

---

# Apache Solr for TYPO3 - File Indexing

Release 12.0.0

Steffen Ritter, Timo Hund, Markus Friedrich, Rafael Kähm

2023-10-17 06:51



# CONTENTS

<b>1</b>	<b>How to use</b>	<b>3</b>
<b>2</b>	<b>Concepts and Background information</b>	<b>5</b>
2.1	Indexing contexts	5
2.1.1	Site exclusive records	5
2.2	Schema fields for files in Solr	5
<b>3</b>	<b>TypoScript Reference</b>	<b>7</b>
3.1	pageContext	7
3.1.1	contentElementTypes	7
3.1.2	fileExtensions	7
3.1.3	enableFields	7
3.1.4	contentEnableFields	7
3.2	storageContext	8
3.2.1	[StorageUid]	8
3.2.2	languages	8
3.2.3	fileExtensions	8
3.2.4	enableFields	8
3.2.5	folders	8
3.2.6	excludeFolders	9
3.3	recordContext	9
3.4	attachments	9
3.5	fields	9
3.6	plugin.tx_solr.index.queue._FILES	9
3.7	default	9
3.8	pageContext	10
3.9	storageContext	10
3.10	default	10
3.11	[StorageUid]	10
3.12	recordContext	10
3.13	default	11
3.14	[__tablename__]	11
<b>4</b>	<b>Available Events</b>	<b>13</b>
4.1	Indexing	13
4.1.1	AfterFileInfoHasBeenAddedToDocumentEvent	13
4.1.2	AfterFileMetaDataHasBeenRetrievedEvent	13
4.1.3	AfterSingleFileDocumentOfItemGroupHasBeenIndexedEvent	13
4.2	Repository	13
4.2.1	BeforeFileQueueItemHasBeenRemovedEvent	13
4.2.2	AfterFileQueueItemHasBeenRemovedEvent	13
4.2.3	BeforeMultipleFileQueueItemsHaveBeenRemovedEvent	14
4.2.4	AfterMultipleFileQueueItemsHaveBeenRemovedEvent	14

<b>5</b>	<b>Releases</b>	<b>15</b>
5.1	Version 2.1	15
	5.1.1 Release 2.1.3	15
	5.1.2 Release 2.1.2	15
	5.1.3 Release 2.1.1	15
5.2	Version 3.1	15
	5.2.1 Release 3.1.0	15
5.3	Version 3.3	15
	5.3.1 Release 3.3.0	15
5.4	Version 3.2	16
	5.4.1 Release 3.2.0	16
5.5	Version 4.1	16
	5.5.1 Release 4.1.0	16
5.6	Version 4.2	16
	5.6.1 Release 4.2.0	16
5.7	Version 5.0	16
	5.7.1 Release 5.0.0	16
5.8	Version 3.0	17
	5.8.1 Release 3.0.0	17
5.9	Version 6.0	17
	5.9.1 Release 6.0.0	17
5.10	Version 4.0	17
	5.10.1 Release 4.0.0	17
5.11	Version 7.0	18
	5.11.1 Release 7.0.0	18
5.12	Version 8.0	18
	5.12.1 Release 8.0.0	18
5.13	Version 10.0	18
	5.13.1 Release 10.0.0	18
5.14	Version 11.0	20
	5.14.1 Release 11.0.1	20
	5.14.2 Release 11.0.0	20
5.15	Version 12.0	22
	5.15.1 Release 12.0.0	22
	5.15.2 Contributors	23
	5.15.3 How to Get Involved	24

EXT:solrfal is an addon to EXT:solr ( $\geq 3.0$ ) which allows to index files using the FileAbstractionLayer of TYPO3. solrfal is meant to work as of TYPO3 6.2 LTS and later.

---

**Note:** Please note that this extension is currently part of the Early Access Program (EAP). The EAP funds the dedicated development of the core extension, the implementation of new solr server features and the maintenance. Please see: <http://www.typo3-solr.com/en/solr-for-typo3/development-model/> if you want to contribute to the financial support.

---



## HOW TO USE

### 1. Setup **EXT:solr >=3.0.0**

- configure page indexing and/or record indexing according to your needs
- make sure you have the compatible EXT:solr version, please see [version compatibility matrix](#)
- EXT:solrfal depends on system extension “filemetadata”, which will be activated automatically while activating the EXT:solrfal extension.

### 2. **Include the Static TypoScript** “Search - FAL File Indexing (solrfal)”

### 3. **Add scheduler task for EXT:solrfal**

- The TASK “File Index Queue Worker” needs to be set up.

### 4. **Check scheduler TASKs for FAL in Core**

- If you have external storages and/or it is possible that files change in a storage in ways, the TYPO3 backend is not involved it is mandatory to have the “File Abstraction Layer: Update storage index” running frequently (a.k.a every 5-10 minutes).
- Solr uses the MetaData of FAL; if you have extractors for MetaData you also might run the “File Abstraction Layer: Extract metadata in storage” frequently.

