes use of the file called 'emimap4.dll', which was used in previous versions. If you have upgraded, this file may still exist in the 'bin' folder and we recommend that it is deleted.

18.2 Enhancements in v5.0

We have made a lot of changes in this version of Autobahn DX; we will discuss these enhancements in this section.

18.2.1 Pause Job

We have now added the ability to resume from Jobs in Autobahn DX if:

1. The Job is Interrupted By a service crash or power failure.
2. If you paused the job from the Autobahn DX GUI.

Note: If you make any changes to the Job when it is in a Paused state the job will start from the beginning.

18.2.2 New Job Steps

In Autobahn DX 5.0, we have added to our long list of job steps. This is to give the user more value and options. For more details, check the section 5.7.2 in the Autobahn DX 5.0 reference guide.

18.2.2.1 Cloud OCR

The optional Cloud OCR module extends Autobahn DX with additional OCR engines from **Microsoft** and **Google**, the main advantages of these OCR engines is their Handwriting recognition capabilities. These OCR engines are available as a SAAS model provided by both vendors. Before you can start using these steps in Autobahn DX, you will need to have a subscription first. See chapter 18 of the reference guide for more details.

We have added two step types to the **Advanced** section of the **Job Designer** tab of Autobahn DX, the steps are named:

- Image to Searchable PDF (Cloud OCR)
- PDF to Searchable PDF (Cloud OCR)

18.2.2.2 Stamp PDF Files

This step can be used to add stamps to PDF pages, we have given the user the ability to customize these stamps extensively in a very simple manner.

Autobahn DX has different ways to apply stamps to a page, this gives the user some level of flexibility.

- **StampTextAsString:** When this operation has selected the text passed as the **StampObject** will be stamped on the PDF document as text.
- **StampPDFText:** When this operation is selected the text passed as the **StampObject** will be stamped on the PDF document as an image.
- **StampPageNumber:** When this operation is selected, every page in the PDF file will be stamped with a page number, starting from the start number. E.g. if **StartNumber** = 6 the first-page number will start from 6.
- **StampPageNumberBates:** When this operation is selected, every page in the PDF file will be stamped with a bates number, starting from the start number. E.g. if **StartNumber** = 6 the first-page number will start from 000006.
- **StampVariable:** This option allows a user to specify a variable like a date, filename or time. The variable specified by the **StampObject** will be stamped on the document. Check the table below for different Stamp variables provided.
- **StampPDFImage:** When this operation is selected the text passed as the **StampObject** is the address of the image to be stamped on the PDF document.

18.2.2.3 Any File to Searchable PDF (Extended)

In previous versions of Autobahn DX, we used to have the **OCR Any File to PDF** (this has changed to **Any File to Searchable PDF (Standard)**) step. This step converted office files to PDF and performed an OCR on image-based files. This step used to be available only for the Standard OCR engine, in version 5.0 we have added a similar step that will use the Extended Engine to OCR image-based files.

18.2.2.4 Azure Storage Download

We added this new step to allow users to download files from an Azure Storage Container to your local machine. This can be used as part of a workflow in Autobahn DX.

18.2.2.5 Azure Storage Upload

We added this new step to allow users to upload files to an Azure Storage Container from your local machine. This can be used as part of a workflow in Autobahn DX.

Using these two steps, you can download files from Azure, process them and upload the outputs back to Azure in a single job.

18.2.2.6 Distributed Polling

This step can be used to implement load balancing in Autobahn DX, it achieves this by copying a fraction of the files from a central input location to the local system where Autobahn DX is running. Multiple Autobahn DX servers can point to one input folder, as a result, the files will be shared across several servers and the processing will be more optimized.

18.2.3 Job API Changes

18.2.3.1 Remote API Enhancement

Previously you had to install Autobahn DX on the client and server machine in order to call a remote API in Autobahn DX. We have changed this so that you will only need to install Autobahn DX on the server computer.

18.2.3.2 GetLastRunDate

We have added the method below to the Job API

Public string GetLastRunDate();

Returns the last Date and Time the job executed.

18.2.4 New Alerts Method

We have changed the way alerts are setup to give the user more control over when to send alerts and what to include in the alerts.

Note: If you are upgrading your jobs from a previous version of Autobahn DX and you have alerts setup for the job, you will have to go to the **Alerts** tab in the **Job Designer** and set up the alerts in your jobs again.

See section 5.2.4 of the reference guide for more details.

Properties	Logging	Schedule	Processing	Alerts
------------	---------	----------	------------	--------

☒ Send Email Alerts on Job Completion

☐ Attach Log File

☐ Attach Job Report

From Email Address

To Email Address

Email Title

%JOBNAME% %JOBSTATUS%!

Test Email

Email Message

Job: '%JOBNAME%' Status: '%JOBSTATUS%'.

Log: %LOGFILE%.
Time: %DATESTAMP% %TIMESTAMP%

Source: %JOBSOURCE%

Target: %JOBTARGET%

Don't Send Email Alerts if

☐ No files were processed

☐ No file errors occurred

☐ Job ran to completion successfully

1.

18.2.5 OCR Updates

18.2.5.1 Extended Engine

Autobahn DX 5.0 now has the latest version of the iDRS engine (iDRS 15.4.2) in the Extended OCR module.

18.2.5.1.1 Default Values

The default values for a few settings have been changed so that it gives good OCR results for different types of documents. These are shown below:

Setting	Changed to
Binarize	true
Binarization Mode	Adaptive
Brightness	128
Smoothing Level	248
Threshold	0
Work Depth	255
Remove Lines	true

18.2.5.1.2 New High-Quality OCR engine

The iDRS™ is updated with I.R.I.S.' brand-new High-Quality OCR: a new OCR engine developed using state of the art concepts from the artificial intelligence research domain.

This new technology brings considerable OCR accuracy improvement especially for bad quality scans, camera images or low-resolution documents, which are affected by common issues such as:

- **Touching characters**

is dressed again! Where have you been, this

- **Broken characters**

'Cold night. Mrs. Corney,' said

- **Distorted characters**

from Barney's hands, who, having delivered another

It will also be suited for recognition of Arabic and Farsi, due to the cursive nature of these languages:

التسهيلات للمستثمر الزراعى وذلك من

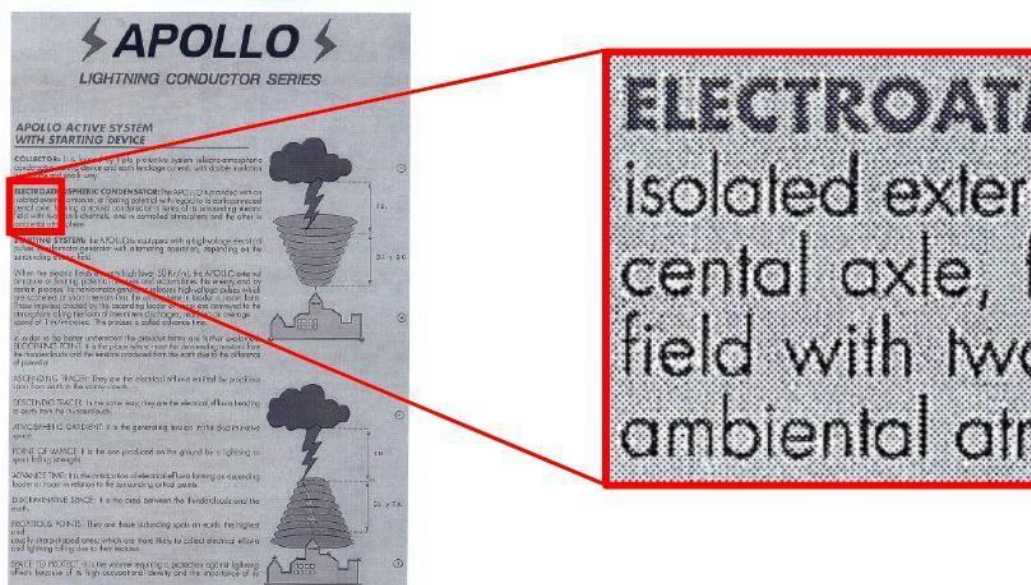
The first release uses High Quality OCR engine for English, Arabic and Farsi languages; further languages will of course be added in future releases.

- For Latin, Cyrillic, Greek, Hebrew and Asian languages, High Quality OCR will be combined with existing OCR engine to use the strengths of both engines.
- For Arabic and Farsi languages, it fully replaces the previous engine, and reaches an unparalleled level of accuracy.

Note that processing time with High Quality OCR engine is expected to increase for low-quality documents: more time will be spent but better accuracy will be reached.

18.2.5.1.3 Recognition of images scanned with dithering

This release exposes an option allowing to improve recognition of color or greyscale images scanned with dithering:



Previous releases would not have properly processed such images: in most cases, the text would simply not have been detected during page analysis step.

How to use

It can be enabled by setting the Undithering property in the Binarization object. Note that you also need to enable smoothing by setting SmoothingLevel to a value greater than '0' to perform undithering.

18.2.5.1.4 Automatic language detection of a single-language page

Extended OCR can now automatically detect the language of an input document.

The aim of this feature is to detect the most probable language of a single-language page.

Supported languages

This release will be able to reliably detect the following scripts/languages:

- **Latin script**

English, German, French, Spanish, Italian, Swedish, Danish, Norwegian, Dutch, Portuguese, Galician, Icelandic, Czech, Hungarian, Polish, Romanian, Slovak, Croatian, Slovenian, Finnish,

Turkish, Estonian, Lithuanian, Latvian, Albanian, Catalan, Irish Gaelic, Scottish Gaelic, Basque, Indonesian, Malay, Swahili, Tagalog, Haitian Creole, Kurdish, Cebuano, Ganda, Kinyarwanda, Malagasy, Maltese, Nyanja, Sotho, Sundanese, Welsh, Javanese, Azeri (Latin), Uzbek, Bosnian (Latin), Afrikaans

- **Cyrillic script**

Serbian, Russian, Byelorussian, Ukrainian, Macedonian, Bulgarian, Kazakh

- **Greek script**

Greek

- **Hebrew script**

Hebrew

Future releases will extend the support to Arabic and Asian scripts.

Note:

- If at least one language has been detected, recognition will be performed in the first language candidate that has been detected, and not in the language(s) set through the **OCR Language x** property.
- If it fails to detect a language, recognition will be performed using the language(s) set through the **OCR Language x** property.

18.2.5.1.5 Punch-hole removal

A new feature has been added to the Extended engine that attempts to remove punch holes from pages. This feature only works when converting images to PDFs or when OCRing PDFs with **Extract Images Method** set to **Convert to TIFF** and with either **Keep Original Image** set to false or **Keep Punch Hole Removal** set to true.

Note: The punch-hole algorithm can be used on images with the following minimum dimensions width: 300px, height: 100px (computed for 300 DPI). The minimum height and width can vary with the image resolution.

18.2.5.1.6 Retain pre-processing settings

You can now retain specific pre-processing in the output PDF documents. For instance, if de-speckling is enabled, speckles are removed from each page to improve the OCR recognition, but this is only done internally and are not reflected in the output PDF document.

In this release, if you want to retain the de-speckling in the output document, set **Keep Despeckled Image** to true. Other pre-processing settings that can be preserved are **deskew**, **dark border removal** and **punch-hole removal**. These can be enabled using **Keep Deskewed Image**, **Keep Dark Border Removal** and **Keep Punch Hole Removal** respectively.

This feature only works when converting images to PDFs or when OCRing PDFs with **Extract Image Method** set to **Convert to TIFF** and with **Keep Original Image** set to false.

18.2.5.1.7 Advanced pre-processing settings

This release has new advanced settings for some existing pre-processing settings of the Extended module. These are:

- AdvancedDeskew
- AdjustmentMode
- ForceDeskew
- AdvancedDespeckle
- Dilate

18.2.5.1.8 New languages available with High-Quality OCR engine

The brand-new technology 'High-Quality OCR' now embeds the 3 following languages:

- Italian
- Spanish
- Portuguese

Note also that variants of already existing High-Quality OCR languages are now supported as well: Afrikaans, Brazilian Portuguese, British, Corsican, Frisian, Luxembourgish, Mexican Spanish, Sardinian, and Swiss-German.

18.2.5.1.9 Performance improved for page orientation detection on Korean documents

The algorithm used for page orientation detection with Korean language has been reviewed, allowing to drastically reduce processing time while improving a bit the accuracy.

On a set of 132 Korean documents, taken in all possible orientations for a total of 528 test cases:

- Older versions:
 - Total time for orientation detection: 5,864 seconds
 - Orientation detection accuracy: 96,0%
- This version:
 - Total time for orientation detection: 971 seconds (divided by a factor 6!)
 - Orientation detection accuracy: 97,3%

18.2.5.1.10 Memory consumption reduced for document conversion

The document output engine includes several optimizations regarding memory consumption when creating an output document. Those changes impact mostly the creation of PDF Image-Text and especially PDF iHQC documents.

In terms of peak memory consumption, considering an input image A4 at 600DPI:

- Older versions:
 - PDF Image-Text: 343 Mb
 - PDF iHQC: 568 Mb
- This version:
 - PDF Image-Text: 238 Mb
 - PDF iHQC: 359 Mb

18.2.5.1.11 Turn off PDF/A validation

In previous versions, PDF/A validation was always performed after converting to PDF/A. However, validating a PDF/A document adds a small performance penalty in terms of the overall processing time. This version allows you to turn off PDF/A validation.

Standard Engine

18.2.5.1.12 Default Values

The default values for a few settings have been changed so that it gives good OCR results for different types of documents. These are shown below:

Setting	Changed to
SavePreDespeckle	true

18.2.6 Step Types that have changed name

For clarity we have changed the names and groupings of our OCR steps in Autobahn DX to represent more clearly what they do. The table below shows the old step names and the corresponding new step.

Old Step Name	New Step Name
Convert TIFF to PDF	Image To Searchable PDF (Standard)
Extended Convert TIFF to PDF	Image To Searchable PDF (Extended)
OCR Image-Only PDF	PDF to Searchable PDF (Standard)
Extended OCR Image PDF	PDF to Searchable PDF (Extended)
OCR Any File to PDF	Any File to Searchable PDF (Standard)
Merge TIFFs to PDF	Merge Image to Searchable PDF (Standard)
Extended Merge TIFF to PDF	Merge Image to Searchable PDF (Extended)

18.2.7 Delete Empty Input Folders

When users select **Delete Input Files** or **Move to Archive after Processing** as the input file post processing action, it is a usual occurrence for a lot of empty folders in the input folder tree to remain. To delete these empty folders, you can use this new setting provided in Autobahn DX 5.0.

Properties Logging Schedule Processing Alerts

Job ID


1001

Job Name

Example : Convert TIFF to Searchable PDF


Source Folder

C:\Aquaforest\Autobahn DX\samples\tiff



Destination Folder

C:\Aquaforest\Autobahn DX\samples\tiff_output



☒ Use Work Folders ☐ Process Sub-Folders ☐ Delete Empty Input Folders

Input Files

Leave input files after Processing

▼

Rename Input Files

%FILENAME%%TIMESTAMP%.%EXT%

Filter Files

Include Files Matching

▼

Filter Expression

*.tif

Batch Size

18.2.8 CPU Core Licensing and Job Control

Your license key will support a specific number of CPU cores. The product will limit the number of concurrent file processing operations to this number and will “throttle” jobs accordingly.

For example, if a 4-core licensed server is currently running a 2-core job and a new job starts that is configured for 4 cores the number of cores allocated to the second job will be reduced accordingly:

Autobahn DX using 2 cores out of 4 allowed.

We will reduce the number of cores in this job from 4 to 2 allowed.

As another example, if a 4-core licensed server is currently running a 4-core job and a new job starts that is configured for 2 cores then the second job will not be able to start until cores are freed up:

Autobahn DX using 4 processors out of 4 allowed.

We will attempt to start the job 18 time(s) over the next 180 seconds.

The retry interval and number of tries is determined by these two config file settings in Autobahn.config (by default this file is in C:\Aquaforest\Autobahn DX\config)

```
<add key="jobqueuetimeout" value="180" />
```

```
<add key="jobqueueinterval" value="10"/>
```

18.2.9 Autobahn DX Directory Changes

We have added a distribution directory to the installation directory of Autobahn DX, this directory will contain the components need for Autobahn DX to function. As a result, we have moved some folders from the top-level folder to the distribution folder, we have also created new folders for other components. The table below shows the details.

Application	Old Directory Path	New Directory Path
Extended OCR	extendedocr	distribution /extendedocr
TIFF Junction	tj	distribution /tj
PDF Junction	pj	distribution /pj
Cloud OCR (new)	-	distribution /cloudocr
SharePoint Connector (new)	-	distribution /sharepoint
Azure Connector (new)	-	distribution /azure
Support Tool	support	distribution /support

18.3 Bug Fixes

18.3.1 [SDK-120] Graphics state

The graphics state was not being restored when processing pages that require rotation in the Standard OCR engine. This caused issues when other applications manipulated the PDF after it had been OCR'd by Aquaforest. This has now been fixed.

18.4 Known Issues

18.4.1 Recognition of accented characters with High-Quality OCR engine (Extended OCR module)

The new Extended OCR module currently has an issue that impacts Latin languages processed with HighQuality OCR engine.

When a character with an accent (like é, è, à, ñ, etc.) is recognized but is not present in the character set (for instance if recognition is performed in English), the OCR engine will output a reject character (U+FFFD).

This is a regression compared to previous versions, where the "base" character would be output instead (e.g. 'e' instead of 'é').

This issue will be fixed with the next release.