bitfarm® ARCHIV

Dokumentenmanagement

System Manual

Version 3.5.0

GPL-Edition

Table of contents

I.   Server installation

1.   Preparations / System Requirements

Please make sure that your system meets the requirements:

- Windows Server 2012/2012 R2, 2016, 2019 including Domain, or
- Windows 7, 8, 8.1, 10 including user administration for minor installations
    - Like GPL 3.4.5 available for older Windows version, but without guarantee nor support

- Functioning name resolution
- Enough storage capacity (600 MB used by the program itself; size of document storage depends on the amount of documents)
- An adequate CPU and enough RAM-storage
- Latest Windows Scripting Host (version 5.6 or higher)
- Configuration of the firewall and anti-virus software, as to allow VBS and MySQL communication through Port TCP 6603
- Data backup
- UPS if possible

2.   Executing the bitfarm-Archiv server setup

a)   Start the bitfarm-Archiv-Server-Setup program.
b)   Select the target folder.

The installer will recognize previous bitfarm-Archiv GPL versions and will upgrade to the latest version; uninstalling the old version is **not necessary**.
After the installation you will see a text file named `Zusammenfassung.txt`, which includes all the necessary information (user name/password for the service user and MySQL) as well as an install.log for troubleshooting in the installation path (usually: %programfiles(x86)%\bitfarm-archiv\).

3.   Updating an older version of bitfarm-Archiv GPL

If an older version of bitfarm-Archiv GPL is already installed, the installer will recognize it and notify the user of an update. Already existing documents will not be affected, although backing up the current system for security purposes is highly recommended.

4. What's new in the GPL version 3.5.0?

- <u>Workflows</u>
- <u>HotSearch</u>
- <u>Microsoft Office Add-Ins</u>
- Updating the OCR-Engine to TesserAct *(see)*
- Procedure documentation (technical part)

5. Working with the bitfarm-Archiv Administrator

a) Start 'AdministratorV3'from the bitfarm-Archiv program group.

b) Select MySQL-Server GPL-3.4 and click on 'connect to SQL-server'.

c) Insert name and password of the MySQL user (sa/sa), created during the installation. Optionally, click on remember login data.

d) Now, click the area that says 'Datenbankname', select 'bitfarmv3' and click on 'Datenbank öffnen'.

e) Create additional users and groups under the 'users and groups' tab. Create new archives and storage folders under the 'Archive' tab and bind new archives beneath a storage folder. Select an archive and move users and groups that get access to the archive from 'free users and groups' to 'users with access'. Select them once again and tick the boxes for the needed authorizations. Additional information on how to create your DMS-structure can be found in the section on *DMS-planning help*.

Now you can create additional users, storage folders and archives with additional panels and information panels as well as allocating authorization to them. (Advice on how to create your DMS-structure is available in the appendix.) Archives should not contain spaces or characters that are incompatible with the system; use _ or – as connectors. Optionally, create a 'bin' folder. For every archive, create a template with the same name. You will also receive a notification by AdministratorV3 for this, in case any changes were made.

Please note that archives always make up the last node in a tree and therefore cannot contain any folder beneath them. Storage folders, on the other hand, can be stacked as many times a necessary. Furthermore, note that every archive-tree has to have a clear root. No archive or storage folder should be without its root.

This concludes the installation of the server. Now you can try starting the client to see if everything is working as it should. Open ViewerV3 from the program group 'bitfarm-Archiv', insert user name and password of the archive-user we have just created and the viewer should open and show the tree structure.

II.   Installation of the client

1. Requirements

To ensure the flawless execution of the bitfarm-Archiv client, you have to meet the following requirements

- Windows 7, 8, 8.1, 10; Windows Server 2012/2012 R2, 2016, 2019
  - Like GPL 3.4.5, available for older Windows version, but without guarantee nor support
- Latest version of Windows Scripting Host (Version 5.6 or newer)
- Access via Windows file share on the DMS-server, preferably via domain
- Anti-virus software should be configured as to allow VB-scripts even when they are executed from a different device in the same network
- The Firewall has to allow access to the MySQL-port of the DMS-server (standard 6603, can differ when configured manually)
- Write-access under `HKEY-CURRENT-USER` has to be granted
- In order to execute scripts of the client without interruption, register the server under 'local intranet' in the format in the security settings of the Internet Explorer. Untick the first box under 'sites' and then tick the three ones beneath it.
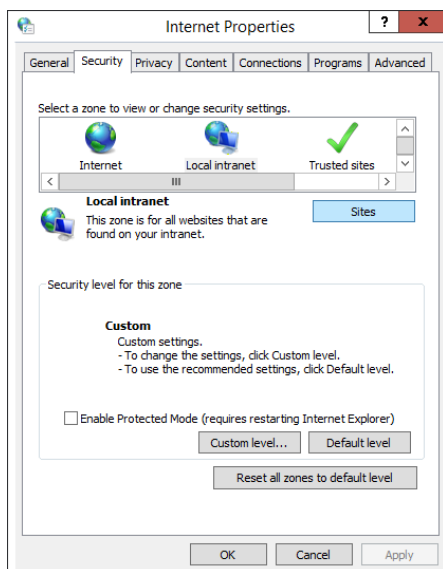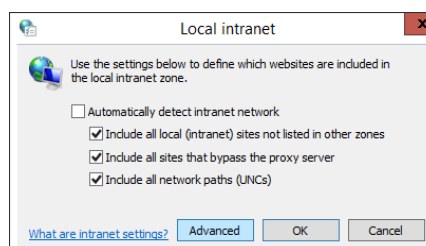