### 5. Do your **individual TypoScript configuration**

### 6. **Setup EXT:tika** to enable search within file contents and metadata.





# CONCEPTS AND BACKGROUND INFORMATION

## 2.1 Indexing contexts

A context describes the “place” where a file is detected to be indexed.

EXT:solrfal knows three contexts of indexing.

- **PageContext:** files which are found during indexation of a frontend-page
- **RecordContext:** files which are attached to records (which are indexed by solr)
- **StorageContext:** files which just reside in a file-system of a storage

### 2.1.1 Site exclusive records

A file in fal can be referenced in any record across multiple sites. Therefore a contexts are checked for changes. In some cases we know, that a change will only effect the contexts of the current site, because the are only related to the current site by nature (e.g. pages, tt\_content, sys\_files\_references). When you want to configure tables to be treated the same way, you can configure the tables in “siteExclusiveRecordTables”.

## 2.2 Schema fields for files in Solr

EXT:solr extends the schema with file specific attributes which are automatically added to the solr document on indexing. See the following list of available and indexed fields:

- **fileStorage:** uid of the Storage the file resides in
- **fileUid:** uid of the file
- **fileMimeType:** the mimetype of the file
- **fileName:** filename (including extensions)
- **fileSize:** file size in bytes
- **fileExtension:** file extension (without .)
- **fileSha1:** content hash of the file (SHA1),
- **filePublicUrl:** publicUrl of the file object
- **fileReferenceType:** table name from which the file was referenced (e.g. tt\_content, or tx\_new)
- **fileReferenceUid:** uid of the record the file was referenced from



## TYPOSCRIPT REFERENCE

### 3.1 pageContext

**Type** Boolean

**TS Path** plugin.tx\_solr.index.enableFileIndexing.pageContext

**Default** 1

**Function** enables indexing for files attached to Content-Element-Records while indexing the frontend pages

#### 3.1.1 contentElementTypes

**Type** Array (key: CType from tt\_content, value: fields in which file attachments should be extracted);

**TS Path** plugin.tx\_solr.index.enableFileIndexing.pageContext.contentElementTypes

**Default** text => bodytext, header\_link; textpic =>bodytext, header\_link; uploads => media

**Function** For different content element types you may want files detected in different fields. The default configuration is to behave like solrfile.

If you have added a custom content element, you may want to configure fields here.

#### 3.1.2 fileExtensions

**Type** String, comma separated values (file extensions without .) or \*

**TS Path** plugin.tx\_solr.index.enableFileIndexing.pageContext.fileExtensions

**Default** \*

**Function** allows to restrict the files being indexed by the file extension

#### 3.1.3 enableFields

**Type** array (of column names in pages)

**TS Path** plugin.tx\_solr.index.enableFileIndexing.pageContext.enableFields

**Allowed keys** endtime, accessGroups

**Function** Use the enableFields from the Page it is referenced at for file

#### 3.1.4 contentEnableFields

**Type** array (of column names in pages)

**TS Path** plugin.tx\_solr.index.enableFileIndexing.pageContext.contentEnableFields

**Allowed keys** accessGroups

**Function** Use the enableFields from the page content elements referencing the file

## 3.2 storageContext

**Type** boolean (1/0);

**TS Path** plugin.tx\_solr.index.enableFileIndexing.storageContext

**Default** 0

**Function** enables indexing of all files in a storage

### 3.2.1 [StorageUid]

**Type** array

**TS Path** plugin.tx\_solr.index.enableFileIndexing.storageContext.[StorageUid]

**Index** affected StorageUid (f.e. fileadmin/ generally 1)

**Function** Enables a detailed indexing configuration per Storage, see properties for details

### 3.2.2 languages

**Type** string, comma separated values (integers list of sys\_language uids)

**TS Path** plugin.tx\_solr.index.enableFileIndexing.storageContext.[StorageUid].languages

**Default** 0

**Function** define for which languages this storage should be indexed

### 3.2.3 fileExtensions

**Type** string, comma separated values (file extensions without .) or \*

**TS Path** plugin.tx\_solr.index.enableFileIndexing.storageContext.[StorageUid].fileExtensions

**Default** \*

**Function** allows to restrict the files being indexed by the file extension

### 3.2.4 enableFields

**Type** array (of column names in sys\_file\_metadata)

**TS Path** plugin.tx\_solr.index.enableFileIndexing.storageContext.[StorageUid].enableFields

**Allowed keys** endtime, accessGroups

**Function** FAL generally does not have enable fields, but metadata ships fields which can be used for that purpose. With this configuration you define “enableFields” just for indexation.

