

# d.veLop

d.3 search: Administrator

# Table of Contents

1. d.3 search: Administrator .....	3
1.1. Introduction .....	3
1.1.1. About d.3 search .....	3
1.1.2. Prerequisites .....	3
1.1.3. Internationalization .....	3
1.1.4. Information on d.3 search .....	3
1.1.5. Architecture .....	4
1.2. Installation .....	8
1.2.1. System requirements .....	8
1.2.2. System recommendations .....	9
1.2.3. Update .....	10
1.2.4. Installing and updating d.3 search .....	11
1.2.5. d.3 gateway - connection .....	12
1.2.6. d.3 SHAFT Agent .....	13
1.2.7. Cluster .....	15
1.2.8. Integration of d.ecs monitor .....	27
1.2.9. SSL encryption .....	27
1.3. Configuration .....	28
1.3.1. Configuration via web interface .....	28
1.3.2. Configuration of the d.3 search server .....	50
1.3.3. Configuration of a document corpus .....	54
1.4. Technical background on the d.3 search server .....	80
1.4.1. Directory structure .....	80
1.4.2. The d.3 search engine .....	82
1.5. Maintenance .....	90
1.5.1. Notes on optimizing d.3 search. ....	90
1.5.2. Notes on optimizing d.3 search on a RAID5 system .....	91
1.5.3. Clues on recovering a document-corpus .....	91
1.5.4. Reindexing a document corpus .....	94
1.6. Errors and issues .....	105
1.6.1. d.3 search .....	105
1.6.2. d.3 search engine .....	109
1.7. Additional information sources and imprint .....	113

# 1. d.3 search: Administrator

## 1.1. Introduction

This topic lists general product information and conventions applied in this product documentation.

### 1.1.1. About d.3 search

The product d.3 search extends the d.3 system with full-text search functionality.

The product consists of the two components d.3 search server and the search engine d.3 search engine (Shaft engine).

The d.3 search server works as the interface between the d.3 server and the search engine d.3 search engine.

### 1.1.2. Prerequisites

This manual is targeted at the d.3 administrators and describes the structure, the installation and the configuration of d.3 search.

Knowledge of the d.3 administration and the administration of the d.ecs rendition service are required.

### 1.1.3. Internationalization

d.3 search supports the following languages:

- German
- English
- Languages from the d.3 national package
- Japanese
- Chinese, simplified

#### Note

Please consider the notes on the [Delimitermode=4](#).

#### Note

For other possible languages, please contact the d.velop AG.

### 1.1.4. Information on d.3 search

#### Common

d.3 search extends the d.3 system with full-text search functionality.

The data to be adopted in the search engine is configured in the rules on the d.3 server administration.

The full-text information (OCR) are also adopted in the search engine d.3 search engine via the d.3 server and the d.3 search server. d.3 search supports the administration of multiple document corpora, so that rules can define which full-text information shall be included in a specific document corpus. This can be advisable to get useful results when using the semantic search.

Optionally, existing full-text information in an index of the d.3 search engine can be adopted with a script as part of a later installation.

If a user performs a full-text search at the client, the d.3 server receives the search query and forwards it to the search engine d.3 search engine (SHAFT) via the d.3 search server. Then the d.3 server receives the result of the search engine and processes it for the delivery to the user.

## Licensing

The product d.3 search is available with two licensing types.

- **d.3 search basic:** With this license, only the property values of document and folders are adopted in the index. Document-contents from the actual file are discarded during the import. Thus, a full-text search is only possible in the property values.
- **d.3 search file-content addon:** Extends the basic license with the data content. All information (properties and document contents) provided by the d.3 server are included in the index.
- **d.3 search enterprise addon:** With this license the features **clustering** and **scaling** are activated. To use d.3 search with a large number of documents and user interactions, a document corpus can be scaled to multiple d.3 search engines (SHAFT).

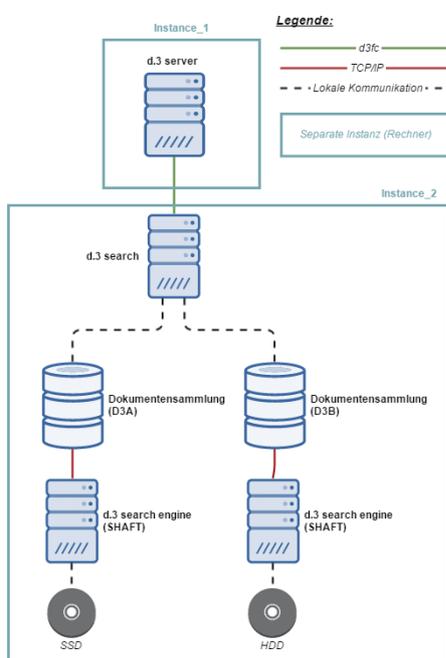
## 1.1.5. Architecture

In the following chapter the individual architecture variants of d.3 search are presented.

### Easy installation

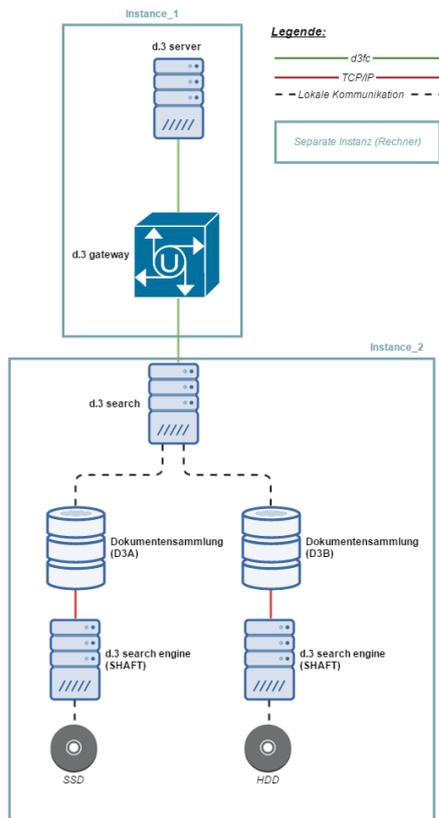
#### Easy installation without d.3 gateway

The following chart shows an easy installation of d.3 search without using the d.3 gateway. Two document corpora are created in the example. The document corpus **D3A** shall display a larger document corpus, which is stored on a SSD-hard disk for performance reasons. The document corpus **D3B** is comparative smaller, so a HDD-hard disk is used. Both document corpora will be addressed by the d.3 server.



#### Easy installation with d.3 gateway

The following chart shows an easy installation of d.3 search with using the d.3 gateway. As shown in the example above two document corpora are created. Both document corpora will be addressed by the d.3 server via the d.3 gateway.

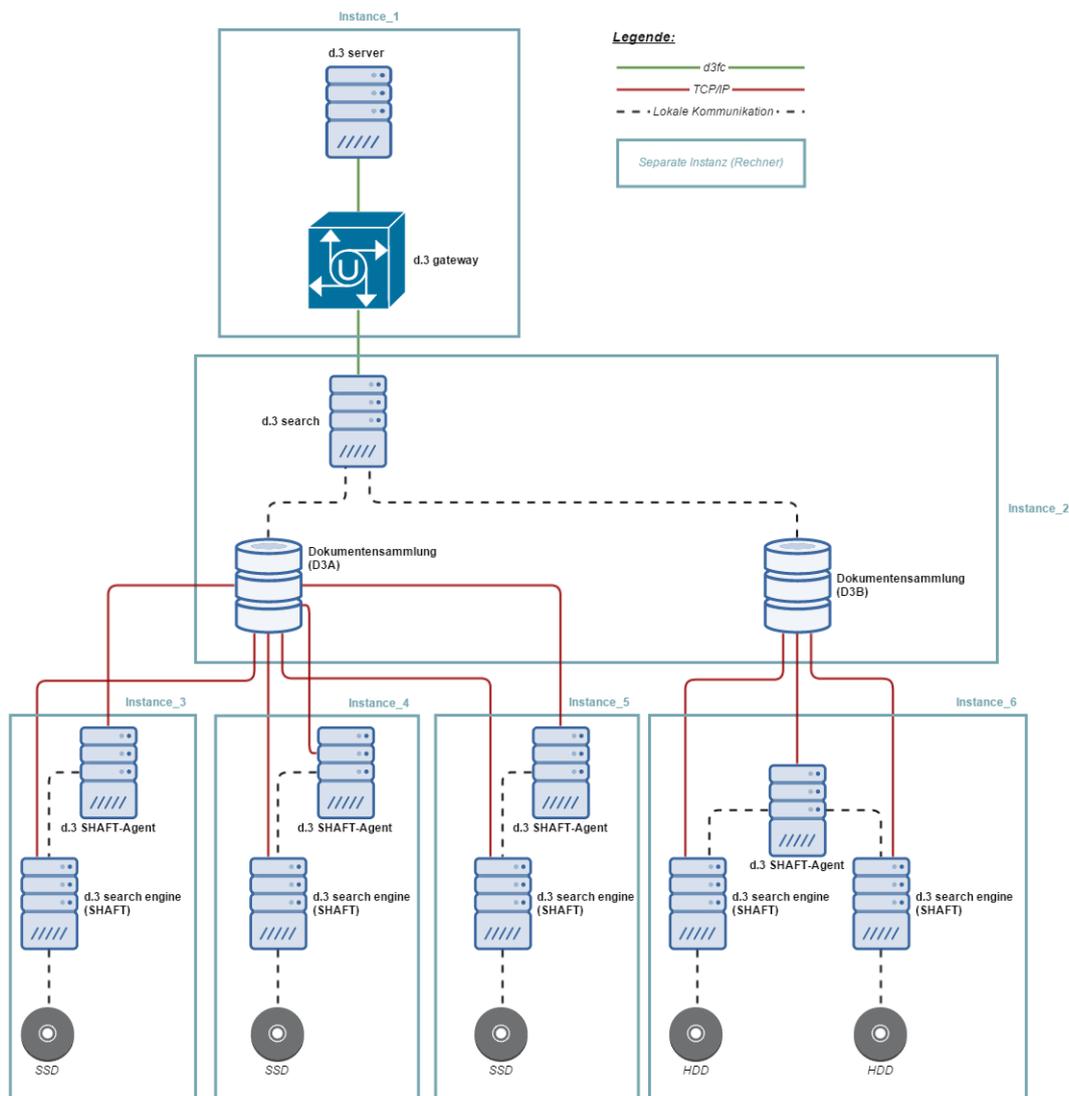


## Scaled installation

During the scaled installation for each document corpus multiple d.3 search engines (SHAFT) will be created. To raise the performance, these can be written to separate instances. To check the processes to the separate instances, the d.3 SHAFT agent will be installed here.

In this example the document corpus D3A is separated to three instances and to three d.3 search engines (SHAFT). A possible scenario could also be a separation to six or more d.3 search engines (SHAFT) but have them installed on a instance previously. This would allow a later scaling or raising the performance. The document corpus D3B was separated to two d.3 search engines (SHAFT). But both processes are located on one instance because the document corpus is smaller than D3A and thus no further scaling is required.

The scaling has the advantage that a large document corpus will be separated to multiple d.3 search engines (SHAFT). This leads to faster search and import procedures. In addition to that separate hard disks can be used with the scaling, which also raises the performance.



The architecture of the scaled installation was displayed here with the d.3 gateway. A direct communication between d.3 server and d.3 search would also be a possible scenario.

## Cluster installation

During the cluster installation multiple, in this case three, redundant d.3 search instances will be installed. All instances communicate via the d.3 gateway with the d.3 server. The d.3 search instances are monitoring themselves to react on a failure. As for the previous examples the document corpora **D3A** and **D3B** are used again.

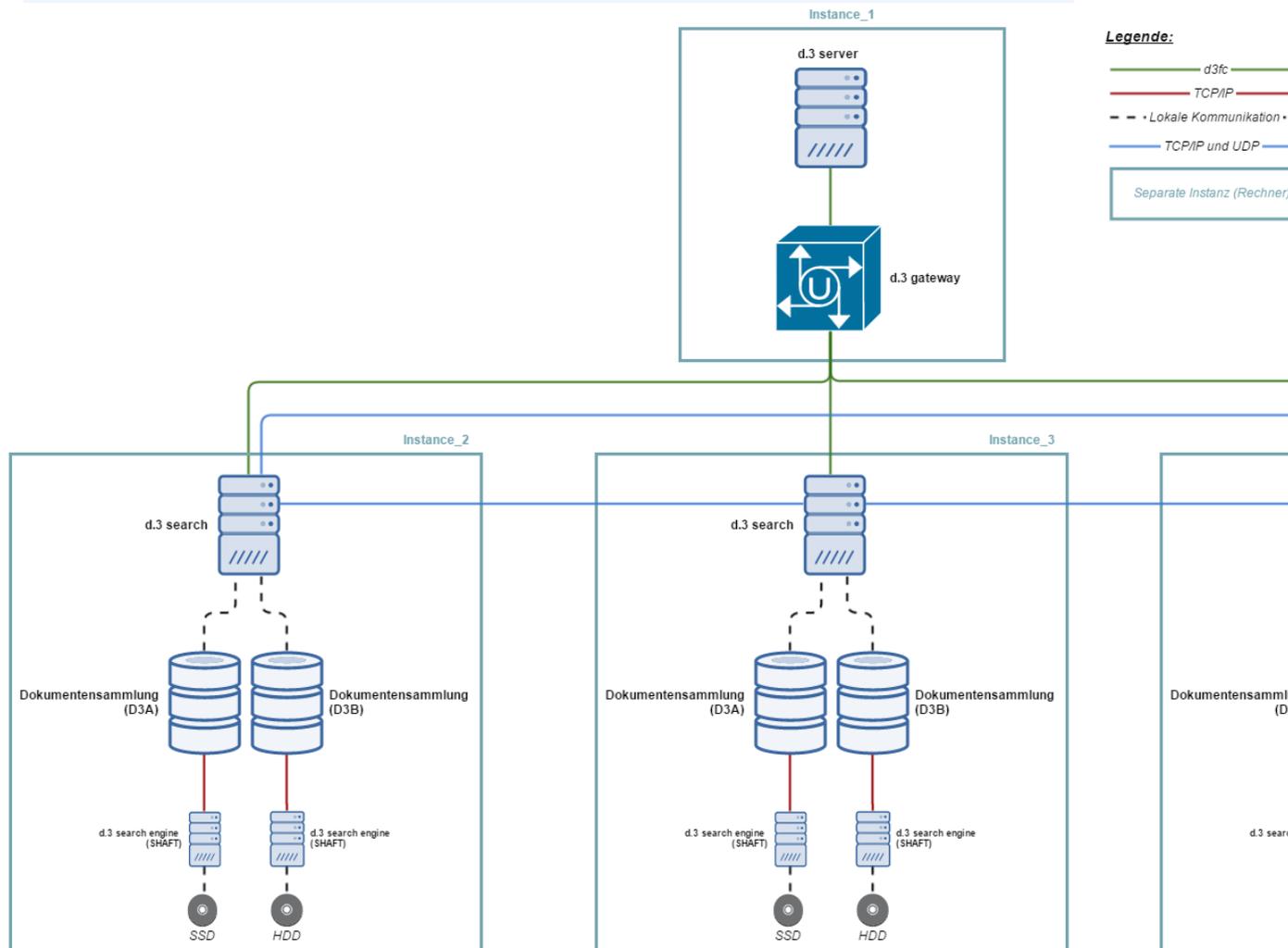
A cluster operation has the advantage that queried tasks can be processed furthermore as long as there is an active instance on which d.3 search is installed. Adding or changing documents in the search index is not possible during a failure because inconsistent conditions could be created in the data.

### Note

If instances two and four fail, instance three can respond furthermore to the search queries of the d.3 server. Document imports by users would not be executed during this.

## Note

For the configuration of a cluster it is required that the communication takes place from the d.3 server to d.3 search via the d.3 gateway. If this is not considered, a failover security is not ensured.



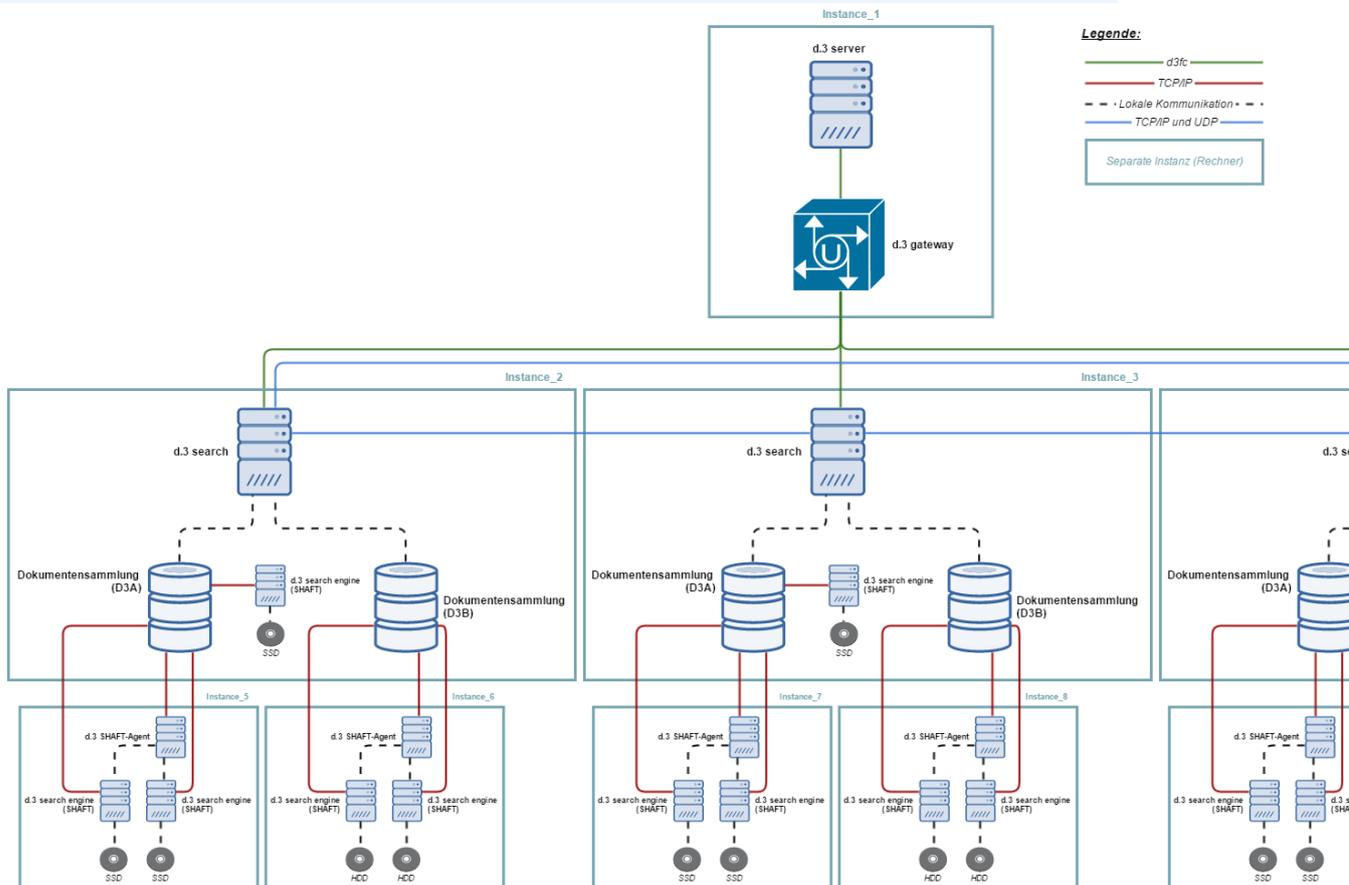
## Scaled and cluster installation

During the scaled cluster installation multiple, in this case three, redundant d.3 search instances will be installed. Each d.3 search instance holds scaled document corpora. For the document corpus D3A a d.3 search engine (SHAFT) was installed on the same instance as d.3 search. The other two d.3 search engines (SHAFT) were configured by the d.3 SHAFT agent on a separate instance. The d.3 search engines (SHAFT) of the document corpus D3B are located on a separate instance. All d.3 search instances communicate via the d.3 gateway with the d.3 server. The d.3 search instances are monitoring themselves to react on a failure.

In this case the search query will be distributed by the d.3 gateway to the d.3 search instances. The document corpus separates to the respective available d.3 search engines (SHAFT). This model provides a optimal failover security as well as an improved performance.

## Note

For the configuration of a cluster it is required that the communication takes place from the d.3 server to d.3 search via the d.3 gateway. If this is not considered, a failover security is not ensured.



## 1.2. Installation

The installation of d.3 search configures the connection between the d.3 server and the d.3 search server to the search engine which is actually managing the full-text information. To enable the process chain you must install the d.3 search server, configure a new document corpus for a d.3 repository and enable the d.3 search support on the d.3 server.

Please consider the additional notes on the operation of d.3 search in the cluster.

A description of the individual steps can be found in the following chapters.

### 1.2.1. System requirements

Please refer to the central [system requirements for d.velop products \(on-premises\)](#). You can find deviating or more extensive system requirements in this documentation.

#### d.3 search engine

Every installation of a d.3 search engine instance requires at least 256 MB available physical main memory. The amount of required space depends on the text volume as the index is approx. 1.4 times larger than the text and requires according disk space. To ensure a good performance, the engine requires as much cache as possible and a fast hard disk. If possible, the cache should amount to 1/4 of the size of the index. The optimum is a cache reaching a match rate of 98%. For the hard disk, the IOPS value with random access to 4 KB blocks ist relevant. If this value lies between 150 and 200, the system is

sufficient for up to 1 million documents. When selecting the disk system for corpus, the required IOPS value should be calculated with the following formula: Minimum IOPS value = Million documents x 200. For 50 million documents, you would thus need a disk system at least reaching 10.000 IOPS.

## 1.2.2. System recommendations

### Introduction

In the section below, system recommendations for the configuration of an environment for d.3 search are provided.

#### Note

Generally speaking a concrete recommendation can not be provided in a documentation because of the individual specifications (number of search queries / number of documents / document size etc.).

This chapter provides information on which things you have to consider and which sizes can be expected. A concrete planning can only be viewed in a specific case.

In general, the resources can be prioritized descending as follows:

1. **Hard disk:** For search engines most of the time is required for the query and writing the files to the hard disk. Thus a dedicated hard disk should be used which is only used by d.3 search in the best case. To be as performant as possible, a SSD hard disk should be used.
2. **Main memory:** To span the hard disk speed the cache of the d.3 search engine (SHAFT) must be at least 1/3 of the data base size.
3. **CPU:** The CPU will not be occupied significantly by the d.3 search. The processor should be up to date and must contain at least 4 kernels.

During the setup of d.3 search you should also consider which directories have need of disk space. Hard disc capacity is required by the following storage directories:

1. **Installation directory** The full installation of d.3 search requires a disk space of ~116MB. This includes all program files, as well as the documentation and the uninstaller.
2. **Database directory:** The database directory contains all files of the database and depends on the following factors:
  - a. **Number of documents:** The number and size of the respective documents
  - b. **Content of the documents: Amount of OCR information of a document**
  - c. **Additional storage of full text content:** If you additionally save the full text contents to enable functions such as phrase search, semantic search, text summary, this requires additional hard disk space
  - d. **Transaction logs:** Saving the transactions in a log is used when restoring the index if the process has been terminated unexpectedly. The size of the log depends on the number of transactions between full backup times.
  - e. **Archive logs:** With the help of the archive log, an index can be rebuilt more quickly. The size depends on the number of transactions between full backup times.
  - f. **Automatic repair:** When automatically repairing an index, the size of the current index is additionally required.
  - g. **Automatic defragmentation:** For the defragmentation of an index an additional storage space in size of the files `link.dat` and `words.dat` is needed.
3. **Backup directory:** The backup directory requires the space that the search index needs at the time of the backup.
4. **Job Directory:** The size of the job directory depends on the number of documents in the queue. This folder contains OCR information as well as information about the jobs

**Note**

The following chapter shows a sample installation of d.3 search.

The example contains mostly commercial documents.

The following properties are given:

Property	Value
Number of documents in the index	2.015.988
Number of different words in the index	51.049.071
Index size	16.3 GB
Of which size of additional storage of full-text content	4.4 GB
Number of active users	~ 200 employees

**Recommended architecture**

There are two orientations in the selection of architecture. One is security, which can be achieved through an active/active cluster, and the other is scaling.

A d.3 search system should be operating in a cluster array, if you want to raise the availability. In addition, a cluster has the advantage of distributing search queries evenly across all instances in the cluster, which provides some degree of load balancing. However, the main aspect of a cluster should be the availability of full-text search.

Scaling of a document collection is required at the latest when an index is larger than 50 GBytes. With the respective hardware an index with 50 GB is the highest recommended size for a document corpus by d.velop. Extrapolated from the above example, this corresponds to approx. 6 million documents.

When creating a document collection, you should already think about the possible number of document in the near future. It is recommended not to let an index become larger than 20 GByte in order to achieve an appropriate performance with current hardware. Since the later conversion to scaling takes some time (several days), you should rather set the scaling higher than needed already when creating a document collection.

**Note**

They have an annual document volume of 2 million documents. Within 10 years, your index will be about 163 GB, based on the above example. To avoid changing the scaling in the next 10 years, it is recommended to set the scaling to 8 d.3 search engines (SHAFT).

**1.2.3. Update**

If you are using d.3 search in a version lower than 2.1, please note the following information:

The structure of the job-files has been changed with d.3 search version 2.1. The new d.3 search server can no longer interpret and process old job-files.

**Warning**

Before you install an update, the job-directory of the d.3 search server must therefore be fully processed since there is no way of migrating the old jobs to the new version.

**Update for existing cluster**Active/Active cluster

If an active/active cluster is already configured with d.3 search, follow the next steps during an update:

- Shut down a d.3 search instance via the d.3 process manager.
- Update the selected system.
- Start the d.3 search instance via the d.3 process manager.
- Repeat the steps above for all other d.3 search instances again.

During the update the search function of d.3 search can still be used. Only changes to the data basis, to the document corpora will not be processed.

#### Active/passive cluster

If an active/passive cluster is already configured with d.3 search, follow the next steps during an update:

- Close the d.3 process manager of the currently **passive** instance.
- Update the selected system.
- Start the d.3 process manager.
- Close the d.3 process manager of the currently **active** instance - the passive instance will take over the active role after a few moments.
- Update the selected system.
- Start the d.3 process manager.

During the update all functions of d.3 search can still be used.

### 1.2.4. Installing and updating d.3 search

You install the software exclusively using d.velop software manager. If an application is required for different products, the corresponding software packages are also installed automatically.

For further information on installing the software, see the d.velop software manager manual.

For the first installation, you need to run the setup (**setup-dvelopSoftwareManager.exe**) from the d.velop service portal to install a product. Once you have installed the d.velop software manager on your server, you can make further changes by starting the application on the server via **d.velop > d.velop software manager** in the Start menu.

#### **Note**

During installation, d.velop software manager always adjusts the following system settings on the server:

- Adjust the IPv4 port range for the TCP protocol from 10,000 to 55,000.
- Adjust the IPv4 port range for the UDP protocol from 10,000 to 55,000.
- Adjust the IPv6 port range for the TCP protocol from 10,000 to 55,000.
- Adjust the IPv6 port range for the UDP protocol from 10,000 to 55,000.
- Adjust the registration key **HKEY\_LOCAL\_MACHINE\SYSTEM\CurrentControlSet\Services\Tcpip\Parameters\TcpTimedWaitDelay** to the value **30 (0x1e)**.

#### **This is how it works**

1. Select a feed.
2. Enter the address of the d.ecs license server if the license server has not been found. You can also install the d.ecs license server directly, but you'll need to restart d.velop software manager.
3. Choose **Install**.
4. Select the products you want to install.
5. Optionally, you can enable the settings for automatic updates.
6. Check the summary and click **Change**.

You can only update the software using d.velop software manager.

For further information on updates, see the d.velop software manager manual.

### 1.2.5. d.3 gateway - connection

#### Introduction

From version 3.0.0 of d.3 search the product can be addressed via the d.3 gateway. The d.3 gateway distributes the requests as the Round-Robin procedure to the signed in d.3 search instances.

As soon as a d.3 search instance fails, it is recognized by the d.3 gateway and sends only incoming requests to the available d.3 search instances. As soon as the d.3 search instance is available again, d.3 gateway sends requests to this instance again. This improves the failover security by the d.3 gateway.

#### Settings in d.3 search

To let d.3 search log on to the d.3 gateway, you have to make some settings in the web interface of d.3 search.

- To do so, open the web interface, which is accessible by default under the address <http://localhost:6133>. You can open the link d.3 search Webinterface from the d.3 search start menu.
- Now open the page **Settings | Global**.
- On the right side you configure the following three configuration parameters, which are located in the section **Gateway** (also see chapter Configuration of the d.3 search server).