Image 1: internet-options
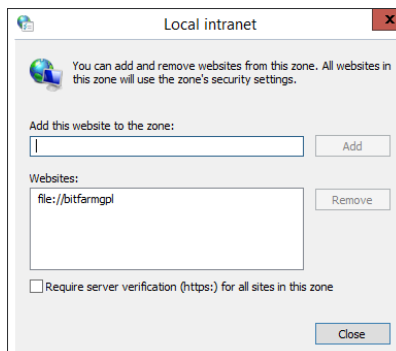

Image 2: Local Intranet - Sites

Image 3: Advanced

2. Installing the bitfarm-Archiv client

1. Log in to the client as an administrator and start the bitfarm-Archiv client setup.
2. After the license-agreement select the NETBIOS-name and IP-address of the bitfarm server. Enter the bitfarm-Archiv file sharing if it differs from the preset: 'bitfarm-archiv$'. Enter the name of your CON-file under 'profile' (extension: .con in the bitfarm-Archiv folder; default: bitfarm)
3. In case you want to set up an additional archive printer, enter the necessary information (a standard archive printer will be set up automatically)
4. The bitfarm-Archiv client installer will install the following components:
   - Shortcuts in the start menu and on the desktop
   - HotSearch (Autostart)
   - Archive printer (including RedMon printer-bypass)
   - Special archive printers
   - 'Send to'-shortcuts
   - Microsoft Office Add-Ins (Word, Excel)
   - .NET Framework

3. HotSearch

If you need to search for documents created by other Windows applications, or if their data is used by other applications, you can use HotSearch to quickly bring up the desired document.

Through HotSearch you can make a quick search in the viewer by using hotkeys. This can be used, for example, to send an invoice-number from an ERP-program, in order to find the related receipts in the bitfarm-Archiv. For this you have multiple options:

- Clipboard: If the desired value exists in text form you can simply copy it via Ctrl-C and search for it using a hotkey
- Screenshot-Tool: Inside the screenshot-tool you can access a selection-tool through a hotkey. Use this to select an area on the screen on which you would like to perform a full-text recognition. The recognized text can then be used to perform a search in the viewer.

- <u>Shortcut-Search:</u> In order to search for shortcuts, activate the shortcut-search
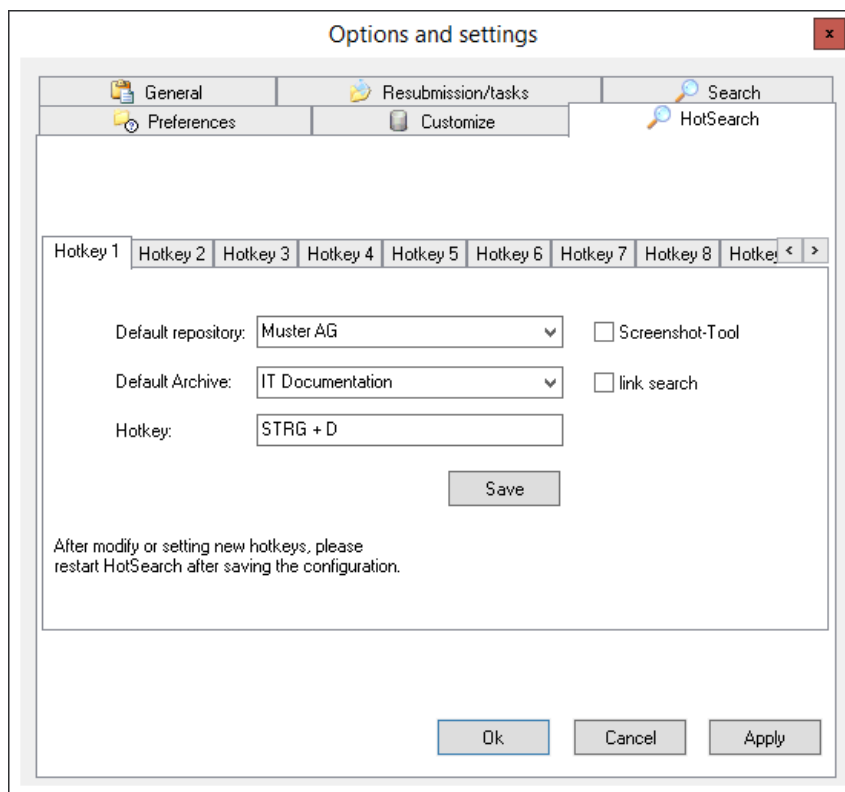
tool.

Attention:	The default hotkey for HotSearch is **Ctrl-B**. The default searching location is the last used archive. The full-text search will, by default, be performed here.
Attention:	To apply changes in the configuration, HotSearch needs to be restarted.
Attention:	'Apply' does not 'Save' the changes!
Hint:	Shortcuts can be used to contextualize documents based on content. For example, commercial documents can be linked to a shortcut based on the order number. The shortcut-search then shows all documents from a single order number from all archives. However, it will take into account the rights the current user has in terms of accessibility. One user may have find documents that are inaccessible to someone who does not have access to certain archives.

4. Installing an additional archive-printer

a) Set up a new printer 'archive printer' as a local printer on a new rerouted connection. For drivers, use a PostScript printer. The default for bitfarm-Archiv is the *Generic MS Publisher Color Printer*.

Attention: It could happen that the connection cannot be established immediately. In this case you have to connect it to an already existing port and set up the rerouted connection via '**properties**'. Now the printer can be connected to the rerouted connection through 'settings' and set up as shown below.

b) Open the properties of the new 'archive printer', select 'connections' and then 'configuration'. Set up the connection as shown. Enter the UNC-path to the bitfarm-Archiv program directory as well as the name of the profile into the server. If no profile is selected, then the profile used for the current viewer session will be selected. The default name is 'bitfarm'. Should the latest viewer login come from a different database, an import panel with that database will open.