### 3.2.5 folders

**Type** string, comma separated values

**TS Path** plugin.tx\_solr.index.enableFileIndexing.storageContext.[StorageUid].folders

**Default** \*

**Function** List of valid directories, relative to storage root directory

**Since** 3.1

### 3.2.6 excludeFolders

**Type** string, comma separated values

**TS Path** plugin.tx\_solr.index.enableFileIndexing.storageContext.[StorageUid].excludeFolders

**Default**

**Function** List of directories to exclude, relative to storage root directory

**Since** 3.1

### 3.3 recordContext

**Type** boolean

**TS Path** plugin.tx\_solr.index.enableFileIndexing.recordContext

**Default** 1

**Function** enables indexing of all file attachment at records;

needs further configuration in the index queue: a, table needs to indexed at all, b, attachment indexation needs to be activated for that table

### 3.4 attachments

**Type** boolean

**TS Path** plugin.tx\_solr.index.queue.[indexingConfiguration].attachments

**Default** 0

**Function** enables file attachment detection when indexing the record

### 3.5 fields

**Type** string, comma separated values (column names of tables)

**TS Path** plugin.tx\_solr.index.queue.[indexingConfiguration].attachments.fields

**Default** \*

**Function** define in which columns of an record files should be detected

### 3.6 plugin.tx\_solr.index.queue.\_FILES

Configuration array to configure the index.queue processing for files. The configuration will be merged. This means that every context specific configuration inherits the default configuration. In addition, if there is a special configuration within context (like per table or storage) these will inherit the base configuration of the context. Each configuration is to be defined like plugin.tx\_solr.index.queue.[indexingConfiguration].fields

### 3.7 default

**Type** array

**TS Path** plugin.tx\_solr.index.queue.\_FILES.default

**Function** Mapping of Solr field names on the left side to database table field names or content objects on the right side. Used for every file indexed

**Default**

```
title = title
description = description
altText_stringS = alternative
width_intS = width
height_intS = height

category_stringM = SOLR_RELATION
category_stringM {
    localField = categories
    foreignLabelField = uid
    enableRecursiveValueResolution = 1
    multiValue = 1
}
```

### 3.8 pageContext

**Type** array

**TS Path** plugin.tx\_solr.index.queue.\_FILES.pageContext

**Function** Additional mapping of Solr field names on the left side to database table field names or content objects on the right side. Used for every file indexed in pageContext

### 3.9 storageContext

**Type** array

**TS Path** plugin.tx\_solr.index.queue.\_FILES.storageContext

**Default** empty

**Function** See introduction and following two entries.

### 3.10 default

**Type** array

**TS Path** plugin.tx\_solr.index.queue.\_FILES.storageContext.default

**Function** Additional mapping of Solr field names on the left side to database table field names or content objects on the right side. Used for every file indexed in storageContext

### 3.11 [StorageUid]

**Type** array

**TS Path** plugin.tx\_solr.index.queue.\_FILES.storageContext.[StorageUid]

**Function** Additional mapping of Solr field names on the left side to database table field names or content objects on the right side. Used for every file of Storage [StorageUid] indexed in storageContext

### 3.12 recordContext

**Type** array

**TS Path** plugin.tx\_solr.index.queue.\_FILES.recordContext

**Default** empty

**Function** See introduction and following two entries.

### 3.13 default

**Type** array

**TS Path** plugin.tx\_solr.index.queue.\_FILES.recordContext.default

**Function** Additional mapping of Solr field names on the left side to database table field names or content objects on the right side. Used for every file indexed as attachment in recordContext

### 3.14 [\_\_tablename\_\_]

**Type** array

**TS Path** plugin.tx\_solr.index.queue.\_FILES.recordContext.[\_\_tablename\_\_]

**Function** Additional mapping of Solr field names on the left side to database table field names or content objects on the right side. Used for every file found attached to record of [\_\_tablename\_\_] indexed in recordContext





## AVAILABLE EVENTS

In case you need to adapt or extend the behaviour of solrfal the following events exist you may consume in your listeners.

### 4.1 Indexing

#### 4.1.1 AfterFileInfoHasBeenAddedToDocumentEvent

The event `ApacheSolrForTypo3\Solrfal\Event\Indexing\AfterFileInfoHasBeenAddedToDocumentEvent` is dispatched after the file information are mapped to Apache Solr document belonging to processed file. This event allows third party extensions to replace or modify the file document after the file info has been added to the document.

#### 4.1.2 AfterFileMetaDataHasBeenRetrievedEvent

The event `ApacheSolrForTypo3\Solrfal\Event\Indexing\AfterFileMetaDataHasBeenRetrievedEvent` is dispatched after the (translated) `MetaData` record of a file is retrieved. The listeners can take the `MetaData` array and may modify it. The modified `MetaData` will then be hand over to the TypoScript “Service”. As an result fields added in the listener to that event can be addressed from your regular TypoScript setup.

#### 4.1.3 AfterSingleFileDocumentOfItemGroupHasBeenIndexedEvent

The event `ApacheSolrForTypo3\Solrfal\Event\Indexing\AfterSingleFileDocumentOfItemGroupHasBeenIndexedEvent` is dispatched after a single Apache Solr document has been indexed. This event allows third party extensions to react on properties provided by this event.

**Note:** This event can not be used to modify provided properties.

### 4.2 Repository

#### 4.2.1 BeforeFileQueueItemHasBeenRemovedEