



Autobahn DX 3.03 Release Notes

14th November 2012

© Copyright 2012 Aquaforest Limited

<http://www.aquaforest.com/>

1 UPGRADING FROM EARLIER VERSIONS

- This release requires both version 3.5 and version 4 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.

To upgrade from earlier versions, uninstall Autobahn DX and install Autobahn DX 3.03.

1.1 Preserving Existing Job Definitions when Upgrading

Existing Version 2.x and 3.0 job definition and status files are compatible with Version 3.03 and are not deleted on uninstallation. It is recommended that Version 2.x job definitions are resaved in Version 3.0 before running.

The job definition files are located in the jobdef and jobstatus folders in the product installation folder. If the new release is installed in a different location to Version 2.x (which by default was installed under C:\Program Files\Aquaforest\Autobahn DX 2.x) these folders may be copied to the new location.

1.2 License Key

A 3.0 license key is required for production use of Version 3.0. Earlier version 2.x keys will not be recognized.

If you have any questions about upgrading to version 3.0 please contact Aquaforest support : support@aquaforest.com

2 OUTLOOK .MSG FILES AND ATTACHMENTS (3.01 ONWARDS)

Direct support for converting Outlook .MSG files has been added. In addition an option is available to convert the email attachments so that the resulting PDF is composed of the email plus each page of each attachment.

Outlook 2007 is sufficient to allow this conversion on 32-bit systems but on 64-bit systems Outlook 2010 is required to convert .msg files.

3 OPEN OFFICE (3.01 ONWARDS)

Open Office formats are now supported. This requires presence of Open Office on the machine where Autobahn DX is installed.

Furthermore, Autobahn DX can now be configured to make use of Open Office to convert Microsoft Office documents rather than require Microsoft Office to be installed.

4 PDF/X SUPPORT (3.01 ONWARDS)

Support has been added for PDF/X-3 and PDF/X-1a for the "Convert any file to PDF" job step.

5 INCLUDE MARKUP IN PDF OUTPUT (3.01 ONWARDS)

A new option is available for the "Convert Any File to PDF" job step. The "Include Document Markups" option allows Word comments and markup to be included in the generated PDF.

6 LOGGED ON USER REQUIREMENT (3.01 ONWARDS)

The requirement for a user to be logged on to allow “Convert any file to PDF” conversion has been eliminated in this release.

7 STAND ALONE IMAGE PDF COMPRESSION (3.01 ONWARDS)

A new step is available that will perform MRC and JBIG2 compression on existing image PDF files without requiring OCR. Note that this is specifically intended for use with image-only PDF files.

8 IMAGE RETENTION IN GENERATED RTF FILES (3.01 ONWARDS)

When analyzing document pages prior to OCR, areas of the document may be identified as being “graphic areas”. When “Treat Image Areas as Text” is set then these graphic areas are treated as text and OCR’ed. From this release, when “Treat Image Areas as Text” is not set, graphic areas will be included in generated RTFs. Note that in this release all the graphics areas are treated as black and white and color images are converted where required.

9 IMPROVEMENTS TO AUTOROTATION (3.01 ONWARDS)

Dictionary lookup has been added to the checks to determine “good” results for rotated pages as well as flip detect for the initial test on whether a page is rotated. These two measures have greatly improved the detection of rotated pages and the determination of the correct rotation to use.

10 OPTIMISED OCR (3.01 ONWARDS)

A new option is available that can be used to run each page OCR with 2 or more different settings and then chose the best set of results.

When the new `-A /optimiseocr` advanced flag is used the OCR and image processing engines will use the settings in the `ImagePreProcessingDefaults` section of the file `Properties.xml` modified by any properties set on the OCR and PreProcessing objects. See section 12 of the reference guide for detailed information.

11 LONG FILE NAME HANDLING (3.01 ONWARDS)

Windows has a file name length of 260 characters and a folder name length limit of 248 characters. There are potentially two areas where this could impact Autobahn DX processing. Autobahn DX 3.0 will automatically detect at the start of a job whether there will be any file that are affected and detailed messages will be provided to enable resolution. Section 2.10 of the reference guide discusses the issues involved.

12 HANDLING OF CORRUPT JOB STATUS OR DEFINITION FILES (3.01 ONWARDS)

Under certain circumstances (such as power failure) it is possible Autobahn DX job status or definition files could become corrupted. Autobahn DX 3.0 has enhanced handling for such cases and will provide improved messaging and where possible will automatically correct any corruption.

13 CORRECTIONS IN 3.01

13.1 Stop Job

Under certain circumstances some sub-processes may have been left running after using “Stop Job” . This has been corrected so that all sub-processes are now correctly terminated.

13.2 API - WaitForExit()

Under certain circumstances such as extreme load the WaitForExit() API call could not return even though the job had completed successfully. This has been corrected.

13.3 Caching of Job Definitions

In certain specific scenarios after making changes, Job Definitions may be cached incorrectly, resulting in an older version being run until the Autobahn DX service was restarted. This has been corrected.

13.4 Run Tab / Job Output

Improvements have been made in the speed and consistency of the log output.

13.5 Process Timeout / Hanging

Under specific extreme load situations it was possible for an OCR process to timeout without restarting, thus causing a job to hang. This has been corrected.

13.6 Job Step Sequencing

Autobahn DX 2.50 incorrectly allowed a “Convert any file to PDF” job step to follow a job step that produced only PDF output files. This has been corrected.

14 EXTENDED OCR MODULE

The new optional Extended OCR module extends Autobahn DX 3.0 with an additional OCR engine (see section 15 of the reference guide for complete details)

The additional engine has the following benefits over and above the standard Aquaforest OCR engine

- IRIS OCR Engine providing enhanced recognition
- DOCX Output option
- Improved RTF Output
- CSV and SpreadsheetML output options
- Additional Western Language Support
- Optional Asian Language Support
- Optional iHQC Compression (see section 16 of the reference guide)
- Support for multiple languages within a single document from the same alphabet - eg French + German + Italian
- Certain advanced pre-processing options such as black border removal

15 MULTI-CORE MODULE

Most modern computers have multiple CPU cores and yet most software is not designed to fully take full advantage of the available CPU power.

Previous versions of Autobahn DX have been able to exploit multiple CPU cores by running multiple concurrent jobs or by using the "threads" option which could split large documents into 2 or 4 segments and process each segment in parallel.

The new Multi-Core Module allows a single job to process multiple files in parallel and so is ideally suited for processing large volumes of documents of any size.

Job steps that have a "Number of Cores" setting can be set to use a specific number of cores. See section 17 of the reference guide for further details.

16 TRIAL LICENSING

To evaluate the product, select the product level that you wish to trial from the list below and enter the appropriate trial license key under the "Modules and Licensing" tab.

Trial License Key	Functionality
SERVERTRIAL	This is the default and enables all functionality except for that provided by the optional modules.
WORKSTATIONTRIAL	This sets the product in Workstation mode and will disable all tabs except for "Quick Job". It will also disable command line use.
MULTICORETRIAL	Enables the Multi-Core module.
EXTENDEDOCRTRIAL	This allows the Extended OCR Optional Module to be trialled.
EXTENDEDOCRHIQTRIAL	This enables use of the highly compressed PDF output option.
EXTENDEDOCRASIANTRIAL	This allows the Extended OCR Optional Module to be trialled. Including Asian languages (Japanese, Korean, Simplified and Traditional Chinese).
EXTENDEDOCRASIANHIQTRIAL	This enables use of the highly compressed PDF output option.

17.1 Skip Long File Names

This relates to the issue described in section 11. In the previous release of Autobahn, if a file with a file name which exceeds the file length limit was encountered, the job execution was terminated until the file in question was removed or renamed. In Autobahn 3.03, users have the option to skip those files without hindering job execution. This can be achieved by selecting the “Skip Long File Names” checkbox found on the Designer page.

17.2 Merge Files Flag

This feature is used to give the user a certain level of control over which folders to be included in the merge operations (“Merge Tiffs”, “Merge PDF”, “Merge Tiffs to PDF” and “Extended Merge Tiffs to PDF”). It provides the user with a mechanism of flagging folders that are to be included for processing.

To use this feature, the following steps need to be followed:

- Choose a name you want to set as the flag name. For instance, “_ready”.
- Open Autobahn.config file located at “<InstallDirectory>\Autobahn DX 3.03\config\” and change the value of <add key=“mergeflag” value=“” /> to <add key=“mergeflag” value=“ready.adx.merge” />. *Note: The file extension must be “*.adx.merge” for this to work.*
- Next create the 0 size byte file named “_ready.adx.merge” and copy and paste in each folder that requires processing.
- Restart the service to load the new configuration. *Note: restarting the service will cause all jobs that are running through the job manager to terminate.*

17.3 OCR Process Hanging

On certain machines and under certain very specific extreme load situations when running several jobs simultaneously using multicore processing, it was possible for processes to hang. This has been fixed. Two configurations relating to this fix can be controlled through the Properties.xml file; further information can be obtained from Aquaforest Support. The options are:

- <PipeClientConnectionTimeout> - sets the maximum amount of time the Pipe Server should wait for a client to connect to it
- <CreateProcess> - when set to True, it uses Platform Invocation Services (P/Invoke) to create the process

17.4 Stop In-place Processing File Deletion

In Autobahn 3.02, it was possible to setup an in-place processing job to use “Delete Input Files” in such a way that both the input and output files were deleted. Validation ensures that this can no longer happen.

17.5 Continue On Error in Optimize PDF

Continue on Error option has been added to the Optimize PDF step properties, which was missing in previous versions of Autobahn DX.

17.6 JBIG2 compression

There were certain cases where JBIG2 compression was not working on Windows Server 2008. This has been corrected.

17.7 Convert any files to PDF

The requirement for a user to be logged on to allow file conversion was eliminated in the previous release for Microsoft Word Documents only. This release eliminates this requirement for XLS/XLSX, PUB, VSD and PPT/PPTX documents.

Extra configuration options in the <Autobahn Install Folder>\pj\bin\topdf.exe.config file has also been provided. Some of the new options still convert the same file types but they perform these actions in a different way there by providing a more flexible approach to file conversion.

Below is a table showing the new options:

Extension	Use
<code><add key="AutoExtensionEx" value=" " /></code>	This option automatically converts any Microsoft office Document. An interactive user is not needed to be logged on to the server.
<code><add key="AutoExtensionOpenNative" value="" /></code>	This option needs open office to be installed for it to work. It handles the same file types as the OpenOfficeExtension option, but performs the conversion differently.

17.8 BCL Printer Install/Uninstall

In previous versions of Autobahn, there were issues during the installation and un-installation process caused due to the BCL Printer queue not being empty. This has now been fixed.

17.9 Blank Page Removal in Extended OCR

An additional feature has been added in Extended OCR to allow removal of blank pages (both for source TIFF and source PDF files). See section 14.3 in the Reference Guide for more information.

17.10 Native Processing in Extended OCR

An additional feature has been added in Extended OCR to allow native processing when using Extended OCR Image PDF. See section 14.3 in the Reference Guide for more information.