Hint: Some Windows versions require programs to be run as an administrator (right click, run as Administrator).

Hint: It is also possible to set up a printer that prints the document simultaneously (for that a profile has to be inserted: profile=bitfarm) the name of the printer will be sent as third parameter. The printer has to be set up on the DMS-server which is also possible for printers on a printing-server. In that case use the following:

wscript \\dms-server\bitfarm-archiv$\druckimport.vbs %1 profile printer1

or

wscript   \\dms-server\bitfarm-archiv$\druckimport.vbs   %1   Profil \\Printserver\Drucker1

Attention: The name should not contain a dash; otherwise GhostView will display an 'Error Parsing Commandline' error message!

Attention: Because the import-panel does not show up in terminal sessions, the default printer cannot be used on a terminal sever yet.

Attention: In order to print from a color printer you have to configure GhostView. To do this start GhostView and select 'Print' from the 'File' menu, then select a Postscript-file to access the printer-configuration. Now you can select the printing method 'Windows-GDI Printer' from the settings and change the color-output of GhostView to 'full color'.
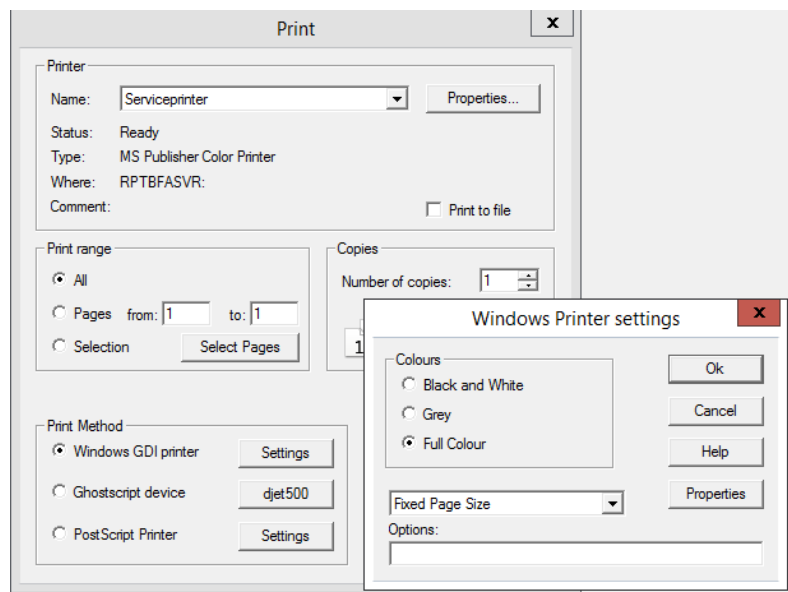
5. Microsoft Office Add-Ins

With Microsoft Add-Ins you can now make your processes even faster. In the GPL version, Add-Ins for Word and Excel are available, whereas in the Enterprise version Outlook-integration is available too. Just select the 'bitfarm-Archiv' tab and click on the icon for archiving the document.
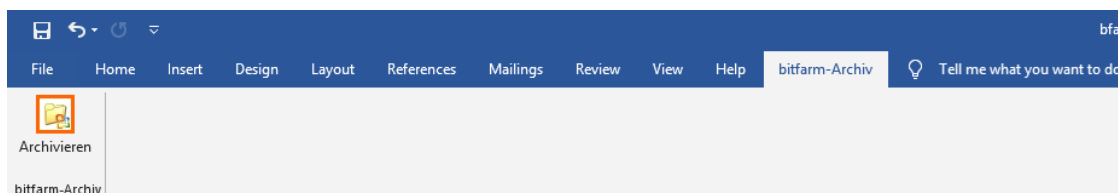
6. Configuring the manual scanning program

Clients that are connected to TWAIN-compatible scanners should always use manual scanning as the default tool for archiving paper-documents. The installation and configuration works as follows:

Start the ManuScanV3-Programm from the bitfarm-Archiv program group. Go to extras, 'options' and then 'Scan-Profile'. Set up the necessary scanning profiles (global) for black-and-white and color scans respectively. Now define the default scanning profile.

Hint:　　　For B/W scanning profiles we recommend 300 dpi and CCITT4-compression, for color 150-300dpi with JPEG or LZW compression. Kodak-scanners have a color-dropout available in the driver. Pay attention to the settings of the ADF (Automatic Document Feeder activated) and Staplescan (Creating a separate document for each page).

Hint:　　　We recommend that first; you scan the document with a high resolution and then scaling it down through a filter to save space. This way the OCR receives a better sample. However, scaling down can elongate the server queue and should not be done in time-sensitive situations.

Attention:　Scanning with the bitfarm-Archiv requires a working TWAIN-driver. Always use the latest manufacturer's driver and not the WIA driver. TWAIN is usually not available for Microsoft terminal-servers.

7. Configuring an automatic scanning-icon

Like printing, the scanning and archiving process can be automated so that no manual action is required. For this create an 'Autoscan' shortcut on the desktop and/or in the program directory under 'All Users'. Open the shortcut's properties and enter the following parameters after 'autoscan.vbs' in the 'target' panel:

1. the target archive,
2. the scanning profile you have created with ManuScan preceded by a -G for global profiles and -U for local profiles,
3. (optional) the graphics-filter used,
4. (optional) the timeout-value for scanner-inactivity (in tenths of seconds) -t:%n
5. (Optional) the threshold for detecting paper jams (data throughput in bytes) -j:%n.

Example:

\\SERVER\bitfarm-Archiv$\autoscan.vbs invoices -G:300dpiSW 300 -Standard -t:50 – j:1000

Attention: If users are not main-users on the clients or on the terminal-server, you have to allow access to the temp-directory (change authorization).