<b>HandlerCount</b>	Specifies the number of connections to the d.3 gateway to be established.
<b>Host</b>	Defines the host name of the d.3 gateway.
<b>Port</b>	Specifies the port via which the d.3 gateway is available.

The following screenshot shows the settings to be made in the web interface:

#### Gateway

<b>G HandlerCount</b>	<input type="text" value="6"/>	
<b>G Host</b>	<input type="text" value="d3-sql-712"/>	
<b>G Port</b>	<input type="text" value="3400"/>	

- Close the settings with a click on the button **Save**.

For the settings to take effect, you must restart the d.3 search server. To use the connection via d.3 gateway, enter d.3 gateway in the settings of d.3 server (d.3 admin).

#### Note

The d.3 search server is still available via the usual port 3406. The connection to the d.3 gateway is additionally displayed.

## Settings in d.3 server (d.3 admin)

To send requests from the d.3 server via the d.3 gateway, the following settings must be applied:

- Start the d.3 admin and open the option **System settings | d.3 config | Full-text-search (d.3 search)**.
- Now, enter the host name of the **d.3 gateway** under the option d.3 search server.
- Afterwards, under **TCP/IP port for d.3 search** enter the port via which the d.3 gateway is accessible. (default: 3400)

After the settings were saved, d.3 search is accessible via d.3 gateway.

### 1.2.6. d.3 SHAFT Agent

#### Introduction

If your current d.3 search instance is busy with regards to CPU and RAM, with the d.3 SHAFT Agent you have the option to outsource the d.3 search engine (SHAFT) on a remote computer. By doing this, resources of the current instance will not be that busy and this leads to a faster processing of requests.

The CPU utilization will be reduced because the requests are processed by the d.3 search engine (SHAFT) on another computer and the used RAM is decreased because the d.3 search engine (SHAFT) holds the search index in the RAM on the separate computer. Thus, the relocation of document corpora to a separate computer leads to lower requirements to the CPU and RAM on the d.3 search instance.

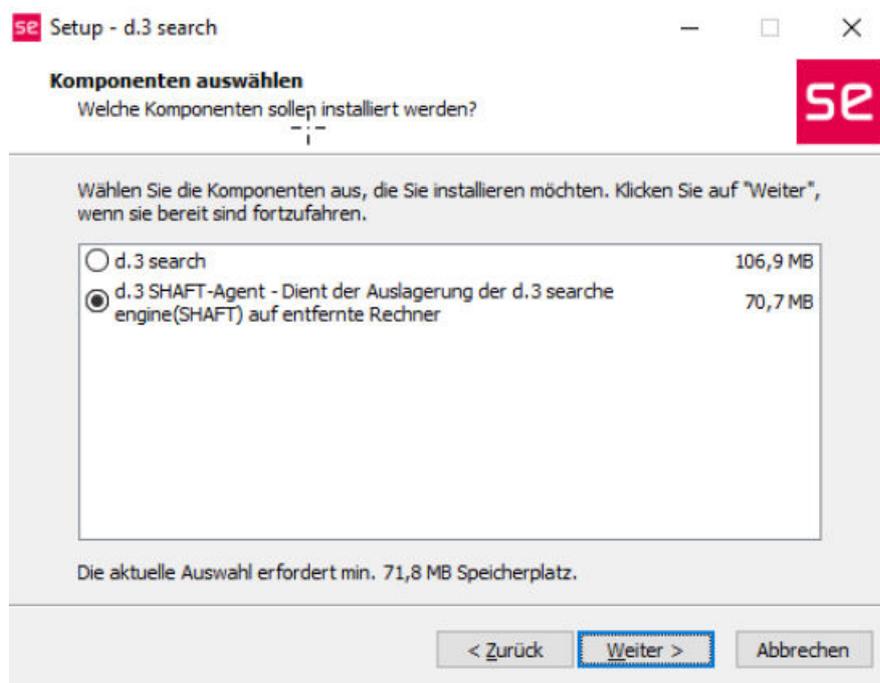
The d.3 SHAFT Agent is used as an interface between d.3 search engine (SHAFT) and the d.3 search server for the outsourcing and the remote control of the processes and data transfer of the search index is taken over by it.

An outsourcing of the d.3 search engine (shaft) should be taken in consideration in the following situations:

- the CPU utilization of the d.3 search instance is permanently over 60%
- there is almost no free RAM available and the PerfLogs inform about the cache ratio is under 80%.

#### Installation and configuration

- Start the setup of d.3 search to prepare the outsourcing of a d.3 search engine (SHAFT).
- Mark the option **d.3 SHAFT Agent** on the side **Select components**.
- Finish the installation the usual way.



- Make sure that the d.3 SHAFT Agent is entered correctly in the d.3 process manager and that it is started.
- Open the web interface on the computer, whose d.3 search engine (SHAFT) shall be outsourced. This can be called by default with the address <http://localhost:6133/> or you can open the link d.3 search Webinterface from the d.3 search start menu.
- There, select the menu option **Settings | Document corpora | <CorpusId> | corpus.ini**, to edit the configuration of the document corpus.
- In the last step you change the specifications in the section **SHAFT1** to the values you want and save these afterwards.

### Note

During the secondary storage, the entire content of the database directory of the d.3 search engine (SHAFT) is transmitted. This also includes the backup and the transaction- and archive-logs. You can accelerate the copying by performing a manual backup before the secondary storage. This way, the transaction- and archive-logs are removed and do not have to be transmitted.

The downtime can be minimized by manually moving or removing the backup. For this effect, move the backup directory outside the database directory of the d.3 search engine (SHAFT). After the secondary storage allows to move the backup-directory manually to a new machine, or to create a new backup removing the old one.

Having changed the settings you will be prompted to restart d.3 search.

The start of the software can take some minutes, depending on the document corpus. As soon as the document corpus is sent to the d.3 SHAFT Agent, d.3 search is fully functional again.

The current status of the send process can be found in the d.3 log.

## Configuration

The `SHAFTAgent.ini` allows to make settings for the d.3 SHAFT Agent.

Section	Entry	Meaning
[Global]	Corpora-Path	Default: \corpora Specifies the directory, under which the corpora are stored.
	SHAFTPath	Default (64-bit operating systems): shaft\64Bit Default (32-bit operating systems): shaft This specifies the path where the <code>shaft.exe</code> is saved. The selected <code>shaft.exe</code> is used for all document corpora. It is impossible to define the <code>shaft.exe</code> for each document corpus.
	TCPPort	Default: 6133 Specifies the TCP port under which the d.3 SHAFT Agent is accessible.
	LogLevel	Default: 6 Specifies the LogLevel for the d.3 SHAFT Agent. LogLevel=0: only errors are logged in the log viewer LogLevel=6: normal logging of d.3 search API calls LogLevel=9: all messages (including debug entries)

### 1.2.7. Cluster

From version 3.0.0 of the product d.3 search two different type of clusters are supported. Cluster provide the possibility to improve the failover security and the scalability of a system.

Failover security means for d.3 search to ensure that the search functionality is running.

There a two alternatives to build a cluster:

Active/ Active cluster	<p>This build of a cluster provides failover security for the cancellation of a process/computer as well as for errors in the data basis, the document corpora.</p> <p>The user can still use the search function during such a fail. Only the changes to the data will not be processed until all members of the cluster are active again.</p> <p>For an active/active cluster search queries will be distributed to the respective members. As a result, you get a scaling for the processing of search queries.</p> <p>An active/active cluster contains any number of members.</p> <p>For additional information see <a href="#">Active/Active cluster</a></p>
Active/ passive cluster	<p>With an active/passive cluster you achieve a basic failover security, if a computer or the d.3 process manager is no longer available. The data basis can still be changed during a fail of an instance.</p> <p>An automatic pan does not take place, if only the d.3 search process is not available. Moreover this build provides no failover security, if the data basis is defect.</p> <p>The cluster type contains always exact two members.</p> <p>For additional information see <a href="#">Active/passive cluster</a></p>

#### Warning

As errors in the data records in an active/passive cluster may lead to a system standstill, it is recommended to use an active/active cluster.

To build a cluster, you can use an already existing d.3 search instance and prepare it with only a few steps for the cluster.

Before adding a second instance into the cluster, all configuration on the first instance should be finished, so that the settings of the new instance will be immediately active.

#### Active/active cluster

By using an active/active cluster a failover security can be provided for the d.3 search system. The build of a d.3 search cluster has also the following advantages:

- A cluster distributes search queries equally to all involved machines; changes on the data, the document corpora, will be processed synchronously by each d.3 search instance.
- If a member of this cluster fails, the cluster gets inactive - this avoids changes to a data store until the cluster is build up again.
- The search can still be used during this failure, so that the functionality of the clients is not hindered.

For this, a member of the cluster is applied as failed under two conditions:

1. The d.3 search engine (SHAFT) is not started.
2. An instance is not available - the master expects feedback every 5 seconds from the slave. As soon as the feedback is overdue for 10 seconds, the instance will be marked as unavailable.

Members of a cluster are always in one of the following modes:

<b>Master</b>	There is always a master in the operation of a cluster. This one checks the availability and the status of the cluster and synchronizes the settings with all members. If the master machine fails, the cluster assigns automatically the master mode to another cluster member. In the web interface settings can only be done via the web interface of the master. For all others the save-button is disabled and an alert appears.
<b>Slave</b>	All functional members of a cluster not being in master mode are noted as slave. A slave is a normal functional d.3 search instance. The web interface can also be opened for local information on the slave. But for changes on the setting it must be switched to the web interface of the master.
<b>Unknown</b>	A member with the mode unknown is shut down currently or is not responding due to other reasons. As soon as a member gets this status, the cluster is inactive and no changes to the data basis will be done. This means, all jobs of the d.3 async will be answered with the error code 4100011 and must be restarted afterwards by the d.3 async job inspector manually.
<b>Is synchronized</b>	Machines are in this mode for two conditions:  In the first case the cluster function was just activated or the instance is starting now. As soon as the machine is in this mode, it is searching for a possible cluster in the network. This procedure takes max. 15 seconds.  In the second case a member switches into this mode, if it will be synchronized/added by a master in the cluster. By doing this, all configurations and data will be transmitted.

The respective current modes of the machines can be viewed at any time via the d.3 search version 3.0.0 web interface under the menu option **Cluster**. This is accessible by default under the address `http://localhost:6133/`.

### Configure existing d.3 search instance for the cluster

To configure an existing d.3 search instance for a cluster, follow the steps below:

- Open the d.3 search web interface on your existing d.3 search instance. This can be found by default under the address `http://localhost:6133` or you can open the link d.3 search webinterface from the d.3 search start menu.

#### Note

The login data for the web interface must be the same for all cluster members.

- Make sure that the d.3 gateway connection is configured (see chapter [d.3 gateway - connection](#)).

#### Warning

Enter the host name of the computer for the gateway host and never `localhost`!

- Enable the cluster in the web interface under **Settings | Global**.
- In the section **Cluster** you enable the checkbox for **Active**. Specify a group name here. Computer building a cluster together, must have the same group name. Also the port must be identical for every computer of the cluster.

For additional information see chapter [Configuration of the d.3 search server](#).

After this step, d.3 search must be restarted. This is also needed by the web interface.

### Note

The group name should be unique. Computers with activated cluster function and the same group name can find themselves automatically via an UDP-broadcast and can add themselves automatically to an existing cluster. Though, all search indexes will be transmitted, which can contain sensible data.

## Cluster

<b>L</b> Active	<input checked="" type="checkbox"/>	
<b>G</b> Group	<input type="text" value="Testgroup"/>	
<b>G</b> Port	<input type="text" value="6134"/>	

After the activation of the cluster, you can find in the web interface the menu option **cluster**.

- Call the cluster information with this option. At this time in the section **cluster-server** your local computer should be listed with the mode master.

## Einstellungen

Automatische Aktualisierung

## Cluster-Server

IP-Adresse	Rechnername	Modus	Letzte Meldung	Information	Fehlermeldungen
10.3.4.91	wathe01	Slave	16.02.2016 11:26:25		
10.95.0.80	d3-sql-712	Master	16.02.2016 11:26:25		



If the **cluster information** is displayed correctly, you have configured successfully your existing instance cluster-compatible.

Now, adding one or more new instances is missing (see chapter [Configure new d.3 search instance for the cluster](#)).

**Warning**

The configuration and start of the new instance must be done implicitly after the configuration and start of the existing d.3 search instance.

The order, in which the server will be added to the cluster, decides for the roles taking the instances.

**Configure new d.3 search instance for the cluster**

For each new instance the procedure is the same and is similar to the configuration of the existing instance.

- Open the d.3 search web interface on the new instance under the address `http://localhost:6133` or open the application d.3 search web interface from the d.3 search start menu.

**Note**

The login data for the web interface must be the same for all cluster members.

- Now, choose the menu option **Settings | Global**.
- Activate the cluster functionality by confirming the option **Active** in the section **Cluster**.
- The, specify the group name of the cluster.
- Check, if the port is identical with the existing instance.
- In the last step, check if the cluster computers have been already connected automatically.
- To do so, open the menu option **Cluster**. In the section **Cluster-server**, the new instance with the mode **Is synchronized** should be listed.

## Cluster-Server

IP-Adresse	Rechnername	Modus	Letzte Meldung	Information	Fehlermeldungen
10.95.4.231	dsearcht2	Master	11.05.2017 11:21:48		
10.95.5.146	dsearcht1	Slave	11.05.2017 11:21:48		
10.95.2.225	dsearchdc1	Slave	11.05.2017 11:21:48		• 11.05.2017 11:21:48

In the first 15 seconds after starting the new instance, these tries to find an already existing, configured d.3 search instance and to start the exchange of the configurations and document corpora automatically.

**Warning**

For these automatic synchronization it is mandatory to be able to send an UDP-broadcast between the new instance and the existing d.3 search instance.

If this is impossible, then the second instance must be added manually. You can do this, if you open the web interface of your existing instance and confirm the button **Add** under the menu option **Cluster** in the section **Cluster server**. Now, enter either the IP address or the host name of the new instance and confirm it with the **<Return>**-key.

## Instructions

### Failure

A d.3 search instance is regarded as failed:

- if it cannot be addressed via the TCP port of the cluster.
- if one or multiple d.3 search engines (SHAFT) are not started.

You can recognize a failure by the following criteria:

- The d.3 object monitor reports:

```
The Cluster [<Cluster Gruppennamen>] is offline. Error servers: <IP-ADDRESS> (<HOSTNAME>)
```

- The d.3 log contains one of the following messages:

```

21.08 09:49:56,519 DSEARCH 260C2F20 > The Cluster [SearchCluster] is offline. Error servers: d3-re
21.08 09:49:56,519 DSEARCH 260C2F20 > n-sto-712 (d3-ren-sto-712)
21.08 09:51:15,857 D3Server 00000FA0 d.search returned an error (4100011)
21.08 09:51:15,860 D3Server 00000FA0 Error on adding OCR data for 'P000000147' to d.search (4100011)

```

- In the web interface of d.3 search under **Home** the option for the cluster is set to **Active = false**:

### Cluster information

<b>Activ</b>	false
<b>Cluster name</b>	SearchCluster
<b>Server count</b>	2
<b>Status</b>	Is_Master

- In the web interface of d.3 search under **Cluster** the log contains the following messages:

```

21.08.2014 10:45:20: [SEER]: d3-ren-sto-712 - Timeout (IAMALIVEMASTER)
21.08.2014 10:45:32: [SEER]: d3-ren-sto-712 - Timeout (IAMALIVEMASTER)

```

For a failure of the master a new master will be determined automatically in the cluster, who is in charge for changes to the configuration afterwards.

If an instance fails, no jobs from d.3 search can be processed. This means, in this time no changes to the data basis, the document corpora can be done.

If an instance fails for a longer time (several days), it is recommended to remove the instance from the cluster to prevent an accumulation of jobs. More details can be found in chapter [Remove cluster member](#).

As soon as the cluster is inactive, the administrator should get the instance to work again quickly. More details can be found in the chapter [Active/Active cluster is inactive](#).

If you have removed the defective member from the cluster (see chapter [Remove cluster member](#)) and you want to add it to the cluster again, do it as described in the chapter [Configure new d.3 search instance for the cluster](#).

As soon as all cluster members are active and functional again, the cluster will be active again and the jobs will be processed.

You can recognize the change of status by the following messages:

- The d.3 object monitor reports:

```
d.3 search is active since <Datum>. <Anzahl Calls> calls handled. Pending
jobs(<Anzahl Add-Jobs> add, <Anzahl Update-Jobs> update, <Anzahl Delete-
Jobs> delete)
```

- You will find this message in the d.3 log:

```
[i] 21.08 10:49:47,761 DSEARCH 34682F20 > Cluster-Status changed: False ->True (One cluster member is
[i] 21.08 10:49:47,763 DSEARCH 34682F20 > not active)
[i] 21.08 10:49:47,763 DSEARCH 34682F20 The Cluster is Online
```

- In the web interface of d.3 search under **Home** the option for the cluster is set to **Active = true**:

## Cluster Informationen

<b>Aktiv</b>	true
<b>Cluster-Name</b>	SearchCluster
<b>Anzahl Server</b>	2
<b>Status</b>	Is_Master

- In the web interface of d.3 search under **Cluster** the log contains the following messages:

```
21.08.2014 11:34:31: [SEOK]: d3-ren-sto-712 - IAMALIVEMASTER
21.08.2014 11:34:38: [SEOK]: d3-ren-sto-712 - IAMALIVEMASTER
```

- Changes to the data basis, the document corpora will be done again - jobs from d.3 server will be processed successfully.

The failed jobs of the d.3 async can be restarted via the d.3 async job inspector in the d.3 admin.

To do so, follow the steps below:

- Start the d.3 admin.
- Start the d.3 admin and open the option **d.3 components | d.3 async job inspector**
- Enter the following values in the search criteria:

Suchvorgaben:		
Job-Nummer:	<input type="text"/>	Status-ID: <input type="text"/>
Fehler:	<input type="text" value="4100011"/>	
Besitzer:	<input type="text"/>	Job-Typ-ID: OCR001 - (Verschlagwortung)
		Dokument-ID: <input type="text"/>

- Select **Search**.
- Mark all found jobs.
- Then choose the button **Send again**.

## Remove cluster member

If you want to remove a member from the cluster, follow the steps below:

- Stop the d.3 search process to be removed from the cluster via the d.3 process manager.
- Open the web interface of the d.3 search instance holding currently the master mode. This can be found by default under the address `http://localhost:6133/` or you can open the link d.3 search web interface from the d.3 search start menu.
- Navigate to the option **Cluster**.

- Select **Delete** for the respective instance. You can only remove stopped instances.

#### Cluster-Server

IP-Adresse	Rechnername	Modus	Letzte Meldung	Information
10.95.0.80	D3-SQL-712	Master	21.08.2014 11:47:24	<a href="#">Löschen</a>
10.95.0.79	d3-ren-sto-712	Unbekannt		<a href="#">Löschen</a>

- Uninstall the d.3 search instance you have shut down via **Software | Programs and Features | d.3 search**.

### Warning

The instance removed from the cluster must not be restarted again because an inconsistent status would be existent.

Thus, always delete the respective instance and perform an uninstallation.

## Active/passive cluster

Create an active/passive cluster for d.3 search to ensure the failover security for software updates and computer crashes. An active/passive cluster always contains of two d.3 search instances, which check each other with the help of two d.3 process manager and a central network share on which the data is stored. If one of the systems is shut down because of updates, the other instance takes over automatically.

Both d.3 search instances access to the same data and thus, must have a common network share.

It is important for this, that always only one instance is running at the same time.

### Note

The active/passive cluster is a basic type of the cluster. An active/active cluster provides much more stability and functions. If you plan to build up a cluster, we recommend using the active/active structure.

## Setup

The setup assumes two computers, a primary and a secondary one with installed d.3 search. Both instances must be integrated respectively in a d.3 process manager. The document corpus is in best case on an external data volume and must be accessible by both instances.

### Note

The secondary computer can have less resources (RAM, CPU, hard disk capacity) because it is only needed, if the primary system fails.

The following steps must be executed for the setup:

- Close both d.3 search instances.
- Create a new folder, e.g. `C:\d3\corpora\` on the data volume which shall hold the data.
- If document corpora already exist, copy these into the just created folder.
- Delete the existing corpora-directory on both d.3 search computers to create in the next step the network link.
- Now create for both d.3 search computers a link to the network path. For this effect switch to one of both d.3 search computers and open the console. Now, use the command:

## Syntax

```
mklink /d "<LOKALER PFAD>" "<NETZWERKPFAD>"
```

### Example

```
mklink /d "C:\d3\d.3 search\corpora\" "\\D3-CORPORA\C$\d3\corpora\"
```

### Note

The user, who will execute the process d.3 search must have access rights to the network share.

- Now open the settings (**Settings**) of the d.3 process manager on both d.3 search instances consecutively and enable the respective **Enable failover support**.
- Specify the respective other computer and define an identical password for **Remote access password**.

Settings

Enable failover support

Host:

Port:

Remote access user:

Remote access password:

d.velop AG d.3 process manager 4.3.1 (196) © 2001-2013 | D3-CLIENT-712

- Define a secondary d.3 search instance, if you open the d.3 process manager on one of the both machines and configure the settings of the instance with a click on **d.3 search (Server)**.
- Then, select in the section **Settings** the option **Only execute if the secondary system is down**.

Settings

Group:

Caption:

Target:

Working path:

Start delay:  sec.

Min Runtime:  sec.

Execute:

Priority:

Execution times:  Su  Mo  Tu  We  Th  Fr  Sa 00:00:00 - 00:00:00

Su  Mo  Tu  We  Th  Fr  Sa 00:00:00 - 00:00:00

Su  Mo  Tu  We  Th  Fr  Sa 00:00:00 - 00:00:00

Failover:  Only execute if the secondary system is down.

Having saved these settings, a monitor icon appears in front of the d.3 search process entry.



Now you can start the d.3 search process on the primary d.3 search instance.

You have build a functional active/passive cluster.

## Instructions

### Cluster pan

A cluster pan can be applied manually or will be applied automatically by the d.3 process manager, if the second d.3 process manager is no longer accessible in the network.

If you want to update a d.3 search instance, you should perorm a manual cluster pan.

Close the d.3 process manager on the instance you want to update.

For additional information on the update of d.3 search in an active/passive cluster see chapter [Update](#).

A cluster pan will be done automatically, if:

- the d.3 process manager is not accessible via the defined port: for additional information see chapter [Setup](#).
- the d.3 process manager is shut down
- the computer of the d.3 search instance is shut down

An automatic cluster pan means, that there was some form of failure.

More details on how to behave in case of a failure can be found in the chapter [Failure](#).

### Failure

A failure of an active/passive cluster occurs, if the two d.3 process manager instances, in which the d.3 search server processes are entered, can no longer communicate with each other. You can recognize a failure with the following criteria:

- an unexpected automatic cluster pan took place.
- the d.3 object monitor reports that a d.3 process manager is no longer accessible; to configure the d.3 process manager in den d.3 object monitor, look at the manual of d.3 object monitor.
- the respective computer is not started
- the respective d.3 process manager is not started

If a failure occurred, you have to make sure that the d.3 process manager can communicate again. To do so, take care on the following possible error sources:

- The computer on which the d.3 process manager is running, is not started.
- The d.3 process manager is not started on the computer. It can either be started as a service or as a process.
- The host/port, which was configured in the d.3 process manager under settings (**Settings**) is not available externally.

If an instance fails, the secondary system takes all functionality. This means, the user can still search and changes to the data basis can also still be done.

If the faulty instance can be reached by the other d.3 process manager, a cluster pan will be done automatically. This way the secondary system is only enabled, if the primary system is not available.

## Configuration examples

### Two-node-cluster

In the following, the individual configuration examples for a two-node-cluster are listed.

On both nodes the configurations are identical:

<b>Cluster</b>	<input checked="" type="checkbox"/> Active	<input checked="" type="checkbox"/>	
	<input type="text"/> Group	Clustergroup	
	<input type="text"/> Port	6134	
<b>Gateway</b>	<input type="text"/> HandlerCount	6	
	<input type="text"/> Host	d3-sql-712	
	<input type="text"/> Port	3400	
<b>Communication</b>	<input type="text"/> SSLCertFile	server.crt	
	<input type="text"/> SSLKeyFile	server.key	
	<input type="text"/> SSLPassword		
<b>General</b>	<input type="text"/> LogFile		
	<input type="text"/> LogLevel	9	
	<input type="text"/> MinRelevance	0	
	<input type="text"/> OMOMsgDist	10	
	<input type="text"/> Port	3406	
	<input checked="" type="checkbox"/> WantHints		
<b>JobQueue</b>	<input type="text"/> JobThreadCount	100	
	<input type="text"/> MaxJobsForRead	-1	
	<input type="text"/> MinutesBetweenReadJobs	-1	
<b>Settings</b>	<input type="text"/> CorporaPath	.\corpora\	
	<input type="text"/> EnginesPath	.\engines\	
	<input type="text"/> JobsPath	.\Jobs\	
<b>Web interface</b>	<input type="text"/> Passwort	.....	
	<input type="text"/> Port	6133	
	<input type="text"/> SSLPort	6135	
	<input type="text"/> Username	admin	
	<input type="checkbox"/> Use SSL		

Document corpus: d3P

**dsearch.ini**

<b>General</b>	<input checked="" type="checkbox"/> DestroyJobs	<input checked="" type="checkbox"/>	
	<input type="text"/> Engine	shaft	
	<input type="checkbox"/> JobSynchron		
	<input type="text"/> ShaftUTLPath	.\shaft\64Bit\	

**corpus.ini**

<b>General</b>		
<input type="checkbox"/> CurrentVersionsOnly	<input checked="" type="checkbox"/>	<a href="#">i</a>
<input type="checkbox"/> UseFileStream	<input type="checkbox"/>	<a href="#">i</a>
<b>Search</b>		
<input type="checkbox"/> Expand Search	<input type="checkbox"/>	<a href="#">i</a>
<input type="checkbox"/> Expand Search Min Result	<input type="text" value="20"/>	<a href="#">i</a>
<b>Shaft_Global</b>		
<input type="checkbox"/> ForceNoHints	<input type="checkbox"/>	<a href="#">i</a>
<input type="checkbox"/> UTF8Support	<input checked="" type="checkbox"/>	<a href="#">i</a>
<b>SHAFT1</b>		
<input type="text" value="localhost"/>	<input type="text" value="3546"/>	<a href="#">i</a>
<input type="text" value="localhost"/>	<input type="text" value="3546"/>	<a href="#">i</a>
<input type="text" value="localhost"/>	<input type="text" value="6133"/>	<a href="#">i</a>
<b>SHAFT2</b>		
<input type="text" value="localhost"/>	<input type="text" value="3547"/>	<a href="#">i</a>
<input type="text" value="localhost"/>	<input type="text" value="3547"/>	<a href="#">i</a>
<input type="text" value="localhost"/>	<input type="text" value="6133"/>	<a href="#">i</a>
<b>SHAFT3</b>		
<input type="text" value="localhost"/>	<input type="text" value="3548"/>	<a href="#">i</a>
<input type="text" value="localhost"/>	<input type="text" value="3548"/>	<a href="#">i</a>
<input type="text" value="localhost"/>	<input type="text" value="6133"/>	<a href="#">i</a>

**shaft.ini**

<b>Backup</b>		
<input type="checkbox"/> CheckDay	<input type="text" value="0"/>	<a href="#">i</a>
<input type="checkbox"/> CheckDB	<input checked="" type="checkbox"/>	<a href="#">i</a>
<input type="checkbox"/> Time	<input type="text" value="00:00"/>	<a href="#">i</a>

Engine		
<input type="checkbox"/>	ArchiveLogs	<input type="checkbox"/>
<input checked="" type="checkbox"/>	AutoCorrect	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	AutoDefrag	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	AutoRepair	<input checked="" type="checkbox"/>
<input type="text" value="256000"/>	CacheSize	<input type="text" value="256000"/>
<input type="text" value="200"/>	CommitCount	<input type="text" value="200"/>
<input type="text" value="64"/>	Cursor	<input type="text" value="64"/>
<input type="text" value="150"/>	WordRevalue	<input type="text" value="150"/>
<input type="text" value="1"/>	Delimitermode	<input type="text" value="1"/>
<input type="text" value="3"/>	FirstWildcard	<input type="text" value="3"/>
<input type="text" value="-"/>	IcuExcludedDelimiter	<input type="text" value="-"/>
<input checked="" type="checkbox"/>	Journal	<input checked="" type="checkbox"/>
<input type="text" value="9"/>	LogLevel	<input type="text" value="9"/>
<input type="text" value="100"/>	MaxDocsForHints	<input type="text" value="100"/>
<input type="text" value="1000000"/>	MaxFilesPerWord	<input type="text" value="1000000"/>
<input type="text" value="100"/>	MaxWordsForHints	<input type="text" value="100"/>
<input type="text" value="2"/>	MinWordLen	<input type="text" value="2"/>
<input checked="" type="checkbox"/>	Numeric	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	Orderbyrank	<input checked="" type="checkbox"/>
<input type="text" value="1000000"/>	WildcardLimit	<input type="text" value="1000000"/>
<input type="checkbox"/>	SemanticExtend	<input type="checkbox"/>
<input checked="" type="checkbox"/>	StoreText	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	UseMostRecent	<input checked="" type="checkbox"/>
PerfLogs		
<input type="text" value="5000"/>	Abstract	<input type="text" value="5000"/>
<input type="text" value="5000"/>	Delete	<input type="text" value="5000"/>
<input type="text" value="5000"/>	Insert	<input type="text" value="5000"/>
<input type="text" value="5000"/>	Search	<input type="text" value="5000"/>
<input type="text" value="5000"/>	Spellcheck	<input type="text" value="5000"/>

### Two-node-cluster with an outsourced d.3 SHAFT engine

For a two-node-cluster, in which you to outsource a d.3 SHAFT engine, the configuration is mostly the same.

For more information on outsourcing of a d.3 SHAFT engine, read the chapter [SHAFT-Agent \(Outsource a Shaft-Engine\)](#).

For the computer with the engine to be outsourced only a new **Host** will be specified instead of using the value **localhost** for **Host**.

Document corpus: d3P

**corpus.ini**

SHAFT1	Host	d3-sql-712	
	Port	3547	
	SHAFTAgentPort	6133	

## 1.2.8. Integration of d.ecs monitor

### Introduction

If d.ecs monitor agent is installed on your system, regular status messages are sent to d.ecs monitor to inform the administrator about problems.

The following messages will be send by d.3 search:

- Start/stop the d.3 search server
- TCP/IP listen port could not be initialized.
- d.3 search license does not exist.
- Error messages occurring during the initialization of d.3 search server and lead to the shut down of the process.
- In cluster operation
  - The cluster is inactive

For better monitoring, the following monitoring objects are also sent to d.ecs monitor:

#### Refresh interval: 5 minutes

- Number of add/update/delete jobs in the queue

#### Refresh interval: 60 minutes

- Sufficient RAM per d.3 search engine (SHAFT)
- Specifies the required free disk space for the automatic defragmentation of the d.3 search engine (SHAFT).
- Sufficient disk space for automatic defragmentation per d.3 search engine (SHAFT)
- Size of the database per d.3 search engine (SHAFT)
- Duration of the backup per d.3 search engine (SHAFT)
- Fragmentation level of the database per d.3 search engine (SHAFT)
- Hours until last successful backup per d.3 search engine (SHAFT)

### Note

If a document corpus is added/changed/removed, the monitoring objects which refer to a d.3 search engine (SHAFT) are only transferred correctly to d.ecs monitor after a restart of d.3 search.

## 1.2.9. SSL encryption

### Setup

To encrypt all connections which will be established with d.3 search, do the following:

An encrypted communication for direct communication between d.3 server and d.3 search is not possible.

If you want to disable the SSL-encryption again, do the same as for the activation except that you disable the parameter instead of enabling.

→ Encrypt the web interface by enabling the parameter `[Webinterface]UseSSL` in the `dsearch.ini`.

For additional information see chapter [Configuration of d.3 search](#).

→ Encrypt the communication between d.3 server and d.3 search via the d.3 gateway.

Check, if in the file `<Windows.ini` the `[D3]SecretKey` of the d.3 gateway is set.

Copy the `dlink.ini` into the `Windows` directory of each d.3 search instance and restart the instances afterwards.

→ Copy the `public.key`-file from the installation directory of the d.3 gateway via which the d.3 server shall communicate into the installation directory of the d.3 server and restart it. Now, the connection between d.3 server and the first d.3 search instance will be encrypted.

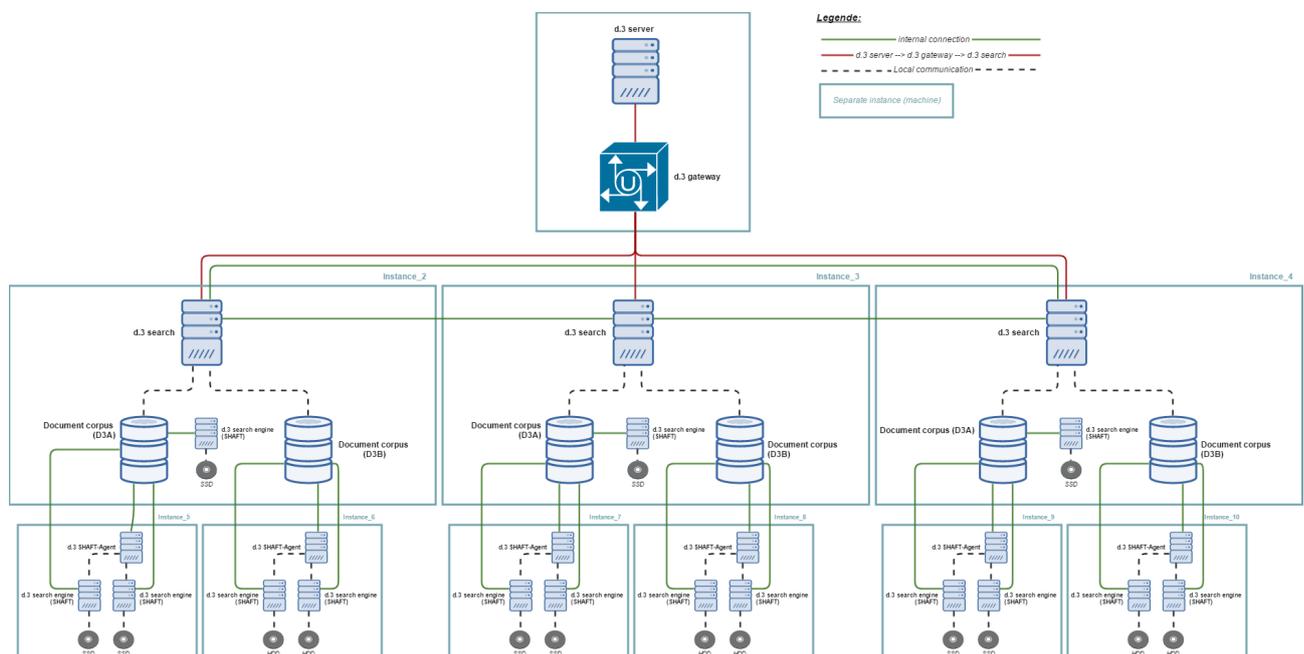
### Note

An encrypted communication for direct communication between d.3 server and d.3 search is not possible.

If you want to disable the SSL-encryption again, do the same as for the activation except that you disable the parameter instead of enabling.

## Architectural chart

The chart shows a scaled installation in the cluster. All green connections represent the connections of point 2.



## 1.3. Configuration

In this chapter you will learn more about the configuration.

### 1.3.1. Configuration via web interface

#### Common

From version 2.4.0 you can apply your settings from a web interface. Furthermore, the user interface provides additional information on the current status of the d.3 search system.

Open the administration interface via `http://host:6133/` or `https://host:6133` if SSL was enabled. A respective link can be found under **Start | Programs | d.velop | d.3 search**.

If the user interface is opened from the host on which it has been installed, then you do not need to enter a user name and password.

### Warning

After the installation of d.3 search you can open the administration interface. Change the credentials for `Username` and `Password` and enable SSL.

### Warning

The administration requires a Microsoft Internet Explorer in version 8, 9, 10 or 11.

From Microsoft Internet Explorer 11 TLS 1.2 must be disabled, if a SSL connection for the web interface will be configured. TLS 1.0 and 1.1 must still be enabled. The settings can be found in the Microsoft Internet Explorer under: **Internet options | Advanced**.

The language of the web interface is depending on the browser language. This means, is this German, then the German user interface will be displayed, else the English.

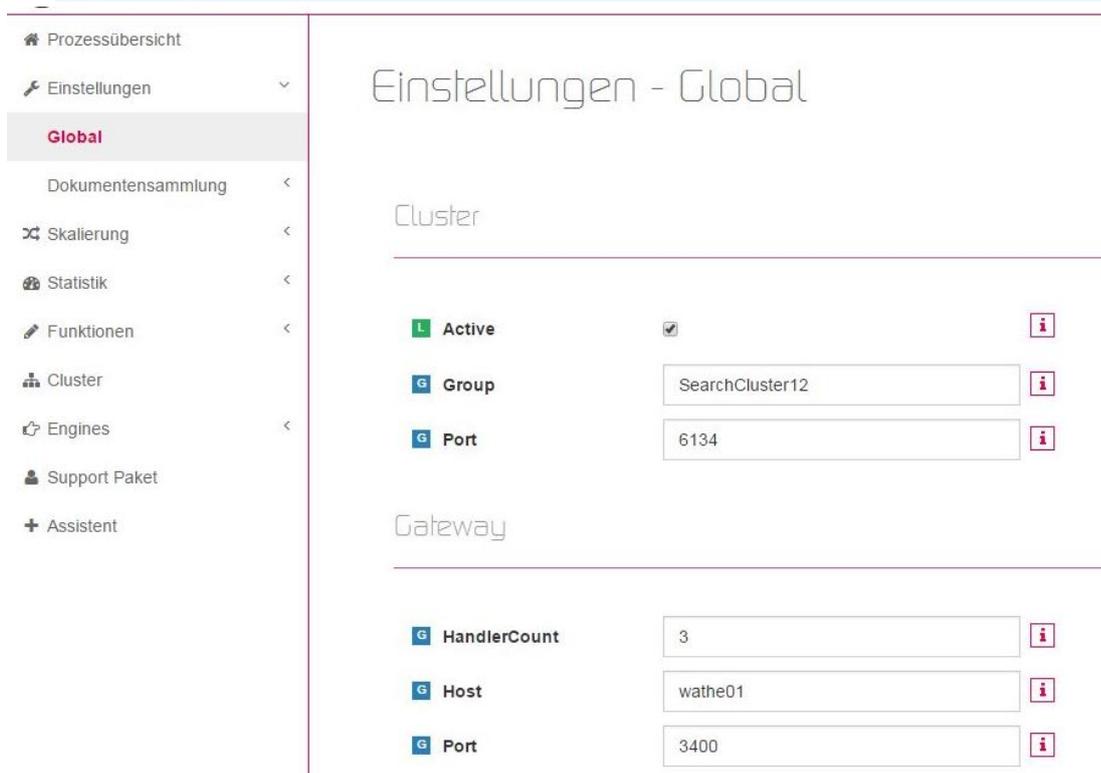
### Operation

The left menu bar allows you to switch between the individual sections of the administration. If you zoom out (minimize) the window, e.g. on a smartphone or tablet, open the menu bar with a click on .

If a menu option such as **Settings** was selected, then the left margin displays a selection of possible sub-sections that can be configured in this section.

### Note

Global or the settings of the document corpora already created.



The screenshot shows the 'Einstellungen - Global' configuration page. The left sidebar menu includes: Prozessübersicht, Einstellungen, Global (selected), Dokumentensammlung, Skalierung, Statistik, Funktionen, Cluster, Engines, Support Paket, and Assistent. The main content area is titled 'Einstellungen - Global' and is divided into two sections: 'Cluster' and 'Gateway'.

**Cluster settings:**

- Active:**  (Information icon)
- Group:** SearchCluster12 (Information icon)
- Port:** 6134 (Information icon)

**Gateway settings:**

- HandlerCount:** 3 (Information icon)
- Host:** wathe01 (Information icon)
- Port:** 3400 (Information icon)

If you need further information to the current input, click on the respective information icon next to the entry.

Afterwards a help text of the documentation appears.

Cluster

---

**Active**  

**Group**  

**Port**  

**Standard:** SearchCluster  
Gibt den Gruppennamen des Clusters an.

 Damit weitere Installationen im Netzwerk nicht automatisiert zum Cluster hinzugefügt werden, sollte der Gruppenname vor der Aktivierung der Clusterfunktion angepasst werden.

The following menu options are available for selection:

Name	Description
<a href="#">Process overview</a>	This section provides current information on the system status.
<a href="#">Setting</a>	In this section you can configure your d.3 search.
<a href="#">Scaling</a>	The configuration of the scaled d.3 search engines (SHAFT) can be done in this section.
<p><b>Note</b> This section is only available, if a scaling was configured.</p>	
<a href="#">Statistics</a>	In this section you can find statistics since the last restart.
<a href="#">Functions</a>	This section provides miscellaneous functions to test or change your document corpora.
<a href="#">Cluster</a>	In this section you can see the current status of your cluster and you can add/remove manually instance to/from the cluster.
<a href="#">Engines</a>	In this section you can find information on the d.3 search engine (SHAFT) and you can start/stop these manually.
<a href="#">Support package</a>	In this section you can create an information package for the d.velop Support in a support case.
<a href="#">Assistant</a>	In this section you can create new document corpora.

## Process overview

This section provides current information on the system status. This is divided in **Global information** and **Cluster information**.

### Note

The section **Cluster information** is only visible, if a cluster is configured.

### Global information

The **Global information** are divided in four further sections: **Status**, **Pending jobs**, **d.3 gateway** and **Directories**.

### Note

The d.3 gateway is only visible, if the connection to the d.3 gateway is configured.

Section	Name	Description
<b>Status</b>	Duration	Defines how much time has passed since the last start.

Section	Name	Description
<b>Pending jobs</b>	Connections	Specifies the number of currently open connections.
	Queries	Specifies the number of all queries accepted by d.3 search since the start.
	Number of document corpora	Specifies the number of all document corpora.
	License information	Specifies, which licenses are valid and loaded.
	Server information	Specifies the computer name and the port under which d.3 search is accessible.
	CPU-load	Specifies the CPU-load of the d.3 search server.
	Add-jobs	Specifies the number of the Add-jobs in the queue.
	Delete-jobs	Specifies the number of the Delete-jobs in the queue.
<b>Job threads</b>	Update-jobs	Specifies the number of the Update-jobs in the queue.
	ID	Specifies the ID of the thread. The threads are in charge for the processing of the job-queue.
<b>d.3 gateway</b>	Status	The status in form of a graphic display shows which status the thread has currently. Possible states are: <ul style="list-style-type: none"> <li>• paused</li> <li>• Waiting for jobs</li> <li>• &lt;JobNr&gt; &lt;DocID&gt; being processed -&gt; &lt;JobFileName&gt;</li> </ul>
	<IP> (<computer name>)	Specifies the IP address and the machine via which will be communicated with the d.3 gateway.
<b>Directories</b>	Illustration	The illustration shows, how many threads are currently processing a query of the d.3 gateway and how many have established a connection with the d.3 gateway. The sum of both numbers results in the overall number of threads, which establish a connection with the d.3 gateway, as specified in the settings.
	Corpora directory	Specifies the directory, under which the document corpora are stored.
	Engine directory	Specifies the directory, under which the search engine libraries are stored.
	Job directory	Specifies the directory, under which the jobs are stored.

## Globale Informationen

---

### Status

Laufzeit	00:01:31
Verbindungen	0
Anfragen	0
Anzahl Dokumentensammlungen	4
Lizenzinformation	d.3 search basic d.3 search Addon: File-Content d.3 search Addon: enterprise (Clustering/Scaling)
Server Information	D3-SQL-712 (3406)
CPU-Auslastung	0.00%

### Wartende Jobs

Add-Jobs	0
Delete-Jobs	0
Update-Jobs	0

### Job Threads

**ID: 0** *pausierend*

**ID: 1** *pausierend*

### d.3 gateway

10.95.0.80 (d3-sql-712)

0 arbeitend

6 anfragend

### Verzeichnisse

Corpora-Verzeichnis	C:\d3\d.3 search\corpora\
Engine-Verzeichnis	C:\d3\d.3 search\Jobs\
Job-Verzeichnis	C:\d3\d.3 search\engines\

## Cluster information

Under the section **Cluster Information** you find the following information:

Name	Description
Active	Specifies, if the cluster is active.
Cluster name	Specifies the name of the cluster array.
Number of servers	Specifies how many servers exist in the cluster.
Status	Specifies the status of the current server in the cluster.

## Cluster Informationen

---

Aktiv	true
Cluster-Name	AndreThesker
Anzahl Server	2
Status	Slave

### d.3 SHAFT agents

Under the section **d.3 SHAFT agents** you find all information to all used d.3 SHAFT agents.

Name	Description
Duration	Defines how much time has passed since the last start.
TCP port	Specifies the TCP-port via which the d.3 SHAFT Agent is available.
Corpora directory	Specifies the directory, under which the document corpora are stored.
SHAFT directory	Specifies the directory from which the d.3 search engine (SHAFT) is used.
LogLevel	Specifies the LogLevel for the d.3 SHAFT agent.
CPU-load	Specifies the CPU-load of the d.3 SHAFT agent process.

## d.3 SHAFT-Agenten

---

### DSEARCH1 (10.95.4.244)

Laufzeit	19:00:33
TCP-Port	6133
Corpora-Verzeichnis	C:\d3\d.3 search\corpora
SHAFT-Verzeichnis	C:\d3\d.3 search\shaft\64Bit\
LogLevel	6
CPUAuslastung	0.15%

For every used d.3 SHAFT agent information about the document corpora will be listed, which lay at the d.3 SHAFT agent.

Name	Description
d.3 search engine for: <document corpus ID>	Specifies, if the engine for the document corpus is currently online or offline.
Engine-ID	Specifies the ID of the engine.
Hostname	Specifies the computer name which uses this engine.
Process-runtime-parameters	Specifies with which start parameter the d.3 search engine (SHAFT) will be started.
Process path	Specifies the started process for the d.3 search engine (SHAFT).
CPU-load	Specifies the CPU-load of the d.3 search engine (SHAFT).

<b>d.3 search engine (SHAFT) für: d3p</b>	<b>Online</b>
→ Engine-ID	shaft1
→ Rechnername	d3-sql-712
→ Prozessaufrufparameter	C:\d3\d.3 search\corpora\d3p\d3-sql-712\shaft1\shaft.ini
→ Prozesspfad	C:\d3\d.3 search\shaft\64Bit\shaft.exe
→ CPUAuslastung	0.00%

## Settings

The menu option **Settings** is divided into the two sections **Global** and **Document corpus**.

In the following it is described which settings you can make on the respective pages.

### Note

If the current computer is in a cluster, the settings will be divided in global  and local  settings.

Global settings have an effect on all instances in the cluster, local settings have an effect only for the current instance.

The global settings can only be changed on the **Master**.

### Global

In this section, the global settings for d.3 search are specified.

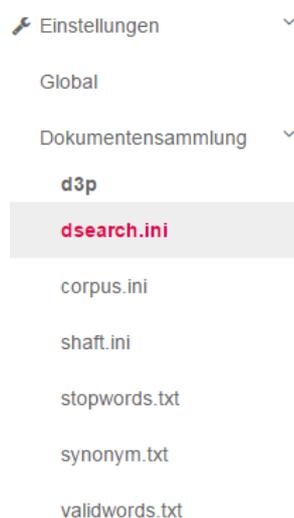
Information on the parameters can be found in chapter [Configuration of the d.3 search server](#)

### Document corpus

In the menu bar below this point you find your created document corpora.

For each document corpus the sub-menu options **dsearch.ini**, **corpus.ini**, **shaft.ini**, **stopwords.txt**, **synonym.txt** and the **validwords.txt** exist.

If the **DelimiterMode=3** is set in the `shaft.ini` an additional option **delimiter.txt** appears.



Information on the parameters can be found in the respective chapters:

Sub-menu option	Information on parameters
dsearch.ini	Configuration of the document corpora
corpus.ini	Configuration of the document corpora
shaft.ini	Configuring the d.3 search engine
stopwords.txt	Managing the stopwords
synonym.txt	Management of the synonym.txt
validwords.txt	Management of the ValidWords
delimiter.txt	Configuring the d.3 search engine

## Scaling

The menu option **Scaling** contains multiple sections. Each section has the name of a document corpus.

With a click on the respective document corpus the settings regarding the scaling of this document corpus can be done.



For each d.3 search engine (SHAFT) an entry can be found on the page. An entry consists of the following information or available settings:

### SHAFT1 - dsearcht3

<b>L Path</b>	C:\d3\d.3 search\corpora\d3r\dsearcht1\sf	
<b>L Database</b>	corpora\d3r\dsearcht1\shaft1\	

In the title you can find the ID of the d.3 search engine (SHAFT) for the current document corpus and behind that, divided by a hyphen, the host name on which the d.3 search engine (SHAFT) is stored.

The following available settings are described below:

Parameters	Description
Path	Storage location for the online backup
Database	Relative path to the storage directory of the index data

**Note**

If this path will be changed, the respective d.3 search engine (SHAFT) must be stopped before. Afterwards, the path can be individually adjusted.

The actual migration of the index data to the new path **must** be done manually.

## Statistics

The menu option **Statistics** is divided into the two sections **Global** and **Document corpus**.

The section itself is divided into

- **Functions**
- **Information (overall)**
- **Information (per d.3 search engine (SHAFT))**

### Functions

The section **Functions** provides information on the called API-functions of d.3 search.

If d.3 search is restarted, all statistics will be resetted.

Furthermore, a reset can be done manually by a click on **Reset statistics**.

This resets only the statistic of the current corpus.

For each API function of d.3 search the following information will be written into the statistics:

Name	Description
Opening	Name of the called API function
Opened connections	Number of open connections disregarding if faulty or successful.  This includes every opened connection to the document corpus, this means every connection to the d.3 search engine (SHAFT). For a scaling the connection to all d.3 search engines (SHAFT) will be count as one connection.
Closed connections	Number of successfully closed connections. Includes average needed time per connection.  This includes every opened connection to the document corpus, this means every connection to the d.3 search engine (SHAFT). For a scaling the connection to all d.3 search engines (SHAFT) will be count as one connection.
Opened connections	Number of currently opened connections. These are all connections currently opened to a document corpus. This means, usually an opened connection to a d.3 search engine (SHAFT).
Connections which encountered errors	Number of connections which have encountered errors. Every attempt to establish a connection to the document corpus (the d.3 search engine (SHAFT)) is counted as an error. So this could lead to a queue of d.3 search with a lot of errors caused by a connection error.

**Note**  
A connection is e.g. faulty, if:

- d.3 search engine (SHAFT) is not available.
- d.3 search engine (SHAFT) is ReadOnly.

### Global

Under the section **Global** all API functions are listed, which can not be assigned to a document corpus.

In the following table you can find the meaning of the listed API functions.

Call	Description
GetEngine	Lists the search engines supported by d.3 search with its functions.
GetServerStatus	Returns the current status of d.3 search server.

# Statistik - Global

## Funktionen

Aufruf	Geöffnete Verbindungen	Geschlossene Verbindungen	Offene Verbindungen	Auf Fehler gelaufene Verbindungen
<b>GetEngine</b>	0	0	0	0
<b>GetServerStatus</b>	0	0	0	0
<b>Summe</b>	0	0	0	0

Reset Statistik (28.05.2015 14:28:48.059)

### Document corpus

Under der section **document corpus** you find a statistics for every document corpus.

The listed API functions refer to the selected document corpus.

In the following table you can find the meaning of the listed API functions.

Call	Description
Add	Adds a document to a document corpus.
Update	Changes a document in a document corpus.
Delete	Deletes a document from a document corpus.
GetCorpora	Lists all created document corpora and its properties.
QueryDocuments	Executes a search query in the document corpus.
GetSimilarWords	Specifies all words which are similar to a specific word.
GetDocument	Returns a list of all words which are saved to a document in the document corpus. The content is in most cases not equivalent to the original document because of stop-word-lists and the removal of duplicate words.
GetCorpusProperties	Determines the specific properties of a document corpus.
SetReadOnlyModus	Sets the document corpus in the mode "read only".
GetSummary	Restores a TextSummary (Abstract) for a document for the specified documents. This means, phrases will be returned, in which the original search (query) was found.

# Statistik - d3p

## Funktionen

Aufruf	Geöffnete Verbindungen	Geschlossene Verbindungen	Offene Verbindungen	Auf Fehler gelaufene Verbindungen
Add	0	0 (0 ms)	0	0
Update	0	0 (0 ms)	0	0
Delete	0	0 (0 ms)	0	0
GetCorpora	0	0 (0 ms)	0	0
QueryDocuments	0	0 (0 ms)	0	0
GetSimilarWords	0	0 (0 ms)	0	0
GetDocument	0	0 (0 ms)	0	0
GetCorpusProperties	0	0 (0 ms)	0	0
SetReadOnlyMode	0	0 (0 ms)	0	0
GetSummary	0	0 (0 ms)	0	0
<b>Summe</b>	0	0 (0 ms)	0	0

Reset Statistik (28.05.2015 14:28:48.059)

### Information (overall)

The section **Information (Complete)** shows you the following values per section.

In the section **Global** the respective values will be cumulated over all d.3 search engines (SHAFT) of all document corpora.

Name	Description
Overall size	Overall size of all d.3 search engine (SHAFT) - databases, which belong to the selected document corpus.
Size of the search index	Size of the d.3 search engine (SHAFT) - databases, which belong to the selected document corpus without StoreText- and nGramm-files.
Size due to use of "FirstWildcard" smaller than 3	Size of the d.3 search engine (SHAFT) database files for the use of a wildcard at first or second position in the search.  Additional information on the parameter <code>FirstWildcard</code> can be found under <a href="#">Configuration of the d.3 search engine</a> .

**Note**  
The percentage behind this values says, how much percent the size is in comparison to the **Size of the search index**.

Name	Description
Size due to additional storage of full-text	Size of the d.3 search engine (SHAFT) database files for the additional storage of the full-text information. Additional information on the parameter <code>StoreText</code> can be found under <a href="#">Configuration of the d.3 search engine</a> .
<p><b>Note</b></p> <p>The percentage behind this values says, how much percent the size is in comparison to the <b>Size of the search index</b>.</p>	
Number of documents	Number of documents within a document corpus
Number of words	Number of distinct words within a document corpus.

### d3p - SHAFT1 (localhost)

#### Datenbankgröße

Gesamtgröße	0.018 GB
Größe des Suchindexes	0.002 GB
Größe aufgrund der Verwendung von "FirstWildcard" kleiner als 3	0.016 GB
Größe aufgrund der zusätzlichen Speicherung des Volltextes	0 GB

#### Datenbankinhalt

Anzahl Dokumente	3
Anzahl Wörter	2

#### Informationen zum Datenbankverzeichnis

Lesemodus	Nein
Datenbankverzeichnis	C:\d3\d.3 search\corporat\d3p\SHAFT1
Speicherkapazität	15.388 GB frei von 44.9 GB
Belegter Speicherplatz	<div style="width: 66%; background-color: #0070c0; height: 10px; display: inline-block;"></div> 66%

### Information (per d.3 search engine (SHAFT))

Below the section **Information (per d.3 search engine (SHAFT))** you can find information on the individual d.3 search engines (SHAFT).

The following values are listed in this section:

Name	Description
Overall size	Full-size of the d.3 search engine (SHAFT) databases.
Size of the search index	Size of the mere d.3 search engine (SHAFT) databases.
Size due to use of "First-WildCard"" smaller than 3	Size of the d.3 search engine (SHAFT) database files for the use of a wildcard at first or second position in the search. Additional information on the parameter <code>FirstWildcard</code> can be found under <a href="#">Configuration of the d.3 search engine</a> .
Size due to additional storage of full-text	Size of the d.3 search engine (SHAFT) database files for the additional storage of the full-text information. Additional information on the parameter <code>StoreText</code> can be found under <a href="#">Configuration of the d.3 search engine</a> .
Number of documents	Number of documents within the d.3 search engine (SHAFT) databases.
Number of words	Number of distinct words within the d.3 search engine (SHAFT) databases.
Read mode	Specifies, if the d.3 search engine (SHAFT) is in read mode. A reason for this could be: Backup is running, engine is being moved.

Name	Description
Database directory	Specifies the path to the database directory of the d.3 search engine (SHAFT).
Storage capacity (database directory)	Specifies the free and total storage capacity of the database directory of the d.3 search engine (SHAFT).
Used disk space (database directory)	Displays graphically how much percent of the storage capacity of the database directory is already used.
Backup in progress	Specifies, if a backup is currently executed.
Backup directory	Specifies the path to the backup directory of the d.3 search engine (SHAFT).
Storage capacity (backup directory)	Specifies the free and total storage capacity of the backup directory of the d.3 search engine (SHAFT).
Used disk space (backup directory)	Displays graphically how much percent of the storage capacity of the backup directory is already used.
Required disk space for the automatic defragmentation	Specifies the required free disk space for the automatic defragmentation of the d.3 search engine (SHAFT).
<p><b>Warning</b> If this value is red, then not enough disk space is available.</p>	
Required disk space for the online backup	Specifies the required free disk space for the online backup of the d.3 search engine (SHAFT).
<p><b>Warning</b> If this value is red, then not enough disk space is available.</p>	
Required disk space for the automatic repair	Specifies, how much free disk space for the automatic repair of the d.3 search engine (SHAFT) is required.
<p><b>Warning</b> If this value is red, then not enough disk space is available.</p>	

#### Informationen zum Backupverzeichnis

Sicherung läuft	Nein
Backupverzeichnis	C:\d3\d.3 search\corpora\d3p\SHAFT1\backup
Speicherkapazität	15.388 GB frei von 44.9 GB
Belegter Speicherplatz	<div style="width: 66%; background-color: #0070c0; color: white; text-align: center; padding: 2px;">66%</div>

#### Benötigter Speicherplatz

Benötigter Speicherplatz für die automatische Defragmentierung	0.5 GB
Benötigter Speicherplatz für das online Backup	0.518 GB
Benötigter Speicherplatz für die automatische Reparatur	0.518 GB

#### Arbeitsspeicher

Arbeitsspeicherkapazität	0.282 GB frei von 4 GB
Verwendeter Arbeitsspeicher	<div style="width: 93%; background-color: #0070c0; color: white; text-align: center; padding: 2px;">93%</div>

## Functions

The menu option **functions** is divided into three sections **Search functions**, **information** and **change document**.

In the following, the individual functions and the respective parameters are described:



### Search functions

This section contains all functions, which can be used for a search for a document.

#### Search

The function **Search** provides you the option to start a search query on the current instance.

The following criteria can be used for this:

Parameters	Description
Corpus	Name of the document corpus to be browsed.
Query	searched content (search query itself).
QueryCodepage	Specifies the code page of the search query.
SimilarDocId	Document ID of a document in the document corpus; if this field is filled, then similar documents to the specified document will be searched.
StartRow	Specifies from which row the results are to be displayed.
MaxRowCount	Specifies how many results are to be displayed.
MinRelevance	Returns the minimum relevance of a search result.
WantHints	Specifies, if search clouds are to be returned.
WantAbstract	Specifies, if an abstract (short text summary of the result) is to be returned.
NoExtend	Specifies, if an extended search based on the hints is to be executed.
QueryByExample	If this option is enabled, then a document can be uploaded. Documents will be searched, which are similar to the uploaded document.
Handle	Handle which can be used to avoid the execution of a renewed search query and instead it is taken from the cache.

## Suchen

---

Korpus	<input type="text" value="d3andre"/>
Query	<input type="text"/>
QueryCodePage	<input type="text"/>
SimilarDocId	<input type="text"/>
StartRow	<input type="text" value="0"/>
MaxRowCount	<input type="text" value="1000"/>
Minimale Relevanz	<input type="text" value="0"/>
WantHints	<input type="checkbox"/>
WantAbstract	<input type="checkbox"/>
NoExtend	<input type="checkbox"/>
QueryByExample	<input type="checkbox"/>
Handle	<input type="text" value="0"/>

### Text conclusion

The function **Text summary** provides you the option to get a text summary (abstract) of a document.

The following criteria can be used for this:

Parameters	Description
Corpus	Name of the document corpus to be browsed.
DocId	Document ID of the document to be combined.
AbstractSize	Specifies how many line breaks are to be returned.
Context	Specifies under which context a summary shall be created. In most cases this is equivalent to the searched word.
HintWords	List of words from the hints, which are to be considered for the summary.

# Textzusammenfassung

---

Korpus	<input type="text" value="d3p"/>
DocId	<input type="text"/>
AbstractSize	<input type="text" value="100"/>
Context	<input type="text"/>
HintWords	<input type="text"/>

## Information

In this section are all functions, which can be used to get information on a document.

### Document-information

The function **Document information** provides you the option to get all saved words of a document on the current instance.

The following criteria can be used for this:

Parameters	Description
Corpus	Name of the document corpus to be browsed.
Document-ID	Document ID of the document whose saved words are to be determined.

**Note**  
If the parameter `CurrentVersionsOnly` is disabled, the document ID must be specified in the format `<document-ID>.<document-version>`. If the parameter is enabled, the format `<document ID>` must be used.

# Dokument-Informationen

---

Korpus	<input type="text" value="d3p"/>
Dokumenten-ID	<input type="text"/>

## Similar words

The function **Similar words** returns comparable/similar words to a specified word.

Parameters	Description
Corpus	Name of the document corpus to be browsed.
Word	the word for which similar words are to be searched
Minimum relevance	Returns the minimum relevance of a search result.

## Ähnliche Wörter

Korpus

Wort

Minimale Relevanz

## Change document

In this section are all functions, which can be used to edit/add/delete documents.

### Add

The function **Add** provides you the option to add a document to the search index on the current instance.

The following criteria can be used for this:

Parameters	Description
Corpus	Name of the document corpus to be browsed.
Document-ID	ID referencing the document to be imported.
Document	the document to be imported

## Hinzufügen

Korpus

Dokumenten-ID

Document

## Change

The function **Change** provides you the option to change a document in the search index on the current instance.

The following criteria can be used for this:

Parameters	Description
Corpus	Name of the document corpus to be browsed.
Document-ID	ID referencing the document to be changed.
Document	the document to be imported

## Ändern

---

Korpus

Dokumenten-ID

Document

## Delete

The function **Delete** provides you the option to delete a document in the search index on the current instance.

The following criteria can be used for this:

Parameters	Description
Corpus	Name of the document corpus to be browsed.
Document-ID	ID referencing the document to be deleted.
Synchronous	Specifies, if the document is to be deleted directly (synchronous) or if a job is to be added in the queue (asynchronous).

## Löschen

---

Korpus

Dokumenten-ID

Synchron

## Cluster

The menu option **Cluster** is only visible, if a cluster is configured. On this page you can find information on cluster, instances to add to a cluster and to remove these again. The cluster section is divided into **Information**, **Setting**, **Cluster server** and **Log**.

### Information

In this section are essential information to the cluster, this includes:

Information	Description
Active	Specifies, if the cluster is active. A cluster is inactive, if an instance in the cluster has an error. If the cluster is not active, no changes on the search index can be done.
Status	Specifies, which mode the current instance has.

## Information

<b>Aktiv</b>	true
<b>Status</b>	Slave

### Settings

In this section you can configure, if the information under **Cluster server** are to be updated continuously.

## Einstellungen

**Automatische Aktualisierung**

### Cluster server

In this section all server are listed, which are in the cluster.

The list contains the following information:

Information	Description
IP address	Specifies the IP address of the server in the current network.
Hostname	Specifies the computer name of the server in the current network.
<b>Warning</b> In the cluster a server is identified by its computer name. A later change of the computer name has effect on the cluster.	
Mode	Specifies the mode, which an instance has at this time (master/slave/unknown).
Last message	Specifies, when the last communication between the current instance and the selected instance happened.
Information	Specifies additional information for a server (e.g. restarting at this moment)
Error messages	Specifies the error messages of an instance (e.g. d.3 search engine (SHAFT) is not started).

The button **Delete** is inactive, if the instance is active. A server can only be removed from the cluster, if the respective instance is turned off. Also compare [Remove cluster member](#). With a click on the button

**Add** an instance can be added to the cluster manually. After the click an additional row appears, in which the computer name or the IP address can be specified. For additional information see [Configure new d.3 search instance for the cluster](#)

## Log

In this section the log outputs of the cluster are displayed. The colors have the following meaning:

Color	Meaning
Green	The request was transmitted successfully.
Red	The request could not be transmitted successfully (possibly timeout).
Blue	Information (e.g. an instance has reported an error).
Grey	The request was received successfully.

## Log

---

```
28.01.2015 13:08:36: [3420] [RECV]: dsearchdc1 - IAMALIVEMASTER
28.01.2015 13:08:42: [3928] [RECV]: dsearchdc1 - IAMALIVEMASTER
28.01.2015 13:08:47: [3868] [RECV]: dsearchdc1 - IAMALIVEMASTER
```

## Engines

Via the menu option **Engines** you can find information on the d.3 search engines (SHAFT). Furthermore, the individual engines can be started, stopped or restarted. The menu option contains for every configured document corpus a separate option.

In the section **Global functions** you can start, stop or restart all engines of a document corpus simultaneously. To do so, click on the respective button.

## Globale Funktionen

---

Alle starten

Alle stoppen

Alle neustarten

Afterwards you find for every d.3 search engine (SHAFT) a separate section.

The section contains the following information:

Information	Description
Status	Specifies, if the process is started or stopped.
ID	Specifies the ID of the engine within the document corpus.
Corpus	Specifies the corpus for the running engine.

Information	Description
Hostname	Specifies the computer name under which the engine is responding.
File name	Specifies the file name of the process. This information is obsolete, if the engine is outsourced.
Command line	Specifies the command line which calls the process. This information is obsolete, if the engine is outsourced.
Process-ID	Specifies the current process-ID of the process. This information is obsolete, if the engine is outsourced.
CPU-load	Specifies the CPU-load for the d.3 search engine (SHAFT).
Start time	Specifies, when the process was started.

## SHAFT2

---

<b>Status</b>	Online
<b>ID</b>	SHAFT2
<b>Korpus</b>	d3p
<b>Rechnername</b>	localhost
<b>Dateiname</b>	C:\d3\d.3 search\shaft\64Bit\shaft.exe
<b>Befehlszeile</b>	C:\d3\d.3 search\corpora\d3p\SHAFT2\shaft.ini
<b>Prozess-ID</b>	4956
<b>CPUAuslastung</b>	0%
<b>Startzeit</b>	16.02.2016 11:40:28.300

Stop

Neu starten

## Service

### Support package

Via the menu option **Support package** you can create a support package.

The contents of the support package can be changed at will. To create the support package, follow the steps below:

- First select the respective document corpora.
- Choose which package size the support package should have and which content should be provided in the support package.
- Create the support package.
- Download the support package.

### Note

For the creation of the support package additional hard disk space is required. Make sure that this is available. A support package will be deleted after the download by the user.

## Service functions

The menu option **Service functions** allows you to execute special functions for individual document corpora.

The following functions can be executed per document corpus:

### Warning

By the execution of the functions you can paralyze the operation of the full-text search. So execute this functions only if you are absolutely certain.

Name	Description
Rename document corpus	<p>You can rename the generated document corpus with this function. This changes the directory structure and the settings. Use this function, for example, if you did not set the parameter <b>Delete document corpus</b> during a parallel regeneration.</p> <p><b>Note</b> Please note that a change may affect the mappings of the document type to the document corpora in the d.3 admin.</p>
Start backup manually	<p>This function allows you to start the backup of the d.3 search engine (SHAFT) manually. Perform this action for security reasons, if you apply changes to the settings of the d.3 search engine (SHAFT).</p> <p><b>Warning</b> During the backup, only read access to the document corpus is possible.</p>
Disable/enable read mode	<p>Set the document corpus to read mode for a flush to take place to allow you to perform a manual backup of the d.3 search engine (SHAFT).</p>
Remove document corpus	<p>You should only delete a document corpus, if you no longer need the document corpus. Use this function, for example, if you did not set the parameter <b>Delete document corpus</b> during a parallel regeneration.</p> <p><b>Warning</b> This process cannot be undone.</p>

# Service Funktionen

d3p

---

Dokumentensammlung umbenennen	
Backup manuell starten	
Lesemodus aktivieren	
Dokumentensammlung entfernen	

## Assistant

Under the menu option **Wizard** you find multiple wizards for the management of d.3 search.

- [The wizard for the creation of a new document corpus](#)
- [The wizard for the reindexing of a document corpus](#)

### 1.3.2. Configuration of the d.3 search server

Global settings for the d.3 search server can be configured in the file `dsearch.ini`.

The file is located in the d.3 search directory (e.g. `D:\d3\d.3 search\`).

The entries in this file have the following meaning:

Section	Entry	Meaning
[General]		
	Port	Default: 3406  The TCP/IP-port, under which the d.3 search server is accessible for the d.3 server.  In case of direct communication, this value must match the server parameter <code>d3fc_dsearch_server_tcpip_port</code> in the administration to allow connection establishment. If a communication via the d.3 gateway is established, the default value should be kept.
	WantHints	(1: yes; 0: no)  Default: 1  If this parameter is set to "0", then no suggestions are provided for the refined search. If this option is set to "1", then words for the refined search are provided.  Disabling this function can increase the performance of the search engine.

Section	Entry	Meaning
	LogLevel	<p>Default: 6</p> <p>LogLevel=0: only errors are logged in the log viewer</p> <p>LogLevel=6: normal logging of d.3 search API calls</p> <p>LogLevel=9: all messages (including debug entries)</p> <p>LogLevel=10: The file prepared for the Shaft engine ( . \$\$\$) remains in the Jobs directory as .tmp.</p> <div style="background-color: #ffe6e6; padding: 5px; margin: 10px 0;"> <p><b>Warning</b></p> <p>If the LogLevel is permanently set to 10, this could yield your hard disk capacity insufficient.</p> </div> <div style="background-color: #e6f2ff; padding: 5px; margin: 10px 0;"> <p><b>Note</b></p> <p>If the d.3 search server window is in the foreground, the loglevel can be set to debug mode with the shortcut key &lt;Strg&gt;+&lt;Alt&gt;+&lt;D&gt;.</p> </div>
	MinRelevance	<p>Default: 0</p> <p>Returns the minimum relevance of a search result. The value is only used, if the search query contains no value for this option.</p> <p>Else, the value from the search query is used.</p>
	LogFile	<p>Defines the file name of a log-file to be used for additional logging. If this option is not set, then no separate log-file is created.</p> <div style="background-color: #e6f2ff; padding: 5px; margin: 10px 0;"> <p><b>Note</b></p> <p>C:\d3\log.txt</p> </div>
	OMOMsgDist	<p>Default: 10</p> <p>Defines the minimum time period between two d.3 object monitor messages in seconds.</p>
[Settings]		
	JobsPath	<p>Default: d.3 search directory\jobs\ (e.g. D:\d3\d.3 search\jobs\)</p> <p>Specifies the directory under which the jobs are stored. To the jobs count add-/update- and delete-jobs, which will be written to the respective sub-directories.</p>
	Corpora-Path	<p>Default: d.3 search directory\corpora\ (e.g. D:\d3\d.3 search\corpora\)</p> <p>Specifies the directory, under which the corpora are stored.</p>
	Engines-Path	<p>Default: default: d.3 search directory\engines\ (e.g. D:\d3\d.3 search\engines\)</p> <p>Specifies the directory, under which the engines are stored. The respective folder contains the DLL files for the comprehension with the respective search engine.</p>
[Webinterface]		
	Username	<p>Standard: admin</p> <p>Specifies the user name to log into the web interface from another machine.</p> <p>When logging via the local machine, no authentication takes place.</p>
	Password	<p>Standard: dvelop</p> <p>Specifies the password to log into the web interface from another machine.</p> <p>When logging via the local machine, no authentication takes place.</p> <div style="background-color: #e6f2ff; padding: 5px; margin: 10px 0;"> <p><b>Note</b></p> <p>The password is stored in encrypted form and can only be set via the web interface.</p> </div> <div style="background-color: #ffe6e6; padding: 5px; margin: 10px 0;"> <p><b>Warning</b></p> <p>Change the password directly after the installation.</p> </div>

Section	Entry	Meaning
	Port	Default: 6133 Specifies the port under which the web interface is accessible.
	UseSSL	(1: yes; 0: no) Default: 0 Specifies, if the web interface is only accessible via an SSL connection. If this parameter is enabled, it can still be accessed to the web page local on the computer via http. An external connection works then only via https. A redirect takes places for this in the browser automatically.  <b>Note</b> If no SSL certificate was stored, then an unsigned certificate is automatically stored with the next restart.  <b>Warning</b> From Microsoft Internet Explorer 11 TLS 1.2 must be disabled, if a SSL connection for the web interface will be configured. TLS 1.0 and 1.1 must still be enabled. The settings can be found in the Microsoft Internet Explorer under: <b>Internet options   Advanced</b> .
	SSLPort	Default: 6135 Specifies the port under which the SSL connection for the web interface shall be established. If SSL is enabled, the unencrypted access for the web interface is always available local. From another computer the browser forwards automatically to the HTTPS connection.  <b>Note</b> The parameter only applies, if UseSSL for the web interface was enabled.
[Communication]		
	SSLCert-File	Specifies the path to the SSL certificate.  <b>Note</b> The SSL certificate is used for the HTTPS connections to the web interface.
	SSLKeyFile	Specifies the file path to the SSL key for the defined SSL certificate.  <b>Note</b> The SSL certificate is used for the HTTPS connections to the web interface.
	SSLPass-word	Specifies the password of the SSL certificate.  <b>Note</b> The SSL certificate is used for the HTTPS connections to the web interface.
	D3FCPass-word	Specifies the password for d3fc communication.  <b>Note</b> The same password must be stored in d.3 admin for the communication with d.3 search.
[Cluster]		

Section	Entry	Meaning
	Active	<p>(1: yes; 0: no)</p> <p>Default: 0</p> <p>This parameter defines, if the cluster functionality is enabled.</p> <div style="background-color: #e6f2ff; padding: 10px; margin-top: 10px;"> <p><b>Note</b></p> <p>If the cluster functionality is enabled, the d.3 search instances with the identical cluster group will be combined automatically to a cluster via an UDP-broadcast.</p> <p>The respective document corpora will be transmitted automatically to the respective instances.</p> </div>
	Group	<p>Default: SearchCluster</p> <p>Specifies the group name of the cluster.</p> <div style="background-color: #e6f2ff; padding: 10px; margin-top: 10px;"> <p><b>Note</b></p> <p>To avoid that further installations in the network will be added to the cluster automatically, the group name should be adjusted before enabling the cluster function.</p> </div> <div style="background-color: #ffe6e6; padding: 10px; margin-top: 10px;"> <p><b>Warning</b></p> <p>The cluster name must not contain an underscore (_). Else, d.3 search can not start.</p> </div>
	Port	<p>Default: 6134</p> <p>Specifies the port via which the cluster communication takes place.</p> <p>d.3 search establishes via this port a TCP- and UDP-server. The UDP-server is for the automatic adding of new instances. The TCP-server controls afterwards the monitoring and communication between the instances.</p> <div style="background-color: #e6f2ff; padding: 10px; margin-top: 10px;"> <p><b>Note</b></p> <p>For the creation of a cluster all instances must have configured the same port.</p> </div>
[Gateway]		
	Host	<p>Default:</p> <p>defines the host name of the d.3 gateway; if you do not specify a host name, the functionality is disabled</p> <div style="background-color: #ffe6e6; padding: 10px; margin-top: 10px;"> <p><b>Warning</b></p> <p>If possible, specify a host name here, which is accessible by other computer systems. Entries like, e.g. localhost or 127.0.0.1 are possible but should be avoided.</p> </div>
	Port	<p>Default: 3400</p> <p>Specifies the port via which the d.3 gateway is available.</p>
	Handler-Count	<p>Default: 3</p> <p>Specifies the number of connections to the d.3 gateway to be established.</p> <div style="background-color: #e6f2ff; padding: 10px; margin-top: 10px;"> <p><b>Note</b></p> <p>If the requests of the d.3 gateway will not be accepted after 30 seconds by d.3 search, then the HandlerCount should be increased.</p> </div>
[JobQueue]		
	JobThread-Count	<p>Default: 2</p> <p>Specifies the number of threads which process the job-queue.</p> <p>The number should be increased, if the queue can not be processed and is permanently growing.</p>

Section	Entry	Meaning
	MaxJobs- ForRead	<p>Default: -1</p> <p>Specifies, how many jobs are to be loaded maximal from the jobs folder at the start.</p> <p>For the value -1 all jobs will be loaded from the jobs folder, else the respective number.</p> <div style="background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p><b>Note</b></p> <p>The usage makes sense, if there are so many jobs in the folder that the start of d.3 search takes more time.</p> <p>By the limitation of the number of jobs which are read at the start, the start process of the d.3 search can be decreased and the full-text search is available earlier.</p> </div> <div style="background-color: #ffe6e6; padding: 10px; margin: 10px 0;"> <p><b>Warning</b></p> <p>Possibly, when using the parameter some jobs can not be read. To load these jobs later, the parameter [JobQueue]MinutesBetweenReadJobs can also be used. Else, the jobs will be loaded at the next restart.</p> </div>
	Minute- sBetween- ReadJobs	<p>Default: -1</p> <p>The parameter specifies after how many minutes the jobs folder shall be searched for new jobs.</p> <p>To keep the job-queue low and to avoid redundant accesses to the hard disk, the jobs will be loaded only if the current job-queue has none entries anymore.</p> <p>For the value -1 the jobs folder will not be searched retroactively.</p> <div style="background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p><b>Note</b></p> <p>The usage makes sense, if the parameter [JobQueue]MaxJobsRead will be used to accelerate the start of d.3 search.</p> </div>

### 1.3.3. Configuration of a document corpus

The d.3 search server manages one or several document corpora.

A document corpus contains a number of text documents (ASCII-format) referenced by a unique identifier.

This identifier is defined by the d.3 server in the form `DOC_ID.VERSION_NUMBER`.

Every document corpus is monitored by a process of the search engine associated to the corpus. As a result, the index data of a search engine are not monitored by the d.3 search server but by an external process. The d.3 search server only forwards the change actions such as the insertion or deletion of a document to this external process.

As the d.3 search server can establish a connection to multiple document corpora, it is not actually necessary that all text information of a repository are stored in one document corpus.

#### Creating a document corpus

The menu option in the web interface **Wizard | Creation** provides the option to create a new document corpus.

For the creation of a document corpus follow the instructions of the wizard.

#### Note

This wizard replaces the previous application **d.3 search wizard**.

## Step 1: Welcome

In the first step you will be prompted to do the following steps before creating a document corpus.

In the cluster operation a successful created document corpus will be automatically distributed to all nodes.

Choose **Next >>** to move to the next tab.

### Warning

Consider that some settings can not be changed retroactively without a reindexing. Depending on the size of the document corpus, this may take a while.

### Note

A document corpus can only be created, if the configured cluster is active.

## Assistent zur Anlage einer Dokumentensammlung

[Willkommen](#)[Basis](#)[Indexierung](#)[Suche](#)[SHAFT-Einstellung](#)[Backup](#)[Fertigstellen](#)

### Schritt 1: Willkommen

Willkommen beim d.3 search Assistenten zur Anlage einer Dokumentensammlung.  
Führen Sie die folgenden Schritte durch, um eine neue Dokumentensammlung zu erstellen.

#### WICHTIG:

Bitte berücksichtigen Sie bei der Konfiguration die d.3 search Dokumentation und bedenken Sie, dass einige Einstellungen später, ohne Reindexierung, nicht geändert werden können.

[<< Zurück](#)[Weiter >>](#)

## Step 2: Basis

In the second step the basic settings of the new document corpus will be defined.

The following basic settings can be made:

Setting	Description
Name of the document corpus	the name of the new document corpus; the maximum length of is 255 characters and it must not contain special characters
	<p><b>Note</b></p> <p>If the d.3 server has not been configured otherwise, all document types are imported to the document corpus D3"XX" where "XX" stands for the repository ID.</p> <p>Additional information can be found here: <a href="#">d.3 server settings for d.3 search</a>.</p>

Setting	Description
<b>Enable mirroring</b>	<p>If this option is enabled, you first have to choose the name of a document corpus in a subordinated menu.</p> <p>For write processes in the mirrored document corpus all changes will be applied to the mirroring document corpus.</p> <div style="background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p><b>Note</b> This applies for import-/delete- and update procedures.</p> </div> <div style="background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p><b>Note</b> This option should only be used for the migration of a document corpus. For additional information on a migration see: <a href="#">Assistant for the reindexing of a document corpus</a></p> </div>
<b>Main memory per SHAFTE engine</b>	The size of the main memory (in KB) for the caching of data. This is used for transactions or search queries by d.3 search engine (SHAFTE).
<b>Number of search cursors</b>	This value thus describes, how many queries are simultaneously processed by the d.3 search engine (SHAFTE). If this number is exceeded, further requests have to wait until the d.3 search engine (SHAFTE) releases a busy search cursor.

After the successful configuration click on **Next >>** to go to the next page.

### Step 3: Indexing

#### Note

Consider that the search index has to be regenerated, if an existing document corpus is changed retroactively.

In the third step specify how the documents are to be added to the search index.

The following settings are available:

Setting	Description
<b>Numeric values</b>	If this setting is enabled, then numeric information of the documents will also be added to the index.
<b>Number of word occurrences</b>	<p>This value defines, when a word is classified as "stopword" and thus will no longer be indexed.</p> <p>All documents containing the word are counted.</p> <div style="background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p><b>Note</b> This threshold can be specified because a search for a word, which is contained in more than 1.000.000 documents makes no sense. On the one hand a result set of 1.000.000 would be an imprecise search for the user and on the other hand the joining of more than 1.000.000 results takes more resources.</p> </div>
<b>Minimum word length</b>	Specifies the minimum length of a word to be included in the index.

Setting	Description
<b>StoreText</b>	Specifies, if the full storage of the transmitted document is to be enabled in the index.  <div style="background-color: #ffe6e6; padding: 5px;"><b>Warning</b> If you disable this function, the following functions of the full-text search engine can not be used:  <ul style="list-style-type: none"> <li>• Semantic search</li> <li>• Phrasal search</li> <li>• Creation of text summaries (abstracts)</li> <li>• Reindexing via the <code>SHAFTUTL.EXE</code></li> </ul> </div> <div style="background-color: #e6f2ff; padding: 5px;"><b>Note</b> The full storage of the text content causes an increased use of disk space.</div>
<b>Only keep the latest version in the index</b>	This option specifies, if more than one version of a document is to be maintained in the index. If you want to clean up the index from the old document versions, the index must be recreated with a reindexing. It is possible to switch this option at runtime, however the behavior then only affects new documents.  <div style="background-color: #ffe6e6; padding: 5px;"><b>Warning</b> For an existing document corpus this value must only be enabled retroactively. Disabling it would have the effect that documents can no longer be found.</div> <div style="background-color: #ffe6e6; padding: 5px;"><b>Warning</b> This option requires the d.3 server version 7.2.1!</div>

After the successful configuration click on **Next >>** to go to the next page.

### Schritt 3: Indexierung

Geben Sie an, welche Textinhalte im Suchindex aufgenommen werden sollen.  
Beachten Sie, dass bei einer nachträglichen Änderung der Suchindex neu aufgebaut werden muss.

Numerische Werte	<input checked="" type="checkbox"/>	<span style="border: 1px solid red; padding: 2px;">i</span>
Minimale Wortlänge	<input type="text" value="2"/>	<span style="border: 1px solid red; padding: 2px;">i</span>
Vollständige Speicherung der Textinhalte	<input checked="" type="checkbox"/>	<span style="border: 1px solid red; padding: 2px;">i</span>
Nur die neueste Version im Index halten	<input checked="" type="checkbox"/>	<span style="border: 1px solid red; padding: 2px;">i</span>

### Step 4: Search

This step allows you to configure the search behavior of the d.3 search engine (SHAFT).

The following settings are available:

Setting	Description
<b>First wildcard position</b>	This defines the character position allowed for the first wildcard in a word.
<b>Auto-correction</b>	Specifies, if it should be searched for a similar word automatically, if a search query returned no results.
<b>Sort by relevance</b>	Specifies, if for each result a relevance value shall be calculated and if the result set shall be sorted by the relevance values.

Setting	Description
<b>Advanced search</b>	Specifies, if the search is to be extended by a wildcard search if the search yields too few matches.
<b>Threshold for the advanced search</b>	Specifies, from which number of matches a wildcard search is to be conducted.

After the successful configuration click on **Next >>** to go to the next page.

### Step 5: SHAFT settings

In this step you can create multiple d.3 search engines (SHAFT) for the document corpus.

These can be distributed immediately or later to separate machines. Thus, it allows you to prepare the document corpus for a later scaling at this point.

The scaling can be enabled by setting the parameter **Enable scaling**. Afterwards a new input field appears where you can enter the number of engines to be created.

#### Note

A search index should have a maximum size of about 40 GB depending on the computing capacity. This is equivalent to 30 Mio. documents (depending on the size of the documents).

For each d.3 search engine (SHAFT) you can define the following settings:

Setting	Description
<b>Host-name</b>	Specifies the hostname of the computer on which the d.3 search engine (SHAFT) is to be executed.  <div style="background-color: #ffe6e6; padding: 5px;"> <p><b>Warning</b> The d.3 SHAFT Agent must be installed on this computer. Additional information can be found here: <a href="#">d.3 SHAFT Agent</a></p> </div>
<b>Port</b>	Specifies the TCP/IP port under which the d.3 search engine receives the requests.  <div style="background-color: #e6f2ff; padding: 5px;"> <p><b>Note</b> Make sure to configure different ports, if your system has more than one active d.3 search engine process.</p> </div>
<b>Backup directory</b>	Specifies the storage location for the online backup.  <div style="background-color: #e6f2ff; padding: 5px;"> <p><b>Note</b> To avoid negative effects on the execution of search queries during the time of the backup, this directory should be on a separate hard disk.</p> </div>

Setting	Description
<b>Database directory</b>	Specifies the directory where the d.3 search engine (SHAFT) saves its full-text data. This working directory later contains the full index.
<p><b>Note</b></p> <p>As a rule of thumb, the uncompressed d.3 search engine (SHAFT) index requires 1.4 times the size of the text to be indexed. This factor can be decreased to under 1, if the storage directory is located on an NTFS partition and you enabled the compression option under Microsoft Windows for this directory.</p>	
<p><b>Warning</b></p> <p>If you enable the compression under Microsoft Windows, this reduces the indexing speed. Thus, it is not recommended to enable the compression. The d.velop AG does not explicitly test the compression scenarios.</p> <p>If a virus scanner is monitoring the directory, this reduces the indexing speed. To receive the best-possible performance, the directory should thus be defined as an exception in the virus scanner.</p>	

## Warning

Consider that the number of d.3 search engines (SHAFT) cannot be changed later. A later scaling is only possible with a parallel reindexing into a new scaled document corpus.

After the successful configuration click on **Next >>** to go to the next page.

## Assistent zur Anlage einer Dokumentensammlung

Willkommen   Basis   Indexierung   Suche   **SHAFT-Einstellung**   Backup   Fertigstellen

---

Schritt 5: SHAFT-Einstellung

---

Sie haben an dieser Stelle die Möglichkeit mehrere SHAFT-Engines für eine Dokumentensammlung zu erzeugen. Diese können sofort oder später auf separaten Rechnern ausgelagert werden, um die Performance zu erhöhen.

Bitte beachten Sie, dass die Anzahl der d.3 search engines (SHAFT) im Nachhinein nicht geändert werden kann. Eine nachträgliche Skalierung ist nur durch eine parallele Nachverschlagerung in eine neue skalierte Dokumentensammlung möglich.

Skalierung aktivieren

SHAFT-Engine 1

---

Rechnername  ⓘ

Port  ⓘ

Backup-Verzeichnis  ⓘ

Datenbank-Verzeichnis  ⓘ

<< Zurück Weiter >>

## Step 6: Backup

This step allows you to configure the backup behavior of the d.3 search engine (SHAFT).

Setting	Description
<b>Journaling</b>	Specifies, if a journal file is to be written or not. The journaling has the effect that changes to the database are first written to a journal file. Using this journal file the index files can be maintained in a consistent state after a process termination.
<b>Time for backup</b>	Specifies the time when the backup will be executed.
<b>Check database before backup</b>	This specifies, if the index is to be checked prior to a backup at all.
<p><b>Note</b></p> <p>The consistency check prior to a backup makes sense and should be executed! Only disable this option, if it is not arguable due to time reasons.</p>	
<b>Day for backup</b>	Specifies the day of the month on which the backup will be executed.

After the successful configuration click on **Next >>** to go to the next page.

## Assistent zur Anlage einer Dokumentensammlung

Willkommen   Basis   Indexierung   Suche   SHAFT-Einstellung   **Backup**   Fertigstellen

### Schritt 6: Backup

Ändern Sie die Parameter, um das Backupverhalten der SHAFT-Engine einzustellen.

<b>Journaling</b>	<input checked="" type="checkbox"/>	
<b>Online Backup</b>		
▶ <b>Zeit für Backup</b>	<input type="text" value="00:00"/>	
▶ <b>Datenbank vor Backup prüfen</b>	<input checked="" type="checkbox"/>	
▶ <b>Tag für Backup</b>	<input type="text" value="Jeden Tag"/>	

### Step 7: Complete

In the final step, you can finish the document corpus. The just configured document corpus will be created and started.

If there is a cluster, then the document corpus will be distributed to the cluster and will also be started on the other instances.

With Google Chrome browser you can download already at this point the settings of the document corpus in JSON format. These settings can be used as backup or for an installation.

With Microsoft Internet Explorer this function is only provided after a click on **Finish**.

## Assistent zur Anlage einer Dokumentensammlung

Willkommen Basis Indexierung Suche SHAFT-Einstellung Backup **Fertigstellen**

### Schritt 7: Fertigstellen

Klicken Sie auf "Fertigstellen", damit der d.3 search Assistent die gewählte Konfiguration vornimmt.  
Einstellungen im JSON-Format herunterladen.

<< Zurück

Fertigstellen

## Assistant for the generation of a document corpus

Welcome Basics Indexing Search SHAFT settings Backup **Finish**

### Step 7: Finish

Document corpus has been successfully created.

This page is automatically refreshed after 11 seconds. Click on this hyperlink to refresh the page immediately.

[Download settings in JSON-format.](#)

#### Note

If d.3 search is used in cluster operation, after the creation of the new document corpus the cluster turns inactive for a moment to synchronize itself.

This can be recognized by the small red warning triangle  next to the menu option cluster.

After a successful creation of the document corpus the following confirmation text appears. You can download the settings in JSON format at this point.

These settings can be used as backup or for an installation.

After 15 seconds this page will be reloaded automatically so that the document corpus can be found in the settings.

### Semantic search

For the semantic search correlations of the requested search term to other documents and terms will be created. The individual found terms will be evaluated for analysis procedures by d.veLop and then returned as so called hints.

**Note**

If a user searches for the word "d.velop", the search engine returns the words "Gescher" and "ECM". This results from the fact that these three words are often used together in letterheads.

As the resources of a computer are finite and timely restricted, the analysis set of the created statistics must be restricted.

The analysis will be restricted by a maximum number of documents (`MaxDocsForHints`). The selection of the documents can be affected by the parameter `UseMostRecent`. If this option is enabled, initially documents will be used which have a high relevance, else the documents will be selected from the result list even distributed. By the setting `MaxWordsForHints` it is defined, how many for the document most relevant words per document will be considered maximal.

**Warning**

The higher the values `MaxDocsForHints` and `MaxWordsForHints` will be chosen, the more computing capacity is required for a semantic search.

Additionally, the parameter `WordRelvalue` can define that words, which are in too many documents, will be filtered. This is justified therefore, that these words get a high relevance only because they are in many different documents but their often occurrences have no value for the search.

If now the option `SemanticExtend` is enabled, the search for hints will be extended automatically with the hint terms.

**Note**

If it is searched for hints and additionally the option `SemanticExtend` is enabled, then the advanced search will effect the results, that there will be documents in which the original searched word does not exist.

**Note**

The search for "d.velop" finds e.g. the terms "Gescher", "Software company", "Schildarpstraße" and "d.forum" as word cloud.

The semantic extension of the search searches also for the terms in the word cloud. This has the effect that more matches will be found.

**Note**

The automatic semantic extension is in most cases only for a clean search useful. This means a user wants to read up on the d.velop AG. He wants a direct result for the d.velop AG but he is also interested in the town and possibly other software companies.

For a targeted search the automatic semantic extension is mostly misleading for the user.

## Warning

If it is searched in a document type for a enabled semantic search, then the semantic extension will not search within this document type.

This means for the search for "d.velop" within the document type "Contacts" the terms from the word cloud, e.g. "Software company", will be searched in other document types.

This has the effect, that the d.3 server must filter many results retroactively.

Furthermore, the enabled automatic semantic extension can lead to more results in the refinement of the search, e.g. "d.velop software company Schildarpstraße", as the search for the word "d.velop". This is depending on the word cloud, which is used for the semantic extension.

## Configuring the d.3 search engine

On starting a d.3 search engine process (`shaft.exe`), the file name of a configuration file is passed as a start parameter (usually `shaft.ini`).

This file contains the following parameters partially requested when [creating a document corpus](#):

Section	Parameters	Meaning									
[Engine]											
	Archive-Logs	<p><a href="#">Store journals</a></p> <p>(1: yes; 0: no)</p> <p>Default: 1</p> <p>Prerequisite: Journal=1</p> <p>This has the effect that every journal is stored in the sub-directory ARCHLOGS, to allow restoring the database in case of emergency (<code>SHAFTUTL {INI-file} -i</code>).</p> <p>The journal files can be backed up with appropriate software.</p>									
	AutoCorrect	<p><a href="#">Automatic correction of unknown expressions</a></p> <p>(0: disabled), (1: enabled)</p> <p>Default: 1</p> <p>If the <b>Automatic correction of unknown expressions</b> is enabled, then search queries that would yield no match will automatically use a term with similar spelling and search for it.</p>									
	Cache-Size	<p><a href="#">Disk space for the caching of data</a></p> <p>Default: 256000</p> <p>This value in KB specifies, how much main memory should be used for the temporary storage of data.</p> <p>Depending on whether the engine is operated with 32-bit or 64-bit, the cache can be configured with different sizes:</p> <table border="1"> <thead> <tr> <th></th> <th>Shaft 32 Bit</th> <th>Shaft 64 Bit</th> </tr> </thead> <tbody> <tr> <td>Microsoft Windows 32 Bit</td> <td>max. 1.5 GB</td> <td>-</td> </tr> <tr> <td>Microsoft Windows 64 Bit</td> <td>max. 3.5 GB</td> <td>max. 16384 GB</td> </tr> </tbody> </table> <p>The cache should not be configured larger than the main memory - 1 GB.</p> <p>For virtual hard disks, the cache should be larger than or of the same size as the index sized.</p> <p>If the cache is too small and this results in some requests taking longer than five seconds, then the engine creates a <code>Perflogs</code> directory and writes logs about the performance of the request there. Among other things, this specifies a percentage value for the cache hit ratio.</p> <p>If this value is lower than 98%, the cache size should be increased.</p>		Shaft 32 Bit	Shaft 64 Bit	Microsoft Windows 32 Bit	max. 1.5 GB	-	Microsoft Windows 64 Bit	max. 3.5 GB	max. 16384 GB
	Shaft 32 Bit	Shaft 64 Bit									
Microsoft Windows 32 Bit	max. 1.5 GB	-									
Microsoft Windows 64 Bit	max. 3.5 GB	max. 16384 GB									

Section	Parameters	Meaning
	Commit-Count	<p>Default: 200</p> <p>This specifies after how many insert operations the index on the hard disk is to be updated.</p>
	Database	<p><a href="#">Storage of index data</a></p> <p>Path to the d.3 search engine database.</p> <p>This path must be on a local disk. Network paths are not accepted by the search engine.</p>
	Logpath	<p>Path to the d.3 search engine transaction log.</p> <p>To improve the performance of the d.3 search engine and to decrease the load on the drive hosting the database, this path should be changed to a different drive.</p>
	Delimitermode	<p>This parameter is by default not entered in the <code>shaft.ini</code>.</p> <p>Standard: for old indexes: 0, for new indexes: 1</p> <p>0: Standard word separation, i.e. all characters except letters or numerics are interpreted as delimiters</p> <p>1: <code>;;_,-/\</code> are not interpreted as delimiters but they are truncated at the beginning and end of words</p> <p>2: <code>;;_,-/\</code> are not interpreted as delimiters but removed from words</p> <p>3: A file named <code>delimiter.txt</code> is expected in the <code>Corpus</code> directory which contains the separators.</p> <p>This file contains one separator per line. You can also define multi-byte delimiters (e.g. UTF8). To do so, simply enter the desired characters in a line.</p> <p>The delimiter mode as well as the <code>delimiter.txt</code> must not be changed at runtime. If the delimiter mode or the <code>delimiter.txt</code> are changed, the index must be rebuilt.</p> <p>A sample <code>delimiter.txt</code> file can be found in the directory <code>Vorlagen</code> (= templates).</p> <p>If <code>Stortext</code> is enabled, this can be executed using <code>shaftutl</code> and the parameter <code>-ri</code>. Else, the index must be entirely rebuilt.</p> <p>4: The BreakIterator of the Opensource component ICU (International Components for Unicode) is used. This mode is required, if, for example, Asian languages are to be used which can only be delimited via dictionaries.</p> <div style="background-color: #ffe6e6; padding: 10px; margin: 10px 0;"> <p><b>Warning</b></p> <p>In this delimiter mode, d.3 search must be operated in Unicode mode, as the Asian language cannot be processed otherwise.</p> </div> <div style="background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p><b>Note</b></p> <p>Additional information on the component ICU can be found under <a href="http://www.icu-project.org">http://www.icu-project.org</a> and <a href="#">Internationalization</a>.</p> </div>
	IcuExcludedDelimiter	<p>This parameter only has a function in combination with the parameter <code>DelimiterMode=4</code>. It specifies, which characters usually interpreted as delimiters by the ICU, are not to be interpreted as such.</p> <p>Default: -</p> <p>Optionally specify here, which characters usually interpreted as delimiters by the ICU, are not to be interpreted as such. If the parameter is not specified, the hyphen/minus character is not interpreted as a delimiter so that hyphenated words are not stored separately. This is especially advisable for product descriptions which often contain many hyphens. If the hyphen is not excluded, the engine would conduct an AND search for two separate words for hyphenated words and possibly find more matches than desired.</p>

Section	Parameters	Meaning
	First- Wildcard	<p><a href="#">Earliest position of wildcard</a></p> <p>Default: 3</p> <p>This defines the character position allowed for the first wildcard in a word.</p> <p>The allowed wildcards are "?" or "*".</p> <p>This parameter is by default set to "3". If this parameter is set to "0" or "1", then the engine creates an additional nGramm index (the files <code>nGramm.idx</code> and <code>nGramm.dat</code>).</p> <p>If the index already contains data, the index is rebuilt after a restart and is then filled with the existing data.</p> <p>This index slows down the import and requires additional disk space but allows the engine to avoid comparing all words in searches with leading wildcards followed by at least three characters.</p> <p>This way, such search queries can be answered significantly faster.</p> <p>If this value is then set to 1 again later, the nGram files are automatically deleted again and the hard disk memory is vacated.</p>
	Journal	<p><a href="#">Write journals (Redo-Log)</a></p> <p>(1: yes; 0: no)</p> <p>Default: 1</p> <p>The value "1" enables the writing of a journal file. The journaling has the effect that changes to the database are first written to a journal file. Using this journal file the index files can be maintained in a consistent state after a process termination.</p>
	LogLevel	<p>Default: 9</p> <p>Defines the height of the loglevel for the transaction-log.</p> <p>With "0" only errors are logged. With "9" everything is logged including the texts passed when inserting into the database.</p> <p>In connection with online backup, this parameter can be easily set to "9" as the log is automatically deleted.</p> <p>If no online backup is used, the directory grows quickly.</p>
	Cursor	<p><a href="#">Max. number of search cursors</a></p> <p>Default: 64</p> <p>The parameter defines the number of search queries to be maintained in the main memory. This value thus describes, how many queries are simultaneously processed by the d.3 search engine.</p> <p>You must enter a value &gt; "0".</p>
	MaxDocs- ForHints	<p>Default: 100</p> <p>Specifies the maximum number of documents in a result set to be considered for the semantic analysis.</p> <p>The parameter <code>MaxDocsForHints</code> is explicitly set to "0", if no <a href="#">semantic analysis</a> is desired. Else, the parameter does not appear in the file.</p>
	Max- WordsFo- rHints	<p>Default: 100</p> <p>Specifies the maximum number of words per document in a result set to be considered for the semantic analysis.</p> <p>The respective most relevant terms are used for each document.</p>
	Numeric	<p><a href="#">Index numeric values including date information</a></p> <p>(1: yes; 0: no)</p> <p>Default: 1</p> <p>Has the effect that numeric information in the document is also indexed in the database.</p>

Section	Parameters	Meaning
	Order-byrank	<p><a href="#">Sort search results by relevance</a></p> <p>(1: enabled), (0: disabled)</p> <p>Default: 1</p> <p>If this option is enabled, then a relevance value is calculated for each document with reference to the current query. Based on this relevance, the result set is then sorted so that the document with the highest relevance is delivered first.</p>
	Port	<p><a href="#">TCP/IP Port for d.3 search</a></p> <p>Default: 3541</p> <p>TCP/IP Port under which the d.3 search engine receives the requests.</p> <p>Make sure that all locally running processes use another port.</p>
	SemanticExtend	<p><a href="#">Extend search terms before search</a></p> <p>(1: enabled), (0: disabled)</p> <p>Default: 0</p> <p>If you enable this option then the semantic cloud is included in the search. This means, the semantic cloud is not only displayed to the user afterwards but the terms are searched immediately, thus increasing the number of results. Depending on their distance to the actual search, the words from the cloud get a lower weight for the calculation of the ranking.</p> <p>Additional information can be found here: <a href="#">Semantic search</a></p>
	StoreText	<p><a href="#">Full storage of text contents</a></p> <p>(1: enabled), (0: disabled)</p> <p>Default: 1</p> <p>StoreText=1 enables the full storage of the text transmitted by the document in the index.</p> <div style="background-color: #ffe6e6; padding: 10px; margin: 10px 0;"> <p><b>Warning</b></p> <p>If you disable this function, the following functions of the full-text search engine can not be used:</p> <ul style="list-style-type: none"> <li>• Phrasal search</li> <li>• Creation of text summaries (abstracts)</li> <li>• Reindexing via the <code>SHAFTUTL.EXE</code></li> </ul> </div> <div style="background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p><b>Note</b></p> <p>The full storage of the text content causes an increased use of disk space.</p> </div>
	UseMostRecent	<p><a href="#">Semantics based on the relevance</a></p> <p>This controls the selection of documents from the result set to be included in the semantic analysis.</p> <p>1: Uses at first the documents with highest relevance to the semantic analysis.</p> <p>0: Determines the documents based on even distribution</p> <p>Default: 1</p>
	WordRelvalue	<p>Default: 150</p> <p>Relative relevance value of a word which must not be under-run.</p> <p>If this value is exceeded for a word, then it is removed from the search suggestions and replaced by another one.</p> <p>The value should be set to "150", if the parameter <a href="#">MaxWordsForHints</a> is set to "100".</p>
	MinWordLen	<p>Default: 2</p> <p>Specifies the minimum length of a word to be included in the index.</p> <p>To remove this limitation, the parameter can be set to "0".</p>

Section	Parameters	Meaning
	AutoRe-pair	<p>Default: 0</p> <p>1: If an error is identified during the backup, then this is resolved automatically. This is only possible, if a transaction log is available and an online backup exists.</p> <div style="background-color: #ffe6e6; padding: 10px; margin: 10px 0;"> <p><b>Warning</b></p> <p>During the defragmentation, additional hard disk memory in the size of the current index is required. The hard disk memory must be available on hard disk where the index is stored.</p> </div>
	WildCardLimit	<p>Default: 1000000</p> <p>Specifies, after how many lines the file is to be split.</p> <p>This means, as soon as the engine has found 1,000,000 matches, it stops searching even if there are other relevant matches.</p>
	AutoDefrag	<p>(1: enabled), (0: disabled)</p> <p>Default: 1</p> <p>This parameter specifies, if the index should be defragmented automatically as part of a backup. The defragmentation is only performed, if the fragmentation of the index exceeds 15%. Depending on the size of the index, a defragmentation may take a while.</p> <div style="background-color: #ffe6e6; padding: 10px; margin: 10px 0;"> <p><b>Warning</b></p> <p>During the defragmentation, additional hard disk space in the size of the files <code>link.dat</code> and <code>words.dat</code> is required. The hard disk memory must be available on hard disk where the index is stored.</p> </div>
[PerfLogs]		
	Abstract	<p><a href="#">Notes on document collections with the d.3 search engine</a></p> <p>Default: 5000</p> <p>Time in milli-seconds a threshold for the execution of the operation <code>Abstract</code>. If the threshold is exceeded, a logfile is written in the directory <code>..\d3\d.3 search\corpora\[CorpusId]\PerfLogs</code>.</p>
	Delete	<p><a href="#">Notes on document collections with the d.3 search engine</a></p> <p>Default: 5000</p> <p>Time in milli-seconds as a threshold for the execution of the operation <code>Delete</code>. If the threshold is exceeded, a logfile is written in the directory <code>..\d3\d.3 search\corpora\[CorpusId]\PerfLogs</code>.</p>
	Search	<p><a href="#">Notes on document collections with the d.3 search engine</a></p> <p>Default: 5000</p> <p>Time in milli-seconds as a threshold for the execution of the operation <code>Search</code>. If the threshold is exceeded, a logfile is written in the directory <code>..\d3\d.3 search\corpora\[CorpusId]\PerfLogs</code>.</p>
	Insert	<p><a href="#">Notes on document collections with the d.3 search engine</a></p> <p>Default: 5000</p> <p>Time in milli-seconds as a threshold for the execution of the operation <code>Insert</code>. If the threshold is exceeded, a logfile is written in the directory <code>..\d3\d.3 search\corpora\[CorpusId]\PerfLogs</code>.</p>
	Spell-check	<p><a href="#">Notes on document collections with the d.3 search engine</a></p> <p>Default: 5000</p> <p>Time in milli-seconds as a threshold for the execution of the operation <code>Spellcheck</code>. If the threshold is exceeded, a logfile is written in the directory <code>..\d3\d.3 search\corpora\[CorpusId]\PerfLogs</code>.</p>
[Backup]		

Section	Parameters	Meaning
	<p>CheckDay</p> <p><a href="#">Perform check ...</a></p> <p>Default: 0</p> <p>This option specifies on which day the backup should be checked.</p> <p>0: everyday</p> <p>1: Sunday, 2: Monday, etc.</p>	<p><b>Note</b></p> <p>The archive logs and the transaction logs are only deleted after a successful check, to be able to restore the database, if the backup was not consistent!</p> <p><b>Note</b></p> <p>During the backup search queries can still be processed. Changes can not be done at this time.</p>
	<p>CheckDB</p> <p><a href="#">The database is checked before a backup</a></p> <p>Default: 1</p> <p>This specifies if the index is to be checked prior to a backup at all.</p>	<p><b>Note</b></p> <p>The check ensures that the backup is consistent afterwards. Disabling this function leads to a faster backup but it does not ensure you that the backup is consistent.</p>
	<p>Path</p> <p><a href="#">Storage in:</a></p> <p>Storage location for the online backup</p>	<p><b>Note</b></p> <p>To avoid a unnecessary and heavily use of the hard disk during the backup, it is recommended to store the backup directory on a separate hard disk.</p> <p><b>Warning</b></p> <p>If this value is not defined, no backup will be executed. Without backup the following functions are not available:</p> <ul style="list-style-type: none"> <li>• Automatic defragmentation</li> <li>• Restoring a database</li> <li>• Regular check of the database</li> </ul> <p>Additionally a reindexing must be executed for a damaged database via the d.3 server.</p>
	<p>Time</p> <p><a href="#">Time</a></p> <p>Default: 00:00</p> <p>Time at which the online-backup is to be performed</p>	<p><b>Note</b></p> <p>It's advisable to choose a time when only a few employees are working with the d.3 system.</p>

**Note**

Example of a `shaft.ini`:

```
[Engine]
Database=C:\d3\d.3 search\corpora\D3P
Port=3541
Journal=1
MaxFilesPerWord=10000000
ArchiveLogs=1
Numeric=1
CacheSize=256000
PhysicalCache=1
MaxCursor=10
WordRelValue=150
MaxWordsForHints=100
FirstWildcard=3
AutoCorrect=1
StoreText=0
Orderbyrank=1
MaxDocsForHints=0
SemanticExtents=0
UseMostRecent=0
LogLevel=9
[Backup]
Path=C:\d3\d.3 search\corpora\D3P\backup
Time=00:00
CheckDB=1
CheckDay=0
```

**Note**

An extract from of a `stopwords.txt` file (this file is located in the sub-directory `corpora \ &lt;d.3 search engine Name&gt;`).

```
im
aber
als
am
an
auch
auf
aus
bei
bin
bis
bist
da
```

**Management of the `stopwords.txt`**

The `stopwords.txt` can be edited via the web interface of the d.3 search instance under the menu option **Settings | Document corpus | <Corpus-ID> | stopwords.txt**.

Via the web interface the `stopwords.txt` is stored in the database directory of the respective d.3 search engines (SHAFT).

Each line represents a stop word. Only after a restart of the d.3 search engine (SHAFT) the changes take effect.

The entry of a word in the stop word list has the following effects:

- The word is not included in the search index for new documents.
- The word is removed from the search query during a search. This means that the stop word also applies to documents that have already been imported.

### Note

#### Example: Ignoring articles

In a search query, articles such as "the" should not be considered, since the articles are negligible in the search.

Stop word:

- the

Search query:

- "Managing the stop words"

Result:

- All documents containing the words "Managing" and "Stop words".
- Among the hits there can be documents WITHOUT "the"

#### Good extensions for the stop word list:

- Words that are not meaningful for a search because they appear in every document.
- For example the very own company name that is contained in every document.

### Warning

The content of the `stopwords.txt` MUST be appended in lower-case letters. This also applies to umlauts.

### Note

If the engine is running in UTF-8 mode, the file must be UTF-8 encoded when it is saved.

To find out, if the d.3 search document corpus is running in the UTF-8 mode, check the `corpus.ini`.

## Management of the `synonym.txt`

You can maintain a catalog of synonyms. This is designed to improve search queries by adding further terms to the search entries by a user, which are maintained in the `synonym.txt`.

### Note

`Synonym.txt` :

Auto Car Vehicle

If, for example, the user is looking for the term Auto, then his search could be automatically extended by the terms Car and Vehicle.

**Note**

If a synonym catalog is maintained, no automatic correction of the search query takes place.

The `synonym.txt` can be edited via the web interface of the d.3 search instance under the menu option **Settings | Document corpus | <Corpus-ID> | synonym.txt**.

One line per term is used. There are two options to create entries:

1. An entry in the catalog looks like this. **Auto Car Vehicle**  
In this case, all three terms are searched, disregarding which one the user has entered.
2. An entry in the catalog looks like this. **Auto = Car Vehicle**  
In this case, all three terms are only searched, if the user has entered the word Auto.

Restart the respective engine so that the settings can take effect.

**Management of the validwords.txt****Warning**

If this catalog is adjusted, the changes only apply to newly indexed documents. The index must be indexed again, if the changes shall take effect retroactively.

More details can be found in chapter [Reindexing a document corpus](#).

You can maintain a catalog with `validWords`. These are for terms which shall not be separated at the text analysis by the search engine, this means these can be found without consideration of the normal delimiter in the search.

**Note**

```
validwords.txt:
```

```
d.velop!AG
```

```
d-velop.de
```

The term `d.velop!AG` will not be separated in three words (`d` and `velop` and `AG`) but is known by the search engine as `d.velop!AG!`

The `validwords.txt` can be edited via the web interface of the d.3 search instance under the menu option **Settings | Document corpus | <Corpus-ID> | validwords.txt**. One line per term is used.

Restart the respective engine so that the settings can take effect.

**Note**

If the `validwords.txt` can not be found at the start of the engine, then the recognition of `ValidWords` will be disabled due to performance reasons. You can find a respective entry about this in the d.3 log.

**Advanced configuration of a document corpus**

For every document corpus, d.3 search creates an individual directory with the configuration- and index data for the search engine.

This directory is usually named `d.3 search directory]\corpora\[Corpus ID]` (e.g. `D:\d3\d.3 search\corpora\d3p`).

The main configuration file for the d.3 search server is the file `dsearch.ini` in this directory.

The entries in this file have the following meaning:

Section	Parameters	Meaning
[Mirrors]	CorpusID	<p>By default, this parameter is not available.</p> <p>If a document corpus ID is set for this parameter, then the document corpus will be mirrored with the current one.</p> <div style="border: 1px solid black; background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p><b>Note</b> Only new documents will be mirrored, a synchronization with the old data does not take place.</p> </div>

The configuration file for the document corpus is the file `corpus.ini`.

The entries in this file have the following meaning:

Section	Parameters	Meaning
[Shaft_Global]	UTF8Support	<p>(1: yes; 0: no)</p> <p>Default: 1</p> <p>This value specifies, if the document corpus supports the character encoding UTF8.</p> <div style="border: 1px solid black; background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p><b>Note</b> If your documents in d.3 are UTF8 encoded, you should enable this parameter.</p> </div>
	ForceNo-Hints	<p>(1: yes; 0: no)</p> <p>Default: 0</p> <p>This parameter disables the creation of word clouds for this document corpus for a search.</p>
[ShaftX]	Host	<p>Standard: localhost</p> <p>Specifies the hostname of the computer on which the d.3 search engine (SHAFT) is to be executed.</p>
	Port	<p>Default: 3541</p> <p>Specifies the port under which the d.3 search engine (SHAFT) is responding.</p>
	SHAFTAgent-Port	<p>Default: 6133</p> <p>Specifies the port via which the d.3 SHAFT Agent is available.</p> <div style="border: 1px solid black; background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p><b>Note</b> This setting is only required, if a d.3 search engine (SHAFT) is outsourced to another machine.</p> </div>
[Search]		

Section	Parameters	Meaning
	Expand-Search	<p>(1: yes; 0: no)</p> <p>Default: 0</p> <p>This parameter specifies, if the search is to be extended by a wildcard search if the search yields too few matches.</p> <p>Then, a * is suffixed to each word to increase the number of matches.</p> <div style="background-color: #e6f2ff; padding: 5px; margin: 10px 0;"> <p><b>Note</b></p> <p>If the automatic wildcard search is conducted, no synchronization with the synonym catalog is performed.</p> </div>
	Expand-SearchMin-Result	<p>Default: 20</p> <p>Specifies, from which number of matches a wildcard search is to be conducted. If the value is set to 0, then a wildcard search is always conducted.</p> <p>Please note that a wildcard search is more time consuming and should only be conducted if necessary.</p>
[General]		
	CurrentVersionsOnly	<p>(1: yes; 0: no)</p> <p>Default: 1</p> <p>This parameter specifies, if more than one version of a document is to be maintained in the index.</p> <p>By default, a document in the index is saved in the format &lt;document_ID&gt;.&lt;versions_ID&gt;. If the parameter is enabled (set to 1), then the documents in the index will be saved in the format &lt;document_ID&gt;, so without version.</p> <p>It is possible to switch this parameter at runtime, however the behavior then only affects new documents.</p> <div style="background-color: #ffe6e6; padding: 5px; margin: 10px 0;"> <p><b>Warning</b></p> <p>The parameter must only be set from 0 to 1 during the runtime. Disabling it would have the effect that documents can no longer be found.</p> </div> <p>If the you want to clear the index from the old document versions, the index must be recreated with a reindexing.</p> <div style="background-color: #e6f2ff; padding: 5px; margin: 10px 0;"> <p><b>Note</b></p> <p>All d.3 server processes must be restarted manually so that the setting takes effect.</p> </div> <div style="background-color: #ffe6e6; padding: 5px; margin: 10px 0;"> <p><b>Warning</b></p> <p>This parameter requires the version 7.2.1 of the d.3 server.</p> </div>
	UseFileStream	<p>(1: yes; 0: no)</p> <p>Default: 0</p> <p>This parameter specifies, if the imported files are to be read via a FileStream. This has the advantage that only parts of the file are loaded into the main memory. However, enabling this option makes the import slower.</p> <p>It is recommended to enable this option only, if the main memory is not sufficient when reading large files.</p>

## Deleting a document corpus

### Warning

To delete a document corpus means that to remove it without any chance of restoring, including configured backup directory.

For large document corpora a regeneration via the d.3 server can take up to several weeks.

Consider this and use this command with care!

### Warning

If d.3 search is used in cluster operation, this command will be forwarded automatically to all cluster members. This leads to the deletion of the document corpus on every node of the cluster array!

The deletion of a document corpus can be started via the web interface. Open the web interface of a d.3 search instance. This can be found by default under the address <http://localhost:6133> or you can open the link `d.3 search webinterface` from the d.3 search start menu.

Ammend the existing URL with the command `/deleteCorpora?Corpora=<Name of the document corpus>`.

### Note

A possible URL for deleting a document corpus would be:

```
http://localhost:6133/deleteCorpora?Corpora=d3A
```

## Renaming a document corpus

### Warning

Renaming a document corpus can have the effect that the assignment document type to document corpus in the d.3 admin is not correct anymore.

Consider this and use this command with care!

### Warning

If d.3 search is used in cluster operation, this command will be forwarded automatically to all cluster members. This leads to the renaming of the document corpus on every node of the cluster array!

The renaming of a document corpus can be started via the web interface. Open the web interface of a d.3 search instance. This can be found by default under the address <http://localhost:6133> or you can open the link `d.3 search webinterface` from the d.3 search start menu.

Ammend the existing URL with the command `/renameCorpus?OldCorpusname=<Name of the existing document corpus>&?NewCorpusname=>Name of the new document corpus>`.

**Note**

A possible URL for deleting a document corpus would be:

```
http://localhost:6133/renamecorpus?OldCorpusName=d3Old&NewCorpus-Name=d3New
```

**Configuration of the d.3 system**

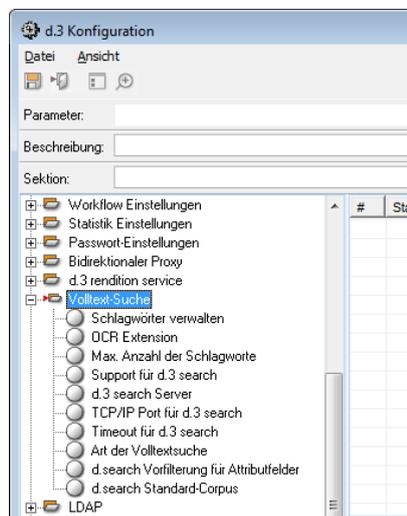
Follow the next steps for the configuration of the d.3 system:

1. [Configuration in the d.3 admin under d.3 config](#)
2. [Configuration in the d.3 admin under d.3 search](#)

**Configuration in the d.3 admin under d.3 config**

To configure the connection of d.3 search to d.3, various parameters must be set in the d.3 configuration via the d.3 admin.

These parameters can be found in the section **Full-text**.



Among them are:

**Keyword management (KEYWORD\_OPTION)**

Default value: 0

The following options are available for the keyword management:

0	no words are adopted in the keyword index on import
1	only the attributes from the comment field are adopted in the keyword index.
2	the document attributes and the words extracted via OCR are inserted into the keyword index
3	the document attributes and the words from the note file are inserted into the keyword index
4	the document attributes, its content (OCR) and the note file are inserted into the keyword index

The keyword index is used for the full-text search

**Note**

If d.3 search is used, this value must at least be set to "1".

**OCR Extension (ocr\_ext)**

Default value: OCR

Ending identifying the d.3 keyword-file.

During the batch-import, a full-text indexing is only executed, if a file with the specified ending exists and the configuration parameter `KEYWORD_OPTION` is set to "2".

The keyword-file should be a pure ASCII-file without special characters.

Documents created with word processing-applications such as .doc-files with Winword are not supported. They must first be converted, i.e. the formatting information must be removed.

#### **Max. number of keywords (OCR\_MAX\_WORDS)**

Default value: 100

Maximum number of words to be included in every d.3 keyword-file.

This number applies to all document types which can be set in the `d3addon.ini`, nothing else was specified under `ocr_dokuart_max_woerter` for the respective document type.

#### **Support for d.3 search (DSEARCH\_SUPPORT)**

Default value: No

Enabling the use of the optional module d.3 search (external full-text search engines) for d.3.

For the import, the parameter `KEYWORD_OPTION` must at least be set to "1".

#### **d.3 search server (d3fc\_dsearch\_server\_name)**

Default value: localhost

Name of the machine where d.3 search is installed.

Using d.3 search external full-text-engines can be integrated.

#### **TCP/IP Port for d.3 search (d3fc\_dsearch\_server\_tcpip\_port)**

Default value: 3406

TCP/IP-port of the machine where d.3 search is installed.

#### **Timeout for d.3 search (d3fc\_dsearch\_server\_timeout)**

Default value: 30

Maximum time in seconds for which d.3 waits for the result once it sent the request to the d.3 search for synchronous processing.

This parameter prevents a timeout of the d.3 processes, if the processing by d.3 search takes unexpectedly long.

#### **Type of full-text search (DSEARCH\_PLUS\_D3)**

Default value: No

Since the search via the dictionary tables is no longer supported, it is not advisable to enter another value than the default here.

In this case, the search is performed using just d.3 search.

### d.search pre-filtering of property fields (DSEARCH\_DIRECT\_DOCFIELD\_SEARCH)

Default value: Yes

Enables the direct d.3 search-search for the document-property fields.

For a combined full-text- and attribute field-search, this reduces the result list of d.3 search and accelerates the further filtering by the d.3 server.

### d.search Standard-Corpus (DSEARCH\_DEFAULT\_CORPUS)

Default value: -

Is used if no corpus for the document type was explicitly specified; when empty, this value is derived from the d.3 gateway server ID (d3[d3fc\_server\_id]).

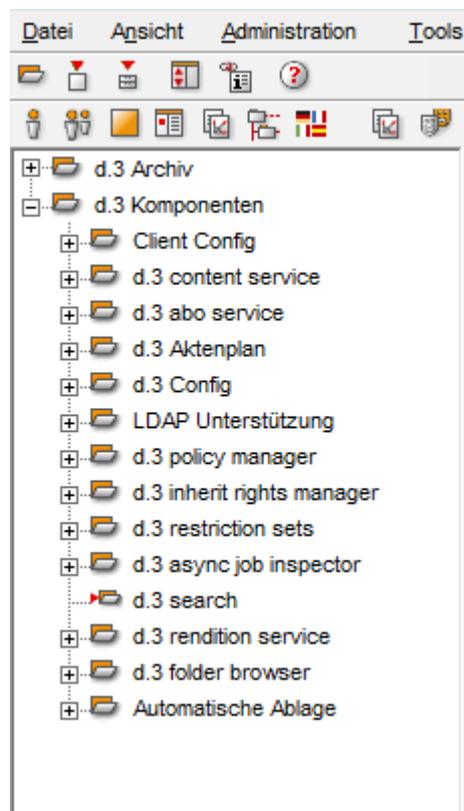
Additional information can be found in the documentation d.3 administration.

More [settings for the d.3 search support](#) can be found in this documentation.

### Configuration in the d.3 admin under d.3 search

If the d.3 search [support for a d.3 archive](#) has been enabled (parameter DSEARCH\_SUPPORT in the **d.3 config** module), you can configure the following settings in the **d.3 search** module of the d.3 administration:

- the d.3 search server status
- the assigned document types
- the properties of the document corpora
- Open the d.3 admin and the choose **More services** with the option **d.3 search**.



The dialog of the d.3 search administration is divided into the three tabs [Server status](#), [Document types](#) and [Document corpora](#).

### The tab **Server status**

The tab **Server status** displays the current properties of the d.3 search server. This view is automatically refreshed every second.

As every update means a query as seen by the d.3 search server, the number of processed requests is also increased by one every second.



The line **d.search server** specifies which d.3 search server processes the d.3 system for the storage of full-text information.

The number in **Current connections** specifies, how many requests have been processed by d.3 search until now (update- or search queries).

The number in **Requests processed** is calculated since the start of the d.3 search server process.

The information on the **Last server start** or the runtime describes how long the d.3 search server process has been active.

In combination with the processed requests and the runtime this allows to determine the system load.

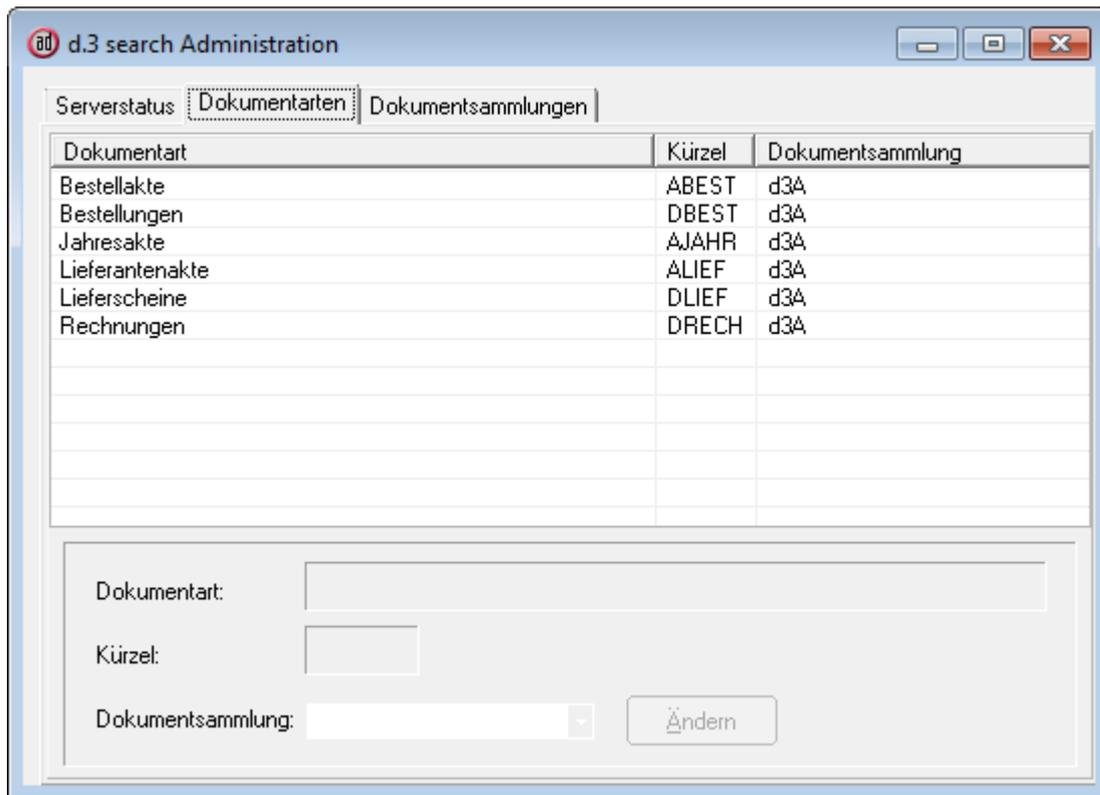
The points **Pending insert operations**, **Pending delete operations** and **Pending updates** specify the number of changes to the document corpora (adding a new document, change or delete a document) that have not yet been processed. Higher values for these numbers signal that a search engine is either not available or d.3 search cannot keep up with the d.3 hostimp in processing the jobs of a batch import.

### The tab **Document types**

This tab allows you to change the mapping between the documents of a document type and the d.3 search document corpus.

The overview displays the "long text" and the ID for all generated document types together with the linked corpus.

The default document corpus assigned by the d.3 server for new document types is determined from the prefix "d3" and the repository ID as the suffix. This means, a repository "DE" would address a default document corpus "d3DE".



- Select a document type to change a mapping of a document type to a corpus

The lower section displays the selected document type with the assigned corpus.

- Choose, if it should remain unindexed or which document type should contain the full-text information for this document type.
- Click **Change** to send your settings to the d.3 server.

### Note

This change affects all documents imported or updated afterwards. The full-text information of the existing documents of this document type should be migrated from the old document corpus into the new one using a d.3 database script.

### Use different document corpora

Basically it is enough to create one document corpus per d.3 repository and have all keywords stored in it. However, there may be circumstances, when it is useful to create different document corpora. Generally you should consider before the configuration, if one or more document corpora are to be created. How many documents do you expect in the future? Are the documents suitable for a full-text search?

Reasons for using several document corpora:

- The document corpus becomes too large
  - With a number of 50 million documents and small file sizes or 30 million documents are large file sizes of the documents, it is advisable to create separate document corpora.
- Using word clouds
  - For some document types, it is not sensible to use word clouds as they contain content that is not intended for the public. Thus, the user should not receive any suggestions in the client pointing to them. All other document types, however, should be included in the suggestions.

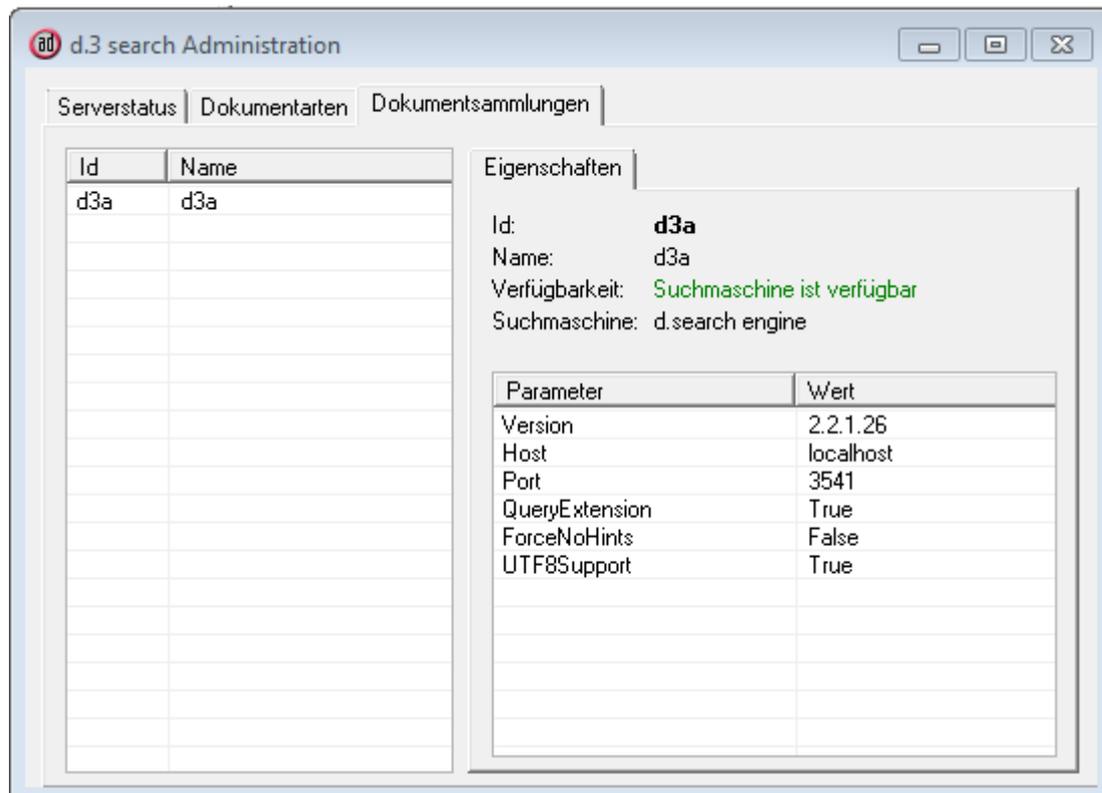
- For some document types, it is not sensible to provide a semantic assistance, as it is impossible to generate useful word clouds from their content.

### The tab Document corpora

The tab **Document corpora** in the d.3 search administration shows an overview of all document corpora available on the d.3 search server.

- Select a document corpus in the list to display additional properties on the right.

The view also contains information if a document corpus can actually be accessed or not (property: **Available**).



## 1.4. Technical background on the d.3 search server

The following chapters will give a technical description of the d.3 server which explains the configuration options for the d.3 server in detail.

### 1.4.1. Directory structure

The default directory of the d.3 search server is `.. \d3\d.3 search`. This contains some sub-directories for temporary working data or configuration files.

**Note**

This could be a list of possible directories:

```

corpora\
corpora\d3p\
corpora\d3p\idbackup
d3index
d3index\d.3 server\
d3index\d.3 server\ID\
d3index\d.3 server\ID\dtiff
d3index\d.3 server\ID\Filter\Break
d3index\d.3 server\ID\Filter\Finish
d3index\d.3 server\ID\Filter\Work
d3index\d.3 server\ID\Process
Database
engines\
shaft
shaft\64Bit
jobs\
jobs\corpusID\add
jobs\corpusID\update
jobs\corpusID\delete

```

**The corpora directory**

You find in this sub-directory the configuration- and index data for individual document corpora.

For every document corpus the d.3 search server creates an individual sub-directory. In this sub-directory of the respective document corpus the following files and folders are located:

File / folder	Meaning
SHAFT1	In this folder you can find at the default settings the search index and the configuration of the d.3 search engine (SHAFT). <div style="background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p><b>Note</b></p> <p>For a scaling a further folder with the name SHAFT2, SHAFT3 etc. will be next to it for each d.3 search engines (SHAFT).</p> </div> <div style="background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p><b>Note</b></p> <p>If the d.3 search engine (SHAFT) is outsourced to another machine, then the folder is without content.</p> </div>
corpus.ini	The configuration file describes the document corpus. Additional information can be found here: <a href="#">Advanced configuration of the document corpus</a> .
dsearch.ini	The configuration file contains information, how d.3 search has to behave with the document corpus. Additional information can be found here: <a href="#">Advanced configuration of the document corpus</a> .

**The directory d3index**

All application directories and files of the d.3 search retroactive indexing are installed in the sub-directory d3index within the d.3 search-application directory.

**The engines directory**

This directory contains the interface modules to the individual search engines. In d.3 search version 2.2 the SHAFT interface (shaft.dll) is located under this directory.

## The directory shaft

Contains the program files for the SHAFT engine d.3 search engine. The directory `64Bit` contains the 64 bit version of the engine.

## The jobs directory

The change jobs for the document corpora are temporarily stored under the Job-directory. This means, the d.3 searchserver e.g. receives the text data after a new import of a document from d.3 server and acknowledges the receipt. The job files remain in the `jobs` directory until the text files are transferred to the search engine.

### 1.4.2. The d.3 search engine

The product d.3 search comes together with the full-text search engine d.3 search engine. This search engine internally operates with the so called SHAFT technology designed by d.velop and is thus often named the SHAFT engine.

The search engine consists of the files `SHAFT.EXE` and `SHAFTUTL.EXE`.

## Notes on updating the d.3 search engine

### Warning

If you are updating a d.3 search version lower than 1.2.4, you must re-index (so-called rebuild) the database.

If this is not done, the database will be marked as “corrupt” in a test and the search engine cannot be started any more.

However, you can then still run the re-indexing.

### Note

The SHAFT engine from version 1.4 is provided as a 32-bit and 64-bit version. If the engine is started from the Shaft directory as in the default, then the 32-bit version started. To switch to the 64-bit version, just change the path to `.\SHAFT\64Bit`.

For this effect, the index does not need to be changed or adjusted. You can switch between 32 and 64-bit at any time.

## Notes on document collections with the d.3 search engine

In order to use a search engine with d.3 search you must create a so-called document corpus (sometimes also called document collection) and you need an instance (that is a running process of the `SHAFT.EXE`) accessible via the TCP/IP-protocol.

This is all done with the d.3 search wizard for the [creation of a document corpus](#).

The user can choose in the wizard which search engine should manage the new document corpus.

In case of the d.3 search engine you can select, if a new or an existing instance of the search engine is to be used, or if a new search engine instance should be created for the new document corpus.

### Note

In the latter case you should make sure to create database for the storage of the index data of the search engine instance on a local drive or an equivalent drive.

**Warning**

Please note that several instances of the search engine can run on one machine each managing its own full-text database. Since the TCP/IP protocol is used for the communication between d.3 search and the d.3 search engine, you only have to make sure that each instance of the search engine uses a different port.

**Warning**

The search engine database should never be located on a NAS or something similar such as a NetApp, as this could limit the availability of the database and thus lead to inconsistent data thus resulting in a corrupt database.

Background: Databases of search engines must provide a high performance for response times and should thus always be highly available and not be subject to external caching mechanisms. These mechanisms could decrease performance and also significantly compromise the functionality.

**Performance**

Databases of search engines must provide a high performance for response times and should thus always be highly available and not be subject to external caching mechanisms. These mechanisms could decrease performance and also significantly compromise the functionality.

Benötigt die d.3 search engine für die Ausführung einer Operation (Insert, Delete, Abstract, Search, Spell Check) einen längeren Zeitraum als der definierte Grenzwert (siehe [Konfiguration der d.3 search engine](#)), wird automatisch eine Log-Datei in das Verzeichnis ...\\d3\\d.3 search\\corpora\\[CorpusId]\\PerfLogs geschrieben. This logfile contains details on a call.

**Warning**

As the logfiles are not deleted automatically, this directory must be monitored.

```

2012-12-11-10-25-05-879-5084.txt - Editor
Datei Bearbeiten Format Ansicht ?
Global informations:
Cache Hit:64926086 Miss:4381026 Ratio:93%
Read access: Time:5286,437676s Bytes:21011829355

This operation:
Search+Hints Start:0 Maxhits:1000 Expression: searchterm 7452581818666
Run time:00:00:07.172
Cache Hit:1215784 Miss:15283 Ratio:98%
Disk access: Time:4,779042s Bytes:62592194

Read operations:
F:\d.3 search corpora\d3D\link.dat, Pos:867172, Bytes:4032, Sec:0,000006
F:\d.3 search corpora\d3D\link.dat, Pos:3959693, Bytes:4032, Sec:0,000003
F:\d.3 search corpora\d3D\link.dat, Pos:996345, Bytes:4032, Sec:0,000002
F:\d.3 search corpora\d3D\link.dat, Pos:4009740, Bytes:4032, Sec:0,000002
F:\d.3 search corpora\d3D\link.dat, Pos:4047642, Bytes:4032, Sec:0,000002
F:\d.3 search corpora\d3D\link.dat, Pos:4057741, Bytes:4032, Sec:0,000002
F:\d.3 search corpora\d3D\link.dat, Pos:4110393, Bytes:4032, Sec:0,000003
F:\d.3 search corpora\d3D\Sound.dat, Pos:199347, Bytes:4096, Sec:0,000007
F:\d.3 search corpora\d3D\Sound.dat, Pos:199049, Bytes:4096, Sec:0,000003
F:\d.3 search corpora\d3D\Sound.dat, Pos:199018, Bytes:4096, Sec:0,000003
F:\d.3 search corpora\d3D\Sound.dat, Pos:198983, Bytes:4096, Sec:0,000004
F:\d.3 search corpora\d3D\Sound.dat, Pos:198961, Bytes:4096, Sec:0,000003
F:\d.3 search corpora\d3D\Sound.dat, Pos:187296, Bytes:4096, Sec:0,000003
F:\d.3 search corpora\d3D\words.dat, Pos:27996009, Bytes:4095, Sec:0,000004
F:\d.3 search corpora\d3D\Sound.dat, Pos:187254, Bytes:4096, Sec:0,000004
F:\d.3 search corpora\d3D\words.dat, Pos:27990229, Bytes:4095, Sec:0,000003
F:\d.3 search corpora\d3D\words.dat, Pos:27990797, Bytes:4095, Sec:0,000003
F:\d.3 search corpora\d3D\words.dat, Pos:27991804, Bytes:4095, Sec:0,000003
F:\d.3 search corpora\d3D\words.dat, Pos:27992079, Bytes:4095, Sec:0,000002
F:\d.3 search corpora\d3D\Sound.dat, Pos:187207, Bytes:4096, Sec:0,000003
F:\d.3 search corpora\d3D\words.dat, Pos:27980520, Bytes:4095, Sec:0,000004
F:\d.3 search corpora\d3D\words.dat, Pos:27981435, Bytes:4095, Sec:0,000002
F:\d.3 search corpora\d3D\words.dat, Pos:27981671, Bytes:4095, Sec:0,000003
F:\d.3 search corpora\d3D\words.dat, Pos:27982023, Bytes:4095, Sec:0,000002
F:\d.3 search corpora\d3D\words.dat, Pos:27982887, Bytes:4095, Sec:0,000003
F:\d.3 search corpora\d3D\words.dat, Pos:27983348, Bytes:4095, Sec:0,000003

```

Example of a logfile for such a search-call:

The logfile is divided into three sections:

- **Global Information:** The information in this section refers to the time between the last start of the d.3 search engine up to the execution of the respective operation in the log.

Row Cache	
Hit	specifies how many data accesses could e served from the cache
Miss	specifies how many data accesses could not be served from the cache but have been determined with a reading access to the disk
Ratio	specifies the percentage of accesses served from the cache
Row Read access	
Time	specifies the amount of time specnt reading from the hard disk
Bytes	specifies the amount of data already read from the hard disk

- **This Operation:** The information refers to the logged operation.

Row Search Hints	
Start	specifies the position from which the matches are read
Maxhits	specifies the maximum number of matches to be read from the position
Expression	Search term entered in the user interface by the user

<u>Row Run time</u>	specifies the time required to determine the results
<u>Row Cache</u>	
Hit	specifies the number of read accesses to the cache
Miss	specifies how many data accesses could not be served from the cache but have been determined with a reading access to the disk
Ratio	specifies the percentage of accesses served from the cache
<u>Row disk access</u>	
Time	specifies the time required to read from the hard disk
Bytes	specifies the amount of data read

- **Read operations:** This section logs the details on the read access for the respective file.

If the value for Ratio from the global information is lower than 85%, then this can be a hint that you should allocate more RAM for the cache to the d.3 search engine. If the Ratio value is higher than 85%, then the reason for the delayed operations is rather to be found in the access times of the hard disk.

For a more detailed analysis of the logfiles, please contact your d.velop partner or the d.velop support.

## The search engine database

The following fundamental files are part of a SHAFT instance:

```

DATABASE.DAT
FILES.DAT
FILES.IDX
FILES.WRD
LINK.DAT
SOUND.DAT
SOUND.IDX
WORDS.DAT
WORDS.IDX
SHAFT.INI

```

Depending on the configuration, you may additionally have the file `JOURNAL.DAT`.

For a full backup, the above mentioned files are required except for the files `SHAFT.INI` and `JOURNAL.DAT`.

If the automatic daily backup of the shaft-search engine is enabled, the actually required files are stored in the configured folder daily.

## Start the d.3 search engine

The instances of the d.3 search engine are usually running as a process of the `SHAFT.EXE` in the d.3 process manager.

It is usually named **d.3 search – engine (d3?)**.

Such a process is always started with a specified configuration file with the name `SHAFT.INI` (see call in the d.3 process manager).

### Note

```
shaft c:\d3\d.3 search\corpora\d3x\shaft.ini
```

This file defines the configuration parameters for the `SHAFT` process as well as for the database administration.

Should the specified port be already in use or the database should be marked as corrupt, the search engine starts nevertheless but the database cannot be used.

In this case, stop the search engine and check the transaction protocol (`trans.log` in the transaction protocol directory `TLOG`).

Especially during operation, you must make sure that the process(es) of the search engine are always orderly shut down to avoid respective issues with the associated databases. This is also essential when using external backup mechanisms.

## Backup mechanisms

To ensure that the database is not destroyed by software or hardware faults, it includes several mechanisms to recognize and resolve errors.

The error correction is partially automatic, however, in case of severe faults, the database is marked as corrupt and requires manual interference using the `SHAFTUTL`.

## Online-backup

If possible, the database should not be backed up while the search engine is running as changes to the database could result in backing up an inconsistent database.

Thus, the search engine should either be stopped for the backup or the online backup function should be used.

The online backup function is enabled specifying a target directory and the time of the backup into the database start configuration.

As a result, the database is switched to “read-only” at the specified time and the database structure is checked for consistency. If this check is successful without an error the database is copied to the given directory. If the backup has been successful the archive log and the transaction log are then deleted.

Accordingly, both logs only contain relevant information since the last backup. Backup software can thus save the backup from the specified target directory at any time. It is additionally advisable, to save the archive logs with a backup software directly after they have been created, if the backup software offers this functionality.

## Journal

This is the fundamental mechanism enabled via the configuration parameter `Journal`.

This function writes all changes to a database into the file `journal.dat` before they are applied to the database. Only after this file was written successfully, the changes are applied to the database.

In case of a hardware failure or if the software is terminated, it recognizes on start that the writing to the database has been aborted and repeats the writing. This prevents that a bluescreen or a power failure etc. can destroy the database.

## Transaction-log

The transaction-log is a text-file `trans.log` containing more or less information on the transactions performed on the database, depending on the `loglevel`.

The file is located in the sub-directory `TLOG` of the document corpus directory.

With a `loglevel` of “9”, the content of the indexed files is additionally stored in a sub-directory structure. Using the `SHAFTUTL`, these transactions can later be applied again.

If an error has not been directly recognized and the archive logs thus contain errors, then all logical changes can be applied to the last backup of the database.

If no backup exists and the transaction log was never deleted or disabled, the database can also be fully recreated with this log.

If archive logs and transaction logs exist, then the transaction logs should be applied as they promise to be more successful. However, if a large number of transactions must be restored, then you should additionally use the archive logs as they can be processed faster.

## Archive-log

Here, the journals are not overwritten but stored in a directory.

Thus this directory contains a full change log of the database.

Using the backup and all log files created afterwards you can accordingly recreate the current status of the database.

To associate log files to a backup, they are stored under the file name containing the timestamp of their backup {Date}-{Time}-{Milliseconds}.

The logs are restored with the SHAFTUTL. The program automatically installs these files in the correct order. If no backup of the database exists, it can also be fully recreated provided that all archive logs still exist.

## Maintenance using SHAFTUTL.EXE

The file SHAFTUTL.EXE is provided together with the search engine.

This is a tool allowing to examine and repair or restore database.

The tool is used from the command line. When using it, it always requires the definition of a configuration file (\*.INI) as its first parameter. This is the file SHAFT.INI of the respective search engine database. It is always located in the respective corpus directory of d.3 search.

### Note

```
SHAFTUTL {Pfad+Name der entsprechenden SHAFT.INI}
<ac:macro ac:name="beispiel-doku"
><ac:parameter ac:name="atlassian-macro-output-
type">BLOCK</ac:parameter><ac:rich-text-body><ac:macro
ac:name="code"><ac:plain-text-body><![CDATA[SHAFTUTL "C:\d3\d.3
search\corpora\d3p\shaft.ini"
```

This is followed by the definition of additional parameters described below. If the program is started without parameters it displays an overview of the functions.

```
C:\d3\d.3 search\shaft\64Bit>shaftutl "c:\d3\d.3 search\corpora\d3p\SHAFT1\shaft.ini"
SHAFT tool 3.0.0 (372)
Usage: <DB-Start.cfg> <option> [optionparam]
Options
-c Check DB structure
-i Import archivelogs
  optionparam = Importpath
-is Import synonym file
  optionparam = filename
-r Rebuild
-ri Reindex (Storetext must be enabled?)
-rt Restore and redo transaction log
-d Defrag
-f file dump
-w Word dump
-ws Stopword dump by MaxFilesPerWord
-t redo transaction log
```

## Updates for the d.3 search engine

IN some cases, the search engine d.3 search engine (SHAFT) may have to be updated.

After this update, you may also be prompted to rebuild the index of the search engine database of all document corpora with the program `ShaftUtl.EXE`.

To do so execute the following command:

### Note

```
SHAFTUTL {Pfad+Name der entsprechenden SHAFT.INI} -r
<ac:macro ac:name="beispiel-doku"
><ac:parameter ac:name="atlassian-macro-output-
type">BLOCK</ac:parameter><ac:rich-text-body><ac:macro
ac:name="code"><ac:plain-text-body><![CDATA[ SHAFTUTL
..\corpora\d3p\shaft.ini -r
```

### Note

```
SHAFTUTL "C:\d3\d.3 search\corpora\d3p\shaft.ini" -r
```

If you do not do this, the database may be tagged as corrupt after an update and thus the d.3 search-job processing is halted. However, there is no danger of any loss of data.

## Commandline parameters for the program ShaftUtl.EXE

Reindex	-ri	
		this only works, if the text is stored in the index; in this case, the index is rebuilt entirely.
Restore	-rt	
		this substitutes for the database administrator having to restore the backup and write the transactions again and rather restores the database from the backup directory and then repeats the logged transactions. This reduces possible error sources, to prevent that a backup is restored and then destroyed with an old.
Defrag	-d	
		defragments the index to increase the reading speed of result lists

## Checking the database

To check the consistency of the database, `SHAFTUTL` is started with the option `-c`.

This check can be performed while the search engine is running, however it may happen that an error is found erroneously. This is due to the fact that the search engine does not write changes to the database directly and the check encounters data which appears to be corrupt after it has been changed by the search engine.

To ensure a safe result, the search engine should be stopped. If an error is found in the database, this is marked as "corrupt" and the search engine cannot be started before renewed check does not return any more errors, i.e. after a repair.

## Restoring an archive-log

With the option `-i` you can restore the archive-logs.

For this effect, you specify the path containing the log-files after this parameter.

### Note

```
shaftutl c:\d3\d.3 search\corpora\d3x\shaft.ini -i
c:\backup\archlogs
```

To restore archive-logs, the search engine must be stopped.

If the search engine is running, `SHAFTUTL` aborts with an error. If no backup of the database exists, you should create one before you restore the archive-logs.

### Restore a transaction-log

If you start `SHAFTUTL` with the option `-t`, all recorded insert and delete operations are applied to the database again. This is thus equivalent to a “Reimport” of all files stored in the transaction-log.

Before you restore a transaction-log, you should create a database backup, if you have not already done this as the journaling functions are disabled during the restore.

### Restore database

#### Warning

You should only apply this step, if it is required after an update OR the above mentioned procedures are not applicable.

If you do not have a backup or you have installed a new search engine using different data structures, the option `-r` allows to start a re-indexing (“Rebuild”).

This has the effect that the structure is heaved to a new version, if required, and all links are written again. Before you start a rebuild, you should also create a backup if you have not done so already.

Should the rebuild-process be aborted, it can usually be started over again but this cannot be guaranteed!

### View all words in the database

The option `-w` writes all words stored in the database to the console window. Use the redirection `>` character to write the output to a file.

### View all words in the database which are stop words

The parameter `-ws` writes all words stored in the database and have been changed to stop words by the value `MaxFilesPerWord` to the console window. Use the redirection `>` character to write the output to a file.

### View all filenames of the database

The para `-f` lists all words with the file names stored in the database.

### Re-indexing the database

The parameter `-ri` triggers a re-indexing of the database.

#### Note

This only works, if the text is stored in the index. In this case, the index is rebuilt entirely. This means that all text is extracted and are completely inserted into a new empty index.

### Restoring transactions

The parameter `-rt` initiates a recopying and following restoration of the transactions.

The database is restored from the backup directory and then the logged transactions are repeated.

This reduces possible error sources, to prevent that a backup is restored and then destroyed with an old.

### Defragmenting the database

Starting `SHAFTUTL` with the parameter `-d` initiates a defragmentation of the database.

This defragments the index to allow for a faster reading of result lists.

**Warning**

During the defragmentation, additional hard disk memory in the size of the files link.dat and words.dat is required. The hard disk memory must be available on hard disk where the index is stored.

## 1.5. Maintenance

This chapter contains further information on maintaining the software.

### 1.5.1. Notes on optimizing d.3 search.

It can be advisable to install d.3 search on a dedicated server as populating the database for the document corpora requires a large amount of writing actions.

This server should then have two physical disks so that the write access can be distributed in the best possible way. Furthermore, you should consider the disk speed calculated and logged on starting the engine processes (Disk speed ~61440 KB/s). For a productive environment, this should not fall below 50.000 KB/s.

Moreover, you need a fast network connection to increase the performance. A multi-core processor may also help.

By doing so

- the application directory of the application (. \d.3\d.3 search\ ) and
- the jobs directory of the d.3 search server . \d.3\d.3 search\jobs\ ) and
- the configuration directory of the document corpora ( . \d.3\d.3 search\corpora\ ) and
- the transaction log directory ( . \d.3\d.3 search\{Name of the document corpus}\tlog\ )

should be stored on a physical partition.

The database files of the document corpora (shaft.ini parameter Database under the section [engine]) together with optionally enabled ARCHLOGS-directory (shaft.ini parameter ArchiveLogs=1 under the section [Engine]) should be located on another physical partition.

The backup directory of the document corpora (shaft.ini parameter Path under the section [backup]) should be located on a long-term stable storage medium and can be configured using a UNC-path.

The transaction-log and the archlog-directories for the document corpora are configured in the shaft.ini under the section [engine] with the parameters Logpath={Path to the tlog-directory} and ArchivePath={Path to the ARCHLOGS directory}, which have to be added manually.

The Jobs-directory for the d.3 search server and the definition of the configuration directory are specified in the respective program directory with the manually created configuration file dsearch.ini with the following content:

**Note**

```
[Settings]
JobsPath={Pfadangabe zum JOBS-Ordner}
CorporaPath={Pfadangabe zum Corpora-Ordner}
```

**Note**

If you only move the Database- and Archlogs-directories of the document corpora to another physical partition, you only need to adjust the respective shaft.ini.

## 1.5.2. Notes on optimizing d.3 search on a RAID5 system

### Warning

If you operate the d.3 search server on a RAID5 system, please note the following information:

- RAID5 is rather slow by design and
- with an incorrectly configured RAID5 system, the performance of the d.3 search server is lacking.

The number of disks in RAID5 systems must be chosen to ensure that 1024 / 2048 or 4096 Bytes are distributed to correspond with the NTFS-cluster size. This is the case, if three, five or nine disks are configured as a RAID5 with a parity disk. For this effect, the disks must work with 512 Byte blocks.

### Note

Modern disks often have sectors larger than 512 Byte.

### Note

The RAID5 array should NOT be configured as described below:

Six hard disks are combined to form a RAID5 array.

Accordingly, there are five data disks and one parity disk storing the checksums for the restore process in case of errors.

Now, all sectors across the six disks form a unit.

Since each sector has a size of 512 Byte and five disks in this unit store data, the system always has to write or change clusters of  $5 * 512 = 2.560$  Byte in one piece, as the parity block of the sixth disk has to be calculated for these 2.560 Byte.

NTFS, however, always uses clusters of 4, 8, 16, 32 or 64 KB (usually 4 KB).

When using a 4 KB cluster size and a 2.560 bytes block, there are ~ 1.536 bytes remaining with every write access which are written to a 2.560 bytes block:  $4 \text{ KB} * 1,024 = 4,096$ ,  $2,560 - 4,096 = -1,536$ .

Thus, the RAID controller must read the block before every write access, insert the 1,536 bytes and calculate the new parity to write it back. For free write access, e.g. of Microsoft SQL Server, SHAFT and small files, this behavior leads to a fall of the disk speed down to 25% of the possible write speed.

Solution: Configure a RAID5-controller with 5 disks. In this case, the block would have a size of  $4 * 512 = 2.048$  Byte. The read operation of the controller would then be obsolete as a 4 KB cluster would be equivalent of exactly two blocks.

## 1.5.3. Clues on recovering a document-corpus

There are various ways of recovering a document-corpus or an index.

These depend on the previously used backup strategy (online or offline) and the enabled log-types (ArchivLogs, TLogs, Journal.dat). Moreover, they depend on the loglevel.

Scenario 1	
Online-backup	yes

Scenario 1	
ArchivLogs	yes
Tlogs	yes
Journal.dat	yes
ToDo	<ul style="list-style-type: none"> <li>• Stop the engine</li> <li>• Copy online-backup to Corpus directory (overwrite existing files)</li> <li>• Journal.dat is also located in the Corpus directory</li> <li>• Restore Archivlogs (shaftutl.exe ".\shaft.ini" -i ".\archlogs")</li> <li>• Start the engine</li> </ul>
Alternative	<ul style="list-style-type: none"> <li>• Stop the engine</li> <li>• Delete Journal.dat from the Corpus directory, if it still exists</li> <li>• Make sure that the directory tlog contains several directories with OCR files (LogLevel &gt;= 9)</li> </ul> <p>Method 1:</p> <ul style="list-style-type: none"> <li>• Copy online-backup to Corpus directory (overwrite existing files)</li> <li>• Restore Transaction logs (shaftutl.exe ".\shaft.ini" -t)</li> </ul> <p>Method 2:</p> <ul style="list-style-type: none"> <li>• Restore Transaction logs (shaftutl.exe ".\shaft.ini" -rt) =&gt; Database files will be overwritten automatically</li> <li>• Start the engine</li> </ul>
Difference between the two alternatives	<ul style="list-style-type: none"> <li>• Errors in the archive-logs are also restored</li> <li>• The recovery with the archive-logs is faster</li> <li>• OCR-information must be stored twice when recovering through the transaction-logs (LogLevel &gt;= 9)</li> </ul>

Scenario 2	
Online-backup	yes
ArchiveLogs	yes
Tlogs	yes
Journal.dat	no
ToDo	<ul style="list-style-type: none"> <li>• Stop the engine</li> <li>• Copy online-backup to Corpus directory (overwrite existing files)</li> <li>• Restore Archivlogs (shaftutl.exe ".\shaft.ini" -i ".\archlogs")</li> <li>• Start the engine</li> <li>• Re-index documents since the last ArchLog</li> </ul>
Alternative	<ul style="list-style-type: none"> <li>• Stop the engine</li> <li>• Delete Journal.dat from the Corpus directory, if it still exists</li> <li>• Make sure that the directory tlog contains several directories with OCR files (LogLevel &gt;= 9)</li> </ul> <p>Method 1:</p> <ul style="list-style-type: none"> <li>• Copy online-backup to Corpus directory (overwrite existing files)</li> <li>• Restore Transaction logs (shaftutl.exe ".\shaft.ini" -t)</li> </ul>
Difference between the two alternatives	<ul style="list-style-type: none"> <li>• Errors in the archive-logs are also restored</li> <li>• The recovery with the archive-logs is faster</li> <li>• OCR-information must be stored twice when recovering through the transaction-logs (LogLevel &gt;= 9)</li> </ul>

Scenario 3	
Online-backup	yes
ArchiveLogs	yes
Tlogs	no
Journal.dat	yes
ToDo	<ul style="list-style-type: none"> <li>• Stop the engine</li> <li>• Copy online-backup to Corpus directory (overwrite existing files)</li> <li>• Journal.dat is also located in the Corpus directory</li> <li>• Restore Archivlogs (shaftutl.exe ".\shaft.ini" -i ".\archlogs")</li> <li>• Start the engine</li> </ul>

Scenario 4	
Online-backup	yes
ArchiveLogs	yes
Tlogs	no
Journal.dat	no
ToDo	<ul style="list-style-type: none"> <li>• Stop the engine</li> <li>• Copy online-backup to Corpus directory (overwrite existing files)</li> <li>• Restore Archivlogs (shaftutl.exe ".\shaft.ini" -i ".\archlogs")</li> <li>• Start the engine</li> <li>• Re-index documents since the last ArchLog</li> </ul>
Alternative	Retroactive full-text indexing was not tested

Scenario 5	
Online-backup	yes
ArchiveLogs	no
Tlogs	yes
Journal.dat	yes
ToDo	<ul style="list-style-type: none"> <li>• Download engine</li> <li>• Delete Journal.dat from the Corpus directory, if it still exists</li> <li>• Make sure that the directory tlog contains several directories with OCR files (LogLevel &gt;= 9)</li> </ul> <p>Method 1:</p> <ul style="list-style-type: none"> <li>• Copy online-backup to Corpus directory (overwrite existing files)</li> <li>• Restore Transaction logs (shaftutl.exe ".\shaft.ini" -t)</li> </ul> <p>Method 2:</p> <ul style="list-style-type: none"> <li>• Restore Transaction logs (shaftutl.exe ".\shaft.ini" -rt) =&gt; Database files will be overwritten automatically</li> <li>• Start the engine</li> </ul>
Alternative	more document afterwards than before

Scenario 6	
Online-backup	yes
ArchiveLogs	no
Tlogs	yes
Journal.dat	no
ToDo	<ul style="list-style-type: none"> <li>• Stop the engine</li> <li>• Delete Journal.dat from the Corpus directory, if it still exists</li> <li>• Make sure that the directory tlog contains several directories with OCR files (LogLevel &gt;= 9)</li> </ul> <p>Method 1:</p> <ul style="list-style-type: none"> <li>• Copy online-backup to Corpus directory (overwrite existing files)</li> <li>• Restore Transaction logs (shaftutl.exe ".\shaft.ini" -t)</li> </ul> <p>Method 2:</p> <ul style="list-style-type: none"> <li>• Restore Transaction logs (shaftutl.exe ".\shaft.ini" -rt) =&gt; Database files will be overwritten automatically</li> <li>• Start the engine</li> </ul>
Alternative	more document afterwards than before

Scenario 7	
Online-backup	yes
ArchiveLogs	no
Tlogs	no
Journal.dat	yes

Scenario 7	
<b>ToDo</b>	<ul style="list-style-type: none"> <li>• Stop the engine</li> <li>• Copy online-backup to Corpus directory (overwrite existing files)</li> <li>• Check consistency of online-backup (<code>shaftutl.exe ".\shaft.ini" -c</code>)</li> <li>• Start the engine</li> <li>• Re-index documents since the last online-backup</li> </ul>

Scenario 8	
<b>Online-backup</b>	yes
<b>ArchiveLogs</b>	no
<b>Tlogs</b>	no
<b>Journal.dat</b>	no
<b>ToDo</b>	<ul style="list-style-type: none"> <li>• Stop the engine</li> <li>• Copy online-backup to Corpus directory (overwrite existing files)</li> <li>• Check consistency of online-backup (<code>shaftutl.exe ".\shaft.ini" -c</code>)</li> <li>• Start the engine</li> <li>• Re-index documents since the last online-backup</li> </ul>

Scenario 9	
<b>Online-backup</b>	no
<b>ArchiveLogs</b>	X
<b>Tlogs</b>	X
<b>Journal.dat</b>	X
<b>ToDo</b>	<ul style="list-style-type: none"> <li>• Start the engine</li> <li>• Re-index all documents</li> </ul>

### 1.5.4. Reindexing a document corpus

During the use of a document corpus it can occur that settings will be adjusted, which take effect on the structure of the document corpus. To apply the changes on the complete, already existing document corpus, the search index, the database, must be reindexed.

Possible reasons for adjusting the settings are:

- The delimiter mode will be adjusted for the internationalization.
- The catalog ValidWords will be adjusted company-specific.
- StoreText shall be enabled.

#### Note

For the respective parameters you will be prompted that a reindexing must be executed.

The two possibilities of a reindexing are described below:

### Assistant for the reindexing of a document corpus

The wizard for the reindexing of a document corpus provides you two functions:

- **Parallel regeneration (PR)** of a document corpus into another document corpus. This function is only available, if you specified a mirror when creating a document corpus.
- **Reindexing (R)** of a document corpus. To be able to query missing documents form the d.3 system, you can query all OCR information for a document corpus from the d.3 system again.

#### Warning

This wizard can be used from d.3 server version 7.2.2.7. For older versions, this function is not available.

**Note**

Keep in mind that for a reindexing always only the current version of a document will be reindexed. Older versions will not be applied.

**Configuring reindexing**

In the following it is described which settings you can make for the reindexing.

Common

## Allgemein

<b>Dokumentensammlung</b>	Migrieren von <input type="text" value="d3p (M)"/> in <b>d3ps</b>	
<b>Anzahl d.3 server-Prozesse</b>	<input type="text" value="1"/>	
<b>Zeitraum der Reindizierung</b>	Vom <input type="text" value="01.02.2015"/>  bis zum <input type="text" value="20.02.2015"/> 	
<b>Anzahl der Dokumente pro Call</b>	<input type="text" value="500"/>	
<b>OCR-Datei generieren</b>	<input type="checkbox"/>	
<b>Dokumentensammlung löschen</b>	<input type="checkbox"/>	

Setting	Description
<b>Document corpus</b>	<p>Select the respective document corpus for which a reindexing shall be executed.</p> <p>A <b>(PR)</b> specifies that a parallel regeneration of this document corpus into another takes place. Accordingly, a mirroring was configured for this document corpus. A <b>(R)</b> specifies that a reindexing takes place for this document corpus. This means that all documents are retrieved from the d.3 system and are reindexed again.</p>
<b>Number of d.3 server processes</b>	<p>Specify the maximum number of d.3 server processes to be used simultaneously for the reindexing.</p> <div style="background-color: #ffe6e6; padding: 5px;"> <p><b>Warning</b></p> <p>The specified number of d.3 server processes is not available for the productive system anymore. Please consider this and optionally configure additional d.3 server processes for the time of the reindexing.</p> </div>
<b>Time range for the reindexing</b>	<p>If you do not want to reindex all documents in the document corpora, you can configure a time range here.</p> <div style="background-color: #e6f2ff; padding: 5px;"> <p><b>Note</b></p> <p>The time is the last change date of the document properties.</p> </div>
<b>Number of documents per call</b>	<p>This parameter allows you to specify how many documents are queried from the d.3 server in one cycle. A high number relieves the network. However, the connections remain open for a longer period of time.</p>
<b>Generate OCR-file</b>	<p>If this parameter is enabled, OCR-files are generated, if no OCR information exist for a document, yet.</p> <div style="background-color: #e6f2ff; padding: 5px;"> <p><b>Note</b></p> <p>If this option is not enabled and a document exists without OCR information, these are later listed as errors.</p> </div>

Setting	Description
<b>Delete document corpus</b>	<p>This parameter defines, if the document corpus will be deleted after a successful parallel regeneration.</p> <div style="background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p><b>Note</b></p> <p>This means for a parallel regeneration from d3P to d3PS the following:</p> <ul style="list-style-type: none"> <li>• d3P will be deleted</li> <li>• d3PS will be renamed to d3P</li> </ul> </div> <p>This saves you from a retroactive manual configuration in the d.3 admin for the document corpus.</p> <div style="background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p><b>Note</b></p> <p>For a retroactive reindexing, this option is not available.</p> </div>

### d.3 user data

Setting	Description
<b>d.3 user name</b>	<p>Enter the d.3 user here, who has read access to all document types in the document corpus. Additionally the user must have administrative rights.</p> <div style="background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p><b>Note</b></p> <p>A refinement of the documents to be reindexed can be done by the rights system of the specified user.</p> </div> <div style="background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p><b>Note</b></p> <p>The administration rights are required to read the job-queue in d.3. The query of the job-queue shall avoid that the queue runs full with TIF001-jobs and TIF002-jobs by the reindexing.</p> </div>
<b>d.3 password</b>	Specify the d.3 password for the selected d.3 user.

### d.3 server data

#### d.3 Serverdaten

<b>Rechnername</b>	<input type="text" value="localhost"/>	
<b>d.3-Server-ID</b>	<input type="text" value="P"/>	
<b>Port</b>	<input type="text" value="3400"/>	

Setting	Description
<b>Hostname</b>	Specify the computer name on which the d.3 gateway is configured.
<b>d.3 server ID</b>	Enter the d.3 server ID for your d.3 system which is to be reindexed.
<b>Port</b>	Specify the port via which the d.3 gateway is available.

### Runtimes

This section allows you to specify up to three execution times. First, select the days for which the execution time is to apply. Then specify the start and end time.

To keep the load on the productive system as low as possible, it is recommended not to run the parallel regeneration during the main working hours.

### Warning

The backup times of the d.3 search engine (SHAFT) should be considered to avoid unnecessary errors.

### Note

To set a runtime over 00:00 (e.g. from 22:00 - 05:00) do as follows:

1. **Time 1** from 22:00 - 23:59
2. **Time 2** from 00:00 - 05:00

### Important notes

Before you start a reindexing you should check the following:

1. Make sure that the chosen user has rights for all documents, which you want to migrate. The user rights are relevant because the reindexing is based on a d.3 search.
2. Under circumstances the job-queue of d.3 search can run full by the reindexing. So make sure that you configure at least as many **job-threads** as you use **d.3 server processes** for the reindexing. Keep in mind, that productive queries will still be processed by the **job-threads**. Thus, the number of **jobs-threads** should always be higher than the number of **d.3 server processes** for the reindexing.
3. If have you have connected d.3 search with the d.3 gateway, set the **HandlerCount** for the d.3 gateway connections to the same number as the number of **job-threads**.

### Monitor reindexing

This overview displays the current status of the reindexing.

## Status

## Status

<b>Status</b>	Reindizierung läuft
<b>Startzeit</b>	23.4.2015 14:10
<b>Vorraussichtlich fertig in</b>	00:00:25
<b>Bearbeitete Dokumente</b>	17
<b>Erstellte OCR-Jobs</b>	3
<b>Fehlerhafte Dokumente</b>	Geamtanzahl: 0
<b>Warnungen für Dokumente</b>	Geamtanzahl: 0
<b>Zeitlich erstes Dokument</b>	18.2.2015 12:47
<b>Zeitlich letztes Dokument</b>	23.4.2015 23:59
<b>Zeitlich aktuelles Dokument</b>	23.4.2015 23:59
<b>Aktuelle Dokumentart</b>	Bestellung (3 von 40)

Fortschritt aktuelle Dokumentart

100%

Gesamtfortschritt

8%

Property	Description
<b>Status</b>	<p>This property describes the current status of the reindexing. The following status can occur:</p> <ul style="list-style-type: none"> <li>• <b>Reindexing in progress</b> The reindexing is currently executed.</li> <li>• <b>Program-related pause</b> The reindexing is not within the defined runtimes.</li> <li>• <b>Manual pause</b> The user has stopped the reindexing manually.</li> <li>• <b>Reindexing is paused...</b> The reindexing will be stopped as soon as it is possible.</li> <li>• <b>Reindexing paused due to errors.</b> The reindexing was stopped because errors occurred.</li> <li>• <b>Reindexing successfully completed</b> The reindexing is successfully finished.</li> <li>• <b>Reindexing completed with errors</b> The reindexing is successfully finished but erroneous document exist.</li> </ul>

Property	Description
<b>Start time</b>	Specifies, when the process was started.  <b>Note</b> All documents until this date will be reindexed. Newer documents will not be reindexed because these will be adopted to the document corpus by the usual way.
<b>Approximately completed in</b>	Specifies, how long it takes until the migration is finished.  <b>Note</b> This specification is an estimated value based on the overall progress in relation to the already passed time.
<b>Edited documents</b>	This specifies how many documents were edited so far.
<b>Generated OCR-jobs</b>	Specifies, how many OCR-jobs were already generated.  <b>Note</b> Consider that OCR-jobs will be processed asynchronously by the d.3 async.
<b>Faulty documents</b>	This specifies how many documents were faulty. Additionally it returns information, how often which error code occurred.
<b>Warnings for documents</b>	Specifies for how many document a warning exists.  A warning for a document exists, if none OCR-information exist and none OCR-files shall be created automatically for the reindexing.
<b>First document (in time)</b>	Specifies, which document is the first in time in the current document type. With help of this information the progress for the document type will be generated.
<b>Last document (in time)</b>	Specifies, which document is the last in time in the current document type. With help of this information the progress for the document type will be generated.
<b>Current document (in time)</b>	Specifies, which document is the current (in time) one in the current document type. With help of this information the progress for the document type will be generated.
<b>Current document type</b>	Specifies, which document type is currently reindexed.
<b>Progress of the current document type</b>	Specifies the progress for the current document type. The progress defines that X-percent of the documents from the current document type were processed.  <b>Note</b> If an OCR-job was created for document, then this document applies as edited, also if the OCR-job will be processed asynchronously.
<b>Overall progress</b>	Specifies the overall progress for the reindexing. The progress defines that X-percent of the documents were processed.  <b>Note</b> If an OCR-job was created for document, then this document applies as edited, also if the OCR-job will be processed asynchronously.

## Settings

In this section the settings are displayed which were set previously for this reindexing. The description of the settings can be found in the following chapter.

## Einstellungen

<b>Dokumentensammlung</b>	Migrieren von <b>d3p</b> in <b>d3ps</b>								
<b>Anzahl d.3 server-Prozesse</b>	<input type="text" value="1"/>								
<b>Zeitraum der Reindizierung</b>	Vom <input type="text" value="01.02.2015"/> bis zum <input type="text" value="20.02.2015"/>								
<b>Anzahl der Dokumente pro Call</b>	<input type="text" value="500"/>								
<b>OCR-Datei generieren</b>	<input type="checkbox"/>								
<b>Dokumentensammlung löschen</b>	<input type="checkbox"/>								
<b>d.3 Benutzername</b>	<input type="text" value="dvelop"/>								
<b>Rechnername</b>	<input type="text" value="10.95.1.166"/>								
<b>d.3-Server-ID</b>	<input type="text" value="P"/>								
<b>Port</b>	<input type="text" value="3400"/>								
	<b>Mo</b>	<b>Di</b>	<b>Mi</b>	<b>Do</b>	<b>Fr</b>	<b>Sa</b>	<b>So</b>	<b>Startzeit</b>	<b>Endzeit</b>
<b>Zeit 1</b>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="text" value="00:00"/>	<input type="text" value="23:59"/>
<b>Zeit 2</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="text" value="00:00"/>	<input type="text" value="23:59"/>
<b>Zeit 3</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="text" value="00:00"/>	<input type="text" value="23:59"/>

**Result protocol**

After a successful reindexing the reindexing status changes respectively to the **Reindexing successfully completed**. If errors occurred for the reindexing, which apply individual documents, these will be listed below the section **Settings** under **Error list** in form of **document ID** and **error code**. The meaning of the error codes can be found in the d.3 server documentation.

**Note**

The error code -23 means that none OCR-information exist to this document. If the parameter **Generate OCR-file** was not set, this document will be displayed as faulty.

## Fehlerliste

---

Dokumenten-ID	Fehlercode
P000000001	-23
P000000002	-23
P000000153	-23
P000000154	-23

### Reindexing via the shaftutl.exe

#### Warning

A fundamental requirement to do so is that you have enabled the parameter `StoreText` for the respective document corpus since the creation of the database.

The reindexing via `StoreText` provides the advantage that the documents must not be requested by the d.3 server. Consider that the document corpus is not available for the duration of the reindexing.

#### Warning

Depending on the size of the document corpus, this may take a while (hours/days).

If you don't want that the document corpus is unavailable for the duration of the reindexing, read the chapter [.Reindexing via the d.3 server v3.0.0.](#)

To execute the reindexing do as follows:

- Navigate to the web interface of the respective d.3 search instance.
- Now select in the menu bar the option **Engines** and in the sub-menu the **Corpus ID** of the engine to be reindexed.
- Afterwards click the button **Stop** in der section **Functions**.
- Open the command line (CMD console, `cmd.exe`) and navigate to the directory of the d.3 search installation.

#### Note

```
C:\Users\Administrator> cd C:\d3\d.3 search
```

- Enter the following command: `shaft\shaftutl.exe corpora\\shaft.ini -ri`

#### Note

```
C:\d3\d.3 search> shaft\shaftutl.exe corpora\d3p\shaft.ini -ri
```

- Wait until `shaftutl.exe` is finished successfully.
- Now navigate to the web interface of d.3 search again and start the respective engine (**Engines** | **<Corpus-ID>**) and then under **Functions** the button **Start**).

#### Note

If you are in cluster mode, execute the steps 1 - 4 for each individual node.

## Reindexing via the d.3 server

If you don't have the opportunity to do a reindexing with the `shaftutl.exe`, you can use the reindexing via the d.3 server.

Consider that the reindexing will take more time overall.

### Note

Use, if possible, the wizard for the reindexing in the web interface.

Additional information on the [wizard for the migration of a document corpus](#) can be found here.

### Reindexing

Follow these steps to execute a reindexing via the d.3 server:

- Navigate to the web interface of the respective d.3 search instance.
- Now select in the menu bar the option **Engines** and in the sub-menu the **Corpus ID** of the engine to be reindexed.
- Afterwards click the button **Functions | Stop**.
- Navigate to the directory of the document corpus (by default: `C:\d3\d.3 search\corpora\<Corpus-ID>`).

### Note

Directory for the corpus d3p: `C:\d3\d.3 search\corpora\d3p\`

- Delete all files and folders in this directory except:
  - \*.ini files
  - \*.txt files
- Navigate to the web interface of the respective d.3 search instance.
- Now select in the menu bar the option **Engines** and in the sub-menu the **Corpus ID** of the engine to be reindexed.
- Afterwards click the button **Functions | Start**.
- Navigate to the installation directory of d.3 search.
- Get the script for the reindexing of documents (`reindex_fulltext_corpus.jpl` or `dokumente_nachverschlagworten.jpl`) and copy it into the `ext_jpl` directory of the d.3 server.

### Note

You can request the script from the d.velop support.

### Note

Copy from `C:\%Path_to_Script%\dokumente_nachverschlagworten.jpl` to `C:\d3\d3server.prg\ext_jpl`.

- Adjust the jpl script before using it.
- To do so, navigate to the `line 131` and specify the ID of the document corpus, if you don't want to use the settings of the d.3 server.

```
## !! Die Dokumentensammlung brauch nicht mehr angepasst werden! Es
werden die Einstellungen des d.3 server verwendet! Geben Sie die ID als
```

5. Parameter an, wenn Sie Dokumente in eine andere Dokumentensammlung hinzufügen wollen.

```
error = dsearch_add_document(dok_dok_id, "", 1, "", "", 1)
```

Additionally you can customize the select command for the determination of the document ID's from line 100.

```
#####
# Select zur Ermittlung der DokumentIds
# Kann bei Bedarf angepasst werden.
#####

DBMS DECLARE CURSOR work_cursor FOR \
  SELECT doku_id \
  FROM phys_datei :lock_hint \
  WHERE (dokuart <> '$LOG$') \
  :DOKU_ID_TERM \
  ORDER BY doku_nr DESC

#####
```

- Open the respective d.3 server interface of the repository from which the documents are to be imported to your document corpus.
- Open under **Programs** the copied JPL script `dokumente_nachverschlagworten.jpl` from the `ext_jpl` directory.
- Wait until all documents haven been indexed.

### Note

If you are in cluster mode, the cluster needs to be active after step 4. Check this in the web interface via the menu option **Cluster**.

### Reindexing of the note files

The note files will not be considered by the script `dokumente_nachverschlagworten.jpl`. For this, you can use the following script `Notizdateien_nachverschlagworten.jpl`.

Follow these steps to execute a reindexing of the note files via the d.3 server:

- Navigate to the installation directory of d.3 search.
- Get the script for the reindexing of note files (`Notizdateien_nachverschlagworten.jpl`) and copy it into the `ext_jpl` directory of the d.3 server.

### Note

You can request the script from the d.velop support.

### Note

Copy from `C:\%Path_to_Script%\dokumente_nachverschlagworten.jpl` to `C:\d3\d3server.prg\ext_jpl`.

- Adjust the jpl script before using it.
- To do so, navigate to the `line 38` and specify the ID of the document corpus, if you don't want to use the settings of the d.3 server.

```
// Konfiguration
vars DSEARCH_CORPUS = "d3P";
```

Additionally you can customize the select command for the determination of the document ID's from line 62.

```
// Doku-IDs aller Notizdateien ermitteln
DBMS DECLARE CURSOR work_cursor FOR \
  SELECT doku_id \
  FROM dependent_documents \
  WHERE UPPER(doc_extension) = 'NOT' \
  ORDER BY doku_id DESC
```

- Open the respective d.3 server interface of the repository from which the documents are to be imported to your document corpus.
- Open under **Programs** the copied JPL script `Notizdateien_nachverschlagworten.jpl` from the `ext_jpl` directory.
- Wait until all documents haven been indexed.

### Note

If you are in cluster mode, the cluster needs to be active. Check this in the web interface via the menu option **Cluster**.

## Generate OCR-files in the d.3 server

If no OCR-files exist in your d.3 system because you have not generated any OCR-files during the import, you can generate these by the script `nach_vertiffen.jpl`.

### Note

Use, if possible, the wizard for the reindexing in the web interface. Additional information on the [wizard for the migration of a document corpus](#) can be found here.

Follow these steps to generate the OCR-files via the d.3 server:

- Navigate to the installation directory of d.3 search.
- Get the script for the retroactive conversion to TIFF (`nach_vertiffen.jpl`) and copy it into the `ext_jpl` directory of the d.3 server.

### Note

Copy from `C:\%Path_to_Script%\nach_vertiffen.jpl` to `C:\d3\d3server.prg\ext_jpl`.

- Adjust the jpl script before using it.
- To do so, navigate to the line 101 and specify the table, which contains the respective document ID's. By default the table `phys_datei` is specified, which contains all documents.

```
// Tabelle mit den DokumentIDs in der Spalte doku_id
vars table_dokuid = "phys_datei"
```

Additionally you can specify the user under whose name the classes will be determined and the OCR-file will be generated from line 103. If the entry remains empty, the owner of the document will be used or the editor for the status "Processing".

```
// User, der für das Vertiffen "benutzt" werden soll
vars user_script = ""
```

- Open the respective d.3 server interface of the repository from which the documents are to be imported to your document corpus.
- Open under **Programs** the copied JPL script `nach_vertiffen.jpl` from the `ext_jpl` directory.
- Wait until for all documents an OCR-file is created.

### Note

If you are in cluster mode, the cluster needs to be active. Check this in the web interface via the menu option **Cluster**.

## 1.6. Errors and issues

If issues occur during the operation of the d.3 search engine,

- open the Log-Viewer on your d.3 server (**d.velop | d.3 gateway | Log-Viewer**) and check the log.

All faulty processes in connection with the d.3 system are displayed here.

All d.3 search log messages are returned with the module ID DSEARCH.

Date	Source	ID	Message
18.05 15:04:13,121	DSEARCH	00000E68	CALL 42: 410100 - d3_admin 049 - DSEARCHGetCorpora
18.05 15:04:13,121	DSEARCH	00000E68	Number: 1
18.05 15:04:13,121	DSEARCH	00000E68	ReturnCaption: SUCCESS
18.05 15:04:13,121	DSEARCH	00000E68	Header->retcode: 0
18.05 15:04:13,121	DSEARCH	00000E68	Function terminated normaly.
18.05 15:04:13,121	DSEARCH	00000E68	#####
18.05 15:04:13,137	DSEARCH	00000E68	ISM_CLOSE
18.05 15:04:13,731	DSEARCH	00000E68	Connecting 10.5.0.119
18.05 15:04:13,731	DSEARCH	00000E68	#####
18.05 15:04:13,731	DSEARCH	00000E68	CALL 43: 410100 - d3_admin 049 - DSEARCHGetServerStatus
18.05 15:04:13,731	DSEARCH	00000E68	ReturnCaption: SUCCESS
18.05 15:04:13,731	DSEARCH	00000E68	Header->retcode: 0
18.05 15:04:13,731	DSEARCH	00000E68	Function terminated normaly.
18.05 15:04:13,731	DSEARCH	00000E68	#####

### 1.6.1. d.3 search

#### The active/active cluster is inactive

If your active/active cluster is inactive, check the following settings which could disturb the operation of the cluster:

- Check, if all desired d.3 search engine (SHAFT) processes are started. This can be checked in the web interface, accessible by default under the address <http://localhost:6133/>, under the menu option **Engines | <CorpusId>**:

## Einstellungen

<b>Status</b>	Online
<b>Korpus</b>	d3p
<b>Rechnername</b>	localhost
<b>Dateiname</b>	C:\d3\d.3 search\shaft\64Bit\shaft.exe
<b>Befehlszeile</b>	C:\d3\d.3 search\corpora\d3p\shaft.ini
<b>Prozess-ID</b>	5452
<b>Startzeit</b>	20.08.2014 15:23:32.281

- Check, if your d.3 search server process is entered in the d.3 process manager and active:

Process	Process ID	Memory	Virtual Memory	Starttime	Starts
d.3 search (server)	5788 (0x0000169C)	11736 K	10692 K	20.08.2014 14:59:51	3

- Check, if the cluster members are communicating by navigating to the d.3 search web interface, which can be accessed by default under the address <http://localhost:6133/> and check the **Log**. This is located under the menu option **Cluster**. Here, you should find entries in the following form, if the connection between the cluster members is successful:

```
21.08.2014 11:34:31: [SEOK]: [REDACTED] - IAMALIVEMASTER
21.08.2014 11:34:38: [SEOK]: [REDACTED] - IAMALIVEMASTER
```

### d.3 search starts and is stopped immediately afterwards

d.3 search stops automatically, if the configured TCP/IP Port (3406 or the one specified in the `dsearch.ini`) are already used by another server process.

You can also find respective messages about this in the log-viewer.

### d.3 search is slow

In the following sections, you will find different procedures to optimize the performance of d.3 search.

#### Allocate more RAM to the Shaft-engine

To reduce the disk access to a minimum, the available RAM should be used completely.

You can allocate more RAM to the Shaft-engine by increasing the `cache` in the `shaft.ini`.

#### Optimize the hard disk access

To find out, if the hard disk access is performed in the best-possible way, follow the steps below:

1. Check, if the Shaft engine has created a PerfLogs directory in the Corpus directory. If this is the case, check if the read access times are significantly slow. To do so, get the respective values from the Perflogs-file ([Notes on document collections with the d.3 search engine](#)) and calculate as follows:

$(\text{Bytes}/1024/1024)/\text{Time} = x \text{ MB/s}$ .

This value ideally be higher than 30 MB/s. If this is not the case then please check the hard disk access. Please consider the notes below.

2. A speed-test is executed when the Shaft engine ist started. This value should ideally be lower than 30 MB/s.

As it is often not possible to keep the entire index in the main memory for lack of resources, the main weakness of the search is the hard disk access.

For this reason, check if other programs are accessing the physical hard disk holding the Shaft database (index).

Ideally, one physical hard disk is exclusively used by the Shaft engine.

### Windows cache-management

Check, if the page faults per second significantly increase when you send search queries.

If this is the case, Windows is swapping large amounts of data to the hard disk. As the access to the hard disk is significantly slower than that to the main memory, this causes long delays. This phenomenon is called swapping.

To avoid or reduce this, the shaft engine must be allocated more main memory which is not swapped.

This can be configured with the tool `SetSystemFileCacheSize`. You can obtain the tool for free from the internet.

Please execute the tool with the following call:

```
SetSystemFileCacheSize.exe MinCache MaxCache
```

where:

`MaxCache` = (Physical cache allocated to the Shaft-engine) - (1024 bis 2048 MB)

`MinCache` = 8 MB

Afterwards, the page faults per second should no longer increase as significantly during a search query and thus no swapping should take place.

### **Active/Active-cluster inactive due to faulty node - preventing accumulation of unprocessed jobs**

If you operate d.3 search in a cluster and the cluster is inactive due to an error, jobs not being processed are not accumulating.

In some cases it is not useful that these will be stopped and handled retroactively in large amount. Then it is recommended to remove the faulty node temporarily from the cluster and to add it again retroactively.

Therefore, follow the steps below:

1. The d.3 search instance on the defective node must be shut down. Remove the faulty node from the cluster by using the button **Delete** in the web interface of the master.

## Cluster-Server

IP-Adresse	Rechnername	Modus	Letzte Meldung	Information	Fehlerme
		Unbekannt			
		Master			

2. Afterwards uninstall d.3 search from the removed computer. Take care, that the node has no tracks to the old d.3 search installation.
3. If the node is ready for operation again, d.3 search will be installed again on it. Add the node to the cluster again on the usual way (see chapter [Configure new d.3 search instance for the cluster](#)).

### Warning

If you do not execute step 2 correctly, then the existing document corpus on the previously defective node will not be updated after the new integration to the cluster.

By the asynchronous states of the nodes the cluster is inconsistent.

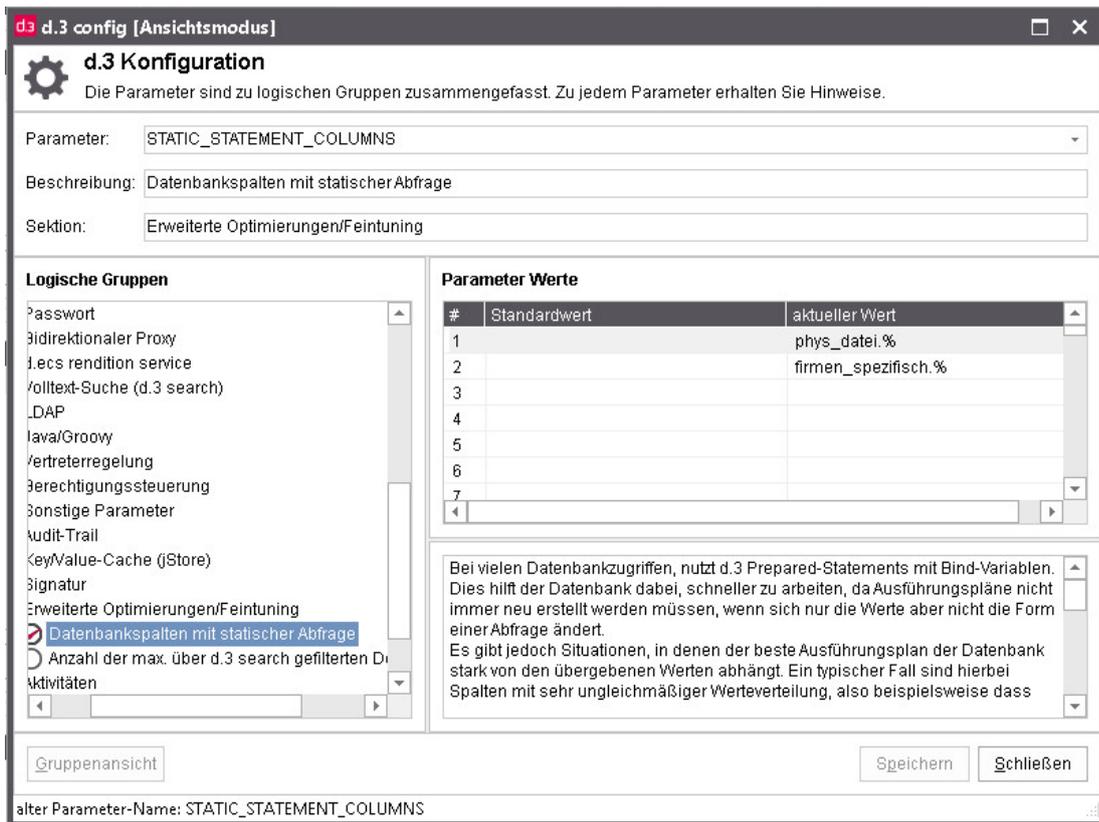
### Slow reindexing by using the wizard

Under certain circumstances, the search can take unnecessary long time for the reindexing by the wizard.

If this is the case, the following parameter can be set to accelerate the search in the database.

- To do so, open the configuration **d.3 config** in the d.3 admin.
- Select the option **Advanced optimizations** and click on the option **Database columns with static querying**. Ente the value: **firmen\_spezifisch.%**.

The change has the effect that the database can better analyze the SQL command and thus, uses the index of the date values.



## 1.6.2. d.3 search engine

### Check list and questionnaire for the support

Questions to be considered when you encounter issues with the d.3 search engine:

- What does the `SHAFT.INI` of the respective document corpora look like?
- Are the journaling, the backup and the logging enabled?
- Did you stop the `SHAFT` processes or are they stopped automatically from time to time?
- Hard or controlled? When at what time?
- When did the problem first occur?
- Can you specify the date and time?
- How do these issues show?
- Process behavior, error messages, disk space usage
- What is document volume to be managed by d.3 search?
- How many and which document types are to be managed?
- By which search engine?
- How many documents do actually have to be managed in one document corpus? (this can be determined with an SQL-query on the d.3 database)
- What does the job processing of d.3 search look like?

- What are the versions of the following files?
- SHAFT . EXE
- SHAFTUTL . EXE
- DSEARCHSVR . EXE
- SHAFT . DLL
- possibly TextUpdate . exe
- How is d.3 search configured in the d.3 administration (screenshots, descriptions)?
- d.3 parameter SUCHWORT\_OPTION (KEYWORD\_OPTION)
- DUMMY document type(s)
- Association of document corpora to document types
- d.3 parameter DSEARCH\_SUPPORT
- Type of full-text search (d.3 search + d.3 keyword tables?)

### Requests to the d.3 search engine take a long time

The processing speed of the d.3 search engine stringly depends on the available main memory.

The parameter `CacheSize` in the `shaft.ini` in the configuration directory of the document corpus should be increased.

The parameter of the `PhysicalCache` should also be enabled ("1"), if the server has enough available main memory.

### Repairing a SHAFT database

#### Repair using an existing backup and transaction-log

##### Assumptions

We look at the document corpus "d3x" with the respective corpus-directory `\d3x` as a sub-directory of the `corpora` directory.

The `shaft`-process is shut down; `SHAFTUTL -C` returns: "Corrupt Database".

A backup exists in the backup directory configured in the `SHAFT.INI`.

All transactions since the last backup (file date) have been logged in the directory `TLOG` (the first entries in the file `trans.log` originate from the time range directly after the backup -> seamless transition).

##### Backup phase

- verify that the d.3 search engine (d3x)-process (SHAFT) is shut down
- verify that the d.3 search server-process is shut down
- Create a backup of the Corpus-directory (e.g. just "d3p", initially without sub-directories)
- Create a backup of the sub-directory `ARCHLOGS`
- Create a backup of the sub-directory `TLOG`

##### Restore phase

- Delete all databases (except for `.INI` files) from the directory `d3x`.

- Restore the content of the content of the backup into the directory of the document corpus (e.g.. d3x).
- Check the restored backup with `SHAFTUTL -C` (the backup remains in the configured backup directory).

### Note

More information on how to open the file `SHAFTUTL.EXE` can be found [here](#).

- Start the SHAFT process with the DOS command.
- Wait until this has started completely (window in the foreground and restored: Database opened).
- Stop the SHAFT process again (close the SHAFT window).
- Repeat the transaction with `SHAFTUTL -t`.
- Check the database with `SHAFTUTL -C` including the restored transactions (and create a backup if necessary).

### Verification phase

- Start the SHAFT process with the DOS command again and wait until it has completely started (database opened, memory usage at least as configured in the `SHAFT.INI`).
- If it is OK, stop it again and start it in the d.3 process manager.
- Wait until the process has started in the d.3 process manager (process is initially active, then remains idle, see chart in the web-interface of the d.3 process manager).
- Check the memory usage.
- Start the d.3 search server again.
- Check the connection of d.3 search to the document corpus in the d.3 administration.
- Check the job processing of d.3 search.

### **Repair using an existing backup and archive-log without a transaction-log**

Perform the same steps as described under the repair with an existing transaction protocol, but in the restore phase, you do NOT use `SHAFTUTL` with the parameter `-t`, but with `-i`.

### **Re-indexing a SHAFT database**

Should the database of the SHAFT search engine be recognized as corrupt and the alternatives under **Repairing a SHAFT-database** are not available or do not work, then you can still re-index database.

This is done with the `SHAFTUTL` parameter `r` and is the final option to repair a corrupt database.

The re-indexing should also be considered, if a large number of documents have been deleted from the document corpus. This is also analyzed when the database is checked with `SHAFTUTL` and the parameter `-c`. In this case, the re-indexing works like a “defragmentation” and thus reduces the size of the database.

### Note

More information on how to open the file `SHAFTUTL.EXE` can be found [here](#).

### **The search engine is not accessible**

#### Issue

You cannot execute any full-text searches in the d.3 smart explorer.

These are aborted with respective error messages:



### Symptoms

The full-text information of (new) documents is provided for the d.3 search server, however, this only increases the number of pending jobs but cannot forward them to the search engine.

The d.3 logviewer shows messages such as:

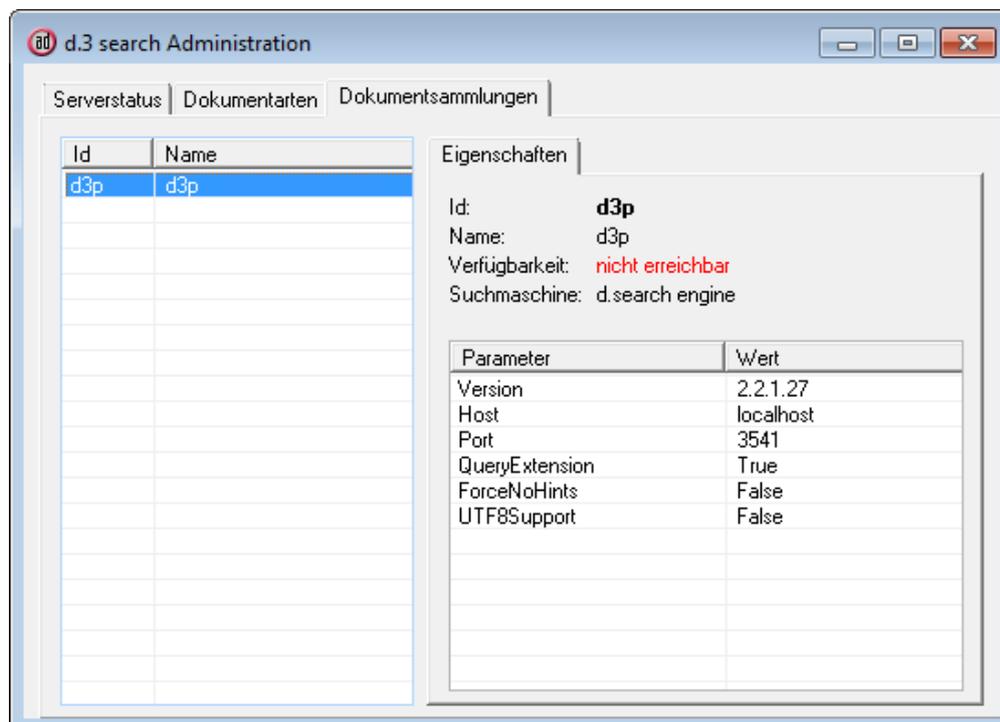
"Exception occurred."

"Context: Shaft - connect"

"Type: EldConnectException"

"Message: Socket-error # 10061#0D#0AConnection rejected."

If you opened the configuration module for d.3 search in the d.3 administration and went to the tab **Document corpora**, you are unable to connect to the respective instance of the search engine with a click on the document corpus:



### Measures

Check, if the respective d.3 search engine process is completely started.

This can be determined by checking the memory usage of the respective process. Also check, if so-called "Zombi-processes" are blocking the processes to be started.

Furthermore, make sure that the TCP/IP connection between d.3 search and the instance of the d.3 search engine can be established. For this effect, the ports in the CORPUS.INI and shaft.ini must match and be available.

## 32 bit-version of the Shaft engine

### Error:

If a 32-bit version of the Shaft engine is used under a 64-bit operating system with a memory allocation of more than 1 GB (cache > 1048576), this can lead to an error message "Open Error (Database or Listenport)" on starting the d.3 search engine in the

d.3

log-file. Then, the process is restarted.

A validation of the document corpus with the Shaftutl tool fails with the message

"WARNING! Database not closed!

Try to close y/[n]?"

and it cannot be closed entering "y".

"Unable to close the database!"

### Cause:

The 32-bit version of the Shaft engine can only work with a maximum allocated memory of 1 GB (CacheSize=1048576).

Only from d.3 search version 2.2.2.0, a Shaft engine version 1.4 is provided, which also contains a 64-bit version capable of working with larger memory allocations.

### Solution:

Please check the allocated cache in the respective configuration file `shaft.ini` of the d.3 search document corpus and correct it if necessary. This is specified in KB.

As an alternative, update d.3 search to a version containing a 64-bit version of the Shaft engine.

## 1.7. Additional information sources and imprint

If you want to deepen your knowledge of d.velop software, visit the d.velop academy digital learning platform at <https://dvelopacademy.keelelearning.de/>.

Our E-learning modules let you develop a more in-depth knowledge and specialist expertise at your own speed. A huge number of E-learning modules are free for you to access without registering beforehand.

Visit our Knowledge Base on the d.velop service portal. In the Knowledge Base, you can find all our latest solutions, answers to frequently asked questions and how-to topics for specific tasks. You can find the Knowledge Base at the following address: <https://kb.d-velop.de/>

Find the central imprint at <https://www.d-velop.com/imprint>.