8. Determining appropriate values for the detection of paper jams and inactivity

In order to determine suitable values for passing parameters, you have to set `scriptdebug=True` in the Autoscan.vbs. If you scan now, the Autoscan.vbs creates a log-file (`%temp%\bfa-scanner.log`). Focus on the second-to-last line which shows the average read-throughput. In case of a paper jam the value will rise abruptly. To identify a paper jam, determine a value that clearly indicates one happening.

10.05.2007 16:10	1396	6	4	2	6	4,5	0
10.05.2007 16:10	1400	4	6	4	2	4	0
10.05.2007 16:10	1404	4	4	6	4	4,5	0
10.05.2007 16:10	1408	4	4	4	6	4,5	0
10.05.2007 16:10	1412	4	4	4	4	4	0
10.05.2007 16:10	1416	4	4	4	4	4	0
10.05.2007 16:10	1423	7	4	4	4	4,75	0
10.05.2007 16:10	1423	0	7	4	4	3,75	0
10.05.2007 16:10	1484	61	0	7	4	18	0
10.05.2007 16:10	2542	1058	61	0	7	281,5	0
10.05.2007 16:10	3992	1450	1058	61	0	642,25	0
10.05.2007 16:10	5616	1624	1450	1058	61	1048,25	0

Example: bfa-scanner.log -abrupt rise of the average read throughput to >1000

For the timeout-parameter, focus on the last line. Here the number of average value=0 are counted in. Please note that, temporarily, data will not be read, even during a scan. Therefore you have to determine a limit, below which the scanner will continue its work indefinitely and send a timeout-notification should it rise above it.

10.05.2007 16:09	91	0	0	0	0	0	27
10.05.2007 16:09	91	0	0	0	0	0	28
10.05.2007 16:09	91	0	0	0	0	0	29
10.05.2007 16:09	91	0	0	0	0	0	30
10.05.2007 16:09	91	0	0	0	0	0	31
10.05.2007 16:09	91	0	0	0	0	0	32
10.05.2007 16:09	91	0	0	0	0	0	33
10.05.2007 16:09	91	0	0	0	0	0	34
10.05.2007 16:09	91	0	0	0	0	0	35
10.05.2007 16:09	98	7	0	0	0	1,75	0
10.05.2007 16:09	99	1	7	0	0	2	0
10.05.2007 16:09	99	0	1	7	0	2	0
10.05.2007 16:09	99	0	0	1	7	2	0
10.05.2007 16:09	99	0	0	0	1	0,25	0
10.05.2007 16:09	99	0	0	0	0	0	1
10.05.2007 16:09	99	0	0	0	0	0	2
10.05.2007 16:09	99	0	0	0	0	0	3
10.05.2007 16:09	99	0	0	0	0	0	4

Example: bfa-scanner.log -Counter for read throughput = 0 reaches a value of 35 while scanning.

Appropriate parameters for the	Kodak i30/i40:	-j:1000 –t:40
	Kodak i1220:	-j:170 –t:60

III.   Further configurations and troubleshooting

1. Configuring watched folders

In order to send documents from a scanner, or other applications to the DMS, you need to create a folder on the DMS-server and enable file storage. In the scripts.ini you can then enter the path through:

`ScannerImportPath=`

and/or

`ExtendedImportPath=`

Now, the spool watches this folder and every folder that has been created within it (only in the scanner-import-path). Files that are copied into the folder are collected by the spool, put into a queue and deleted from the watched folder. This usually happens fast enough that the folder should always appear empty.

Documents from the watched folder are sent to the 'unsorted archive' by default, but can be sent to any target-archive by using sorting-rules (see: section 2 and 3 of this chapter).

Sorting can also be done based on the name of a sub-folder inside the main folder. To do this, put the files into a folder that has the same name as the target-archive. Now you just have to set

`readarchivfromfolder=true`

in the scripts.ini. Furthermore, there has to be an import-sample of the same name (.tpl in the 'templates' folder of the bitfarm-Archiv directory; created in AdministratorV3) related to the archive. Sorting by name of the document is also possible. Another option is to use the folder- or file-name as user-name instead of just 'System'. To do this, enter the appropriate settings into the scripts.ini.

2. Sorting automatically based on the content of documents

Open the file `bitfarm.wfd`. Create a new sorting-section based on the logical sequence of the process according to the following example:

```
Sorting section Macrotron invoice        <-- Name of the section
andindocument=Ingram Micro               <-- Requirements
andindocument=order
andindocument=net
andindocument=due
notindocument=receipt
sortto=supplierinvoices                  <-- sorting-instruction
end section                              <-- End of section
```

With `andindocument=` you can determine which words and word families NEED to appear in the full-text, in order to execute the sorting-instruction. Save the .wfd file and the changes are applied immediately to the next document, without having to restart. You can also use `orindocument=` and `notindocument=` for the logic-functions.

Hint: If documents need to be sent to special archives and therefore exempt from the automatic sorting, define the appropriate target archives with `nosorting=` in a `settings section`. This can, for example, prevent accidentally making classified documents, from management, available to everyone through an automated sorting process.

3. Setting up an automated keywording with trigger-strings

To save certain information from documents into database fields = additional fields, you can set up definitions for an automated keywording process. To do this, edit the .wfd file. Options are shown in the following example:

```
naming section offers          -> Start of the section
archivtabelle=offers           -> Condition: Folder 'offers'
displayname=bitfarm-offers     -> Change of title
searchstring=date              -> search String 'date' (Trigger-string)
zusatzfeld=process-number      -> fill in field 'process-number'
searchsteps=1                  -> ...with string: 1. step...
searchdirection=1              -> ... left from the search string.
searchstring=company           -> next block: search string 'company'
zusatzfeld=customer-name       -> fill in field 'customer-name'
getstrings=3                   -> take 3 strings
```

```
nodeletespaces                          -> do not delete spaces
searchsteps=1                           -> one step from search string ...
searchdirection=2                       -> ... to the right
                                        -> (searchdirection completes the block)
...                                     -> if needed, more blocks
end section                             -> end of the section
```

Hint: To avoid errors, check spelling. Upper- and lowercase may be distinguishing factors! Keep the sequence exactly as shown above.

Hint: `andindocument=` can also, or exclusively, be used as condition.

Hint: Instead of `searchdirection=1`, you can also use `searchdirection=left` and instead of `searchdirection=2`, you can use `searchdirection=right`.

Hint: If user-keywording (for example from ManuScan or Importer) needs to be further edited through a naming section (e.g. if you want to use the user-input as a shortcut), you can use `searchstring=userinput`. Nevertheless, `Zusatzfeld=`, `searchsteps=` and `searchdirection=` have to be entered into the block, even if they do not have a function.

4. Setting up an automated keywording with graphical interfaces

If trigger-strings are not good enough to extract a certain keyword, bitfarm-Archiv also supports graphical extraction as well as targeted OCR in certain sections. They can be set up like this:

a) Save the TIF-file of a sample-document (sample) for this keywording on the hard drive (e.g . . . \bitfarm-Archiv\masks\sample-invoice1.tif). This file needs to be of OCR-compatible resolution. Sample-files from the archive are usually scaled down from the required 300dpi post-OCR. The best solution would be, to stop the archiving and extract a sample document. Now you can take the temp-file from the queue, set up the keywording and use it for testing the process right away.

b) Open `DocCutV3.exe` from the bitfarm-Archiv program-directory, click 'file' and 'load template', then load the sample-document. Click 'new cut-list' for a new keywording, or 'load cut-list' to edit an already existing cut-list. Cut-lists have to be saved in the 'masks' folder of the bitfarm-Archiv program-folder.

c) Use the mouse to draw a box around the keyword that you want to extract. Make sure that the area is free from any undesired elements in the preview.

Make also sure that you leave a certain tolerance-area around it. Depending on the quality of the scanner and its paper transport, the position may change. This is why the tolerance-area needs to be wider to ensure the correct recognition.

d) Name the new cut-mark. It has to correspond to the additional field that needs to be filled in. To save, click on the green tick.

Hint:	Rotating the extracted area is also possible before the OCR. To do this, open the .CUT-file with the editor and add the 'rotation' statement into the related section:

```
[order-number]
Left=1968
Top=2013
Width=60
Height=249
Rotation=1
```
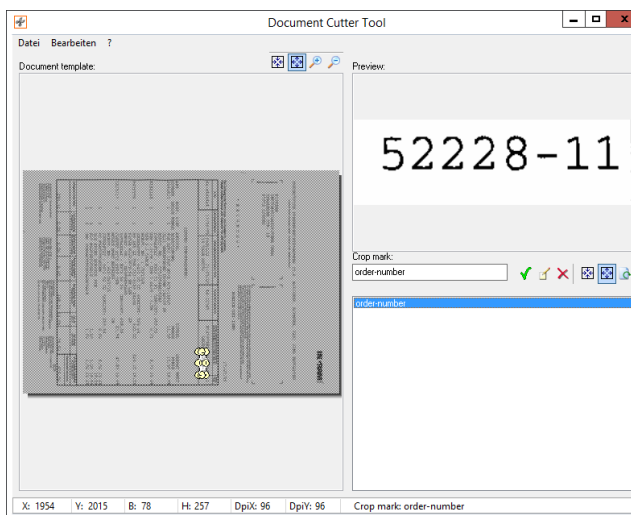

e) Enter the cutlist into the naming section of the workflow-file.

Example:

```
naming section delivery-note1
andindocument=Ingram              -> condition 1
archivtabelle=delivery-notes      -> condition 2
displayname=Ingram_delivery-note
rotation=90                       -> rotate by 90 degrees when archiving
mask=Ingram-LS-Nr                 -> evaluate cut-list
cutawayleft=2                     -> remove 2 symbols to the left (44-)
end section
```

5. Using internal variables for additional fields

At different stages of processing a document, many details are saved into the internal variables of the DMS, which can be used for additional fields.

You can establish which internal variables are copied into which fields, in the 'settings section' of the .wfd file. Fields can, in this case, be identifiers, that do not end up in the target archive. These 'wrong' fields can be used internally to make plugin-information available. The plugin can thus fill in a 'real' additional field.

Example of a setting-section with USE statements:

```
Settings section
nosorting=Dokumentationen
nosorting=GL-Verträge
USE [Richtung] FOR %direction%
USE [Absender] FOR %sender%
USE [Betreff] FOR %title%
USE [Empfangsdatum] FOR %date%
USE [Versanddatum] FOR %date%
USE [Empfaenger] FOR %receiver%
USE [Faxnummer] FOR %receiver%
USE [Originaldatei] FOR %Plannr%
USE [Originaldatei] FOR %Docname%
USE [Titel] for %title%
USE [Herkunft] for %category%
USE [Autor] FOR %author%
USE [Betreff] FOR %Dokname%
end section
```

The internal variables that are mapped to the additional fields that are surrounded by % are in square brackets. The following internal variables are available right now (List is constantly expanding):

Originaldatei	Dateiname ohne Endung
Betreff	E-Mail-Betreff
Absender	E-Mail-Absender
Empfaenger	E-Mail-Empfänger
Richtung	E-Mail, Fax (SBS) Richtung: ein-, ausgehend oder intern
Empfangsdatum	E-Mail, Fax (SBS) Empfangsdatum
Versanddatum	E-Mail, Fax (SBS) Versanddatum
Titel	Word, PDF, Dateieigenschaft „Titel"
Autor	Word, PDF, Dateieigenschaft „Autor"

Herkunft	Scan, Datei, E-Mail, Fax, Druck
Extension	Dateinamen-Erweiterung (doc, pdf etc.)

Hint: %Dokname% is a special additional field. It enters the name of the document into the list of results. See `displayname=` in a naming-section.

Hint: An already existing entry will NOT be overwritten. This way you can map multiple variables into an additional field (the first variable with content fills in the field).

Hint: You can fill in an additional field in a naming section with a with a certain fixed value by using `fixvalue=`fixed value

Attention: You can also send multiple values from virtual additional fields.

Example:
```
….
fixvalue=%Autor%-%Titel%
zusatzfeld=Name des Zusatzfelds
…
```

6. Automatic linking

Links allow you to organize documents based on content; alternatively to the tree-like formal structure of the DMS. They can link all documents of a project together. This creates chains of documents, virtual dossiers that exist beyond the archive structure. The individual documents can be linked to multiple chains at once.

To link documents automatically, you can select one or more keywords from the naming section and use them as linking-numbers. The relevant syntax looks like this:

```
naming section bitfarm-invoice
archivtabelle=customer-invoice        -> condition
displayname=bitfarm invoice
mask=offer                            -> evaluate mask
linkstring                            -> save as shortcut
end section
```

You can create links from several different values, using the `createlink` statement. Example:

```
naming section accountingdocumentation

archivtabelle=accountingdocumentation

createlink=%companynumber%-%accountname%

end section
```

This creates a link from two values conjoined by a dash. Let us suppose that the additional field 'company-number' is '01' and the name of the account '4711'. This would create the following link '01-4711'.

Hint: `createlink` works well for meeting special formatting requirements of an ERP-system, so that you can use a shortcut directly from the ERP.

Hint: Links make the fastest global searches possible, through the 'reference' tab in the viewer.

7. Workflows

Admins also need to establish workflows, which assign certain people or groups to certain work with incoming documents, which they then send to a different person or group for further work. To establish an automatic workflow, you need the WFD-file. Open the file (preferably with Notepad++) and add a `workflow section`, like the following:

```
workflow section invoices
archivtabelle = supplier-invoices
sendmailto = test@companyxyz.com
sendheadr = new incoming invoice
sendmsg = check for correctness
nextworkflow = release
termin = 0
alarm = 3
bearbeiter = #editor
controller = j.smith
end section
```

archivtabelle establishes the condition, that this workflow is only triggered, when the document is archived or sent into this exact archive. With `sendmailto` you can send an E-mail with the subject in `sendheadr` and the message in `sendmsg`, if the SMTP-server is set up. In the field `nextworkflow` you enter the next step of the workflow

which will be set up in the `continue section`. With termin you can set a point in time for resubmission. Here it is 0 days, which means that the workflow triggers as soon as a document is sent to the archive `supplier-invoices`. The command: `alarm (in days)` when the document is due. The document then turns red and you can send a mail to the controller with the command `options=mailonalarm`.

`bearbeiter` stands for the person or group that receives the document in the list of resubmissions, as a task.

Attention: If the document needs to be sent to a group for work, you have to enter the name of the group preceded by a # or a §. `#bearbeiter` means that only one member of the group needs to finish the work to mark it as done and have it disappear from everyone's workbench, while `$bearbeiter` means that every member of the group needs to mark the task as done by clicking on *done/continue.*

In the `continue section` the workflow enters a possible next step. For this you need to enter `continue=`, in which you have to enter the entry from `nextworkflow`. Now you can check additional conditions using, amongst others, CASE-inquiries as shown in the following:

```
continue section invoices
continue = release
CASE Status_geprüft = true
sendmailto = name@companyname.com
sendfile
bearbeiter = $accounting
controller = j.smith
end section
```

In this case the `CASE`-inquiry checks if the status-field is 'checked'. If that is the case, an E-mail with the file (`sendfile`) gets send to the address `sendmailto`. Editor of the list of resubmissions is the group accounting, In which **every member** needs to click on done/continue to finish the workflow.

Attention: you can have multiple continue sections, separated by different conditions.

Hint: To catch bad user-behavior, you can create loops by having a section reference itself and show every possible `CASE`, until the user corrected the error.

Attention: If you use multiple `CASE`-inquiries, all of them have to be met to run the section in the first place. `CASE`-inquiries are therefore AND-linked.

8. Checking for duplicates

The duplicate-check establishes a test value through the full-text of a document or its original file, which then gets compared to the test values of previously archived documents. If the values of the two documents are similar, the new one will be sent to the folder 'duplicates'. The admin can then, if necessary, delete the file completely. The following switch influences the duplicate-check:

In the relevant .con-file there is the statement `dublettenpruefung=false` which can enable or disable the duplicate-check.

In the scripts.ini the statement `preventdoubledocsmemory=35` where the number of recent documents, which the system should remember for the duplicate-check, is set.

Hint: The duplicate-check has a success rate of 70% with scanned-in documents, but can reach 100% with spool-documents or files.

Hint: Test values of the last documents are saved into the '{Profil}-checksums.txt' file in the bitfarm-Archiv program-folder. You can delete this file to restart the duplicate-check.

Hint: Many more features are available in the Enterprise-Version of bitfarm-Archiv:

9. TesserAct-OCR optimization

bitfarm-Archiv contains different configuration-switches for optimization of the TesserAct-OCR. These are:

- `usetesseract=`
  - Default: True – For backwards-compatibility you can use the CuneiForm-OCR through false. *(see)*
- `tesslang=`
  - Default: deu – the language packages being used. If you need additional Sprachpakete, you can download them from GitHub etc.
  - Folder: .\bitfarm-archiv\Teaseract-OCR\tessdata\
  - Example: tesslang=deu+eng
- `tessmaxpagetime=60`
  - Timeout-limit. The timeout-limit should be raised on slower machines; otherwise documents cannot receive full-text.
- `tessprocs=4`
  - The maximum of pages processed at the same time.

10. Deleting documents

bitfarm-Archiv offers a tool to delete documents in accordance with privacy laws. This tool is available at: 'Downloads & Support'-> „Downloadbereich" -> „Weitere Downloads". Documentation is available in the file `bfa_delete_arc_gpl.zip` .

11. Troubleshooting server-problems

The most common errors on servers are documents that are processed incorrectly. It is possible that with 'bad documents' the OCR works in an infinite loop, or generates an error. In this case, the queue is not emptied anymore. To fix this, the 'bad' document needs to be removed from the queue to have it continue with the other documents. This can be done in the following way:

a) Stop the server-clients with 'Stop Server' (bitfarm-Archiv program-menu)
b) delete the files: 'lock', 'directlock', and 'superlock' in the 'temp' folder inside bitfarm-Archiv, if the show up
c) Open the task-manager and end the processes „filterv3.exe", „dbimpv3.exe", „gswin32.exe" and „gsview.exe", should they appear
d) delete the 'bad' document from the queue-folder by deleting the related temp-files
e) restart the bitfarm-Archiv clients ('Start Server')

In general, you should stop everything with 'stop server' and then restart. You can also restart the entire system, if needed. Error-messages are recorded in the application-log with the source WSH, which can be viewed through the Windows management-section.

Hint: If an error-message shows that some documents entered the system without preview, they can be sent through the process once again through the 'refresh document' tool in the viewer. But, this can only be done one-by-one.

Hint: Oftentimes, lack of authorization for the client-user or the user on NTFS- and enabling levels are the reason for not being able to use the program, especially after fresh installations. Errors like these can be easily identified through notifications in the event-log.

12. FAQ-server

Problem: The server does not process and archive documents.

Solution: There are two bitfarm-clients running on the server.

First, check if these run or if the client-user (usually bitfarm) has full access on the program-folder and its sub-folders (default path: C:\Program Files (x86)\Bitfarm-Archiv) as well as the authorizations.

The first client called 'bitfarm-Archiv spool-client' checks if documents (files) are sent by a client. These are usually sent to the transfer-folder \bitfarm-archiv\uebergabe\ or .\bitfarm-archiv\import\, from which the spool-client send them into the queue-folder, together with the related .job-file and a .slw-file (for keywords).

If there are no files sent from the transfer-folders to the queue, there might be a problem with the client. Check if the client is running (Control Panel -> System and Security -> Administration -> Clients) and if the client-user has full access to the transfer-folders (NTFS and Share).

The second client is the 'bitfarm-Archiv archiving-client', which usually start the Archivierung.vbs to process the next document in the queue. If you do not notice any changes in the queue-folder over the course of several minutes, check if the client is running and if the client-user has full access to the entire bitfarm-Archiv program-folder.

Every document gets identified by a .job-file, therefore, check if the .job-file exists or if its content is valid; definitely check, if the entered profile has a corresponding .con-file.

Now, stop the clients through the client management, or via 'stop-server' (in the bitfarm-Archiv program-menu) and start the script: Archivierung.vbs manually.

If, during the process messages appear, it could be them that are responsible for the error of the script, since these messages do not appear when the script runs as a client.

Problem: The OCR-recognition does not work. Documents do not have full-text.

Solution: Check if the .job-files in the queue have an entry: OCR_Typ=CRE.

If you use the CuneiForm-OCR, .NET-Framework 2.0 needs to be installed correctly and the apuma.dll needs to be registered. You can register this manually by entering:

```
regsvr32.exe   /s   "bitfarm-installdir"\Viewer-
Files\bfaOCR\apuma.dll
```

into the command line.

Problem: The automatic keywording does not work anymore after updating.

Solution: With the GPL-update from version 3.4.5 to version 3.5.0 the OCR-engine has been changed. Because of this change, the design of the full text may be different. If an extensive keywording already exists, you can change the full-text recognition back to CuneiForm. For this, you need to set the switch usetesseract=False. After restarting, the full-text recognition should provide full-text again in the format you are used to.

Attention: This switch should only be set to false, if CuneiForm is available through an early installation of bitfarm-Archiv.

Problem: No connection to the MySQL-Database.

Solution: Check if the MySQL-client is running and accessible. By default, Port 6603 is set up for MySQL when performing a fresh installation. This could be different, if you have upgraded from an older version of bitfarm-Archiv.

(look into the my.ini-Datei in .\bitfarm-archiv\mysql-5.1.xx-win32\; the port is defined in the [client]- and [mysqld]-section.)

Make sure that the port can be accessed through the firewall. It also needs to be defined in the same way in the bitfarm.con and in the bitfarm_Demo.con.

13. Troubleshooting client problems

Most client problems can be fixed through a reconfiguration. For this you need to edit the registry through

Regedit. Located in HKEY_CURRENT_USER\Software\bitfarm are all the settings for the behavior of the DMS-parts run by this user. To fix the problem, close

all bitfarm-Archiv programs delete the branches individually or entirely and restart the program. Special configurations might have to be reset.

Other problems might be due to authorization, or lack thereof. If, for example, there are no problems while running the program from an admin-profile, but not as a regular user, then the user might not have authorization for certain activities. Oftentimes, the user does not have the right to read or write a file on the server. Check the authorizations by trying to open files on the DMS-authorizations via 'start' -> 'run' as a user. The required authorizations can be viewed in the manual of the server installation. Also make sure that local or group-guidelines are not hindering the execution of the DMS-programs. Also try to disable or uninstall anti-virus software and firewalls for testing. On most Windows-systems, the firewall has to allow access in order to use the MySQL-ports (usually 6603) that the bitfarm-Archiv requires.

It is always recommend to equip all users of the DMS with main-user-rights on the clients (Move the group of the domain-users to the group of the local main user).

In case of problems with sending documents to the server, check if the user is allowed to write into the folder. Stop and restart the server-clients on the server using `TERMALL.bat` and `STARTALL.bat` (with admin-rights -> right click -> run as an Administrator)

If the viewer tends to crash a lot, you can find reasons for this in the file `ViewerV3.log`, which can be found in the `%temp%` -folder and can be accessed via 'start' -> 'run' -> `%temp%\ViewerV3.log`.

Hint: If some documents do not show up in the viewer and the message „document cannot be read" appears on the bottom of the page, lack of network-rights (NTFS, access) of the current user can be the cause of the problem.

Attention: When using a second database (e.g. 'Bitfarm-Demo-database') you need to log into the target database (ManuScan, Archive-printer without profile details 'send to'), before importing a document, to have the correct import-window open up. Otherwise the incorrect importer-window might open up.

Imports from watched folders (autoscan etc.) that do not open an importer-window, generally target the default archive, regardless from which database you were logged into most recently.

IV. Appendix

1. DMS-planning help: Concepts and data-structures

To achieve the optimal DMS-structure, it is vital that you not only understand the logic behind the processes documents undergo in the company, but also the capabilities AND restraints of the DMS. The basic structure of any given company is fairly similar, whereas it can differ greatly in the details. These details are usually weak points, since they consist of established processes that work best for this particular company and should, in theory, be carried over to the DMS. Our goal is to visualize many different processes within the software. Ideally, there are more than one way to solve tasks with the GPL- and Enterprise-version of bitfarm-Archiv.

Unfortunately, software development suffers from the fact that the absolute ideal can never be accomplished. Thus even bitfarm-Archiv has its fair share of limitations, which cause not every document-concept to be fully realized in the software. This is why, in the following, we will outline the current limits as well as the most important conceptual considerations.

The main idea behind bitfarm-Archiv is that of organizing documents into a tree-like structure. This structure is established by the administrator and cannot be changed by the user. It outlines filing-storage for documents that are formally similar: Invoices in an invoice-folder, delivery notes in delivery-note-folder, E-Mail into one or more E-Mail-folders. Authorization is based on these folders. The status of the process can be visualized by status fields. These, together with the additional fields, are based, in turn, on the folders. Relations in content are visualized, independent from folders, through links between documents. The authorization-structure of the archives only allows users with the necessary rights to view documents that they are allowed to view. To increase speed of the search, you should use links, rather than performing an archive-independent search for additional fields. Make use of the option 'shortcut' (as administrator) even on the additional fields to automatically generate shortcuts via the text input.

Outlook: In the Enterprise-version you can make the same document accessible in multiple archives via 'symbolic-links'. This offers alternative possibilities for creating a 'dossier'.

One of the concepts of bitfarm-Archiv is that documents that are scanned in are immediately put into the archive they belong to for the duration of their existence. This can help avoiding the 'mailbox-problem' which is when, during the standard document-handling, processed documents are put into 'mailboxes' of the employees and cannot be accessed by anyone else. It is best to avoid carrying this particular

practice and its disadvantages into the digital age. Processes can be done in the final archive using: additional fields, annotations, status fields. This also suggests that users should ideally not be able to remove documents from the final archive.

bitfarm-Archiv generates a spreadsheet for each folder (archive) in the MySQL-database. This gives you the ability to handle a big set of data through the DMS. In each archive with up to 500,000 documents a full-text search can be done under 10 seconds. The search via keywords or additional fields (within an archive) is also very small, even with very big data sets, thanks to the binary-tree. This has the advantage that through a smart archive-structure, full-text searches can be kept very short in time. With the OCR-component, you can avoid using a complicated manual keywording.

To make this work, it is important that the search is, in most cases, restricted to just one archive, since this is when the database is the quickest. For performing a full-text search in multiple archives within a superordinate folder, the system needs to generate an interim table and can thus take a lot more time. If possible, you should avoid having to perform these kind of searches regularly by establishing the archive-structure accordingly (This limitation does not exist in the Enterprise-version 3.6 since it makes use of a new data set)

For further categorizations within an archive, you can use the additional fields. This way you can visualize document-categories with, for example, a pre-made drop-down menu (check-boxes). These fields can also be used to improve the overview and the structure of the search-results in the list of results. For the targeted mass-processing (E-Mail, printing, keywording, export), you can use the list of results with Ctrl-A + action or by selecting the results individually while holding the Control key.

If there is no way around regularly searching through multiple archives, you can use the regular keyword-search to reduce time. Then the additional-field-information will also be searched which is usually not available as explicit fields on the storage-layer.

The best way to search globally is by searching for reference (link-names) in the appropriate field (attention: this is only available in global searches!). This is very fast thanks to it having its own table with the reference as the main key. This way many million documents can be searched within fractions of a second.

Authorizations also need to be taken into account. They have to be set on base of the folder (archive). In the Enterprise-version Authorizations can be set WITHIN an archive using global bookmarks. As of right now, this is not possible in the GPL-version.

Additional- and status fields also only apply to a single archive. When sending a document into a different archive, the related information also becomes obsolete/ unspecified, unless there are additional- and status fields of the same name in the target-archive.

2. Further sources of information on the internet

Descriptions of the bitfarm-Archiv interfaces:

Differences between the GPL/Enterprise-versions of bitfarm-Archiv:

Free and live online-presentations of the Enterprise-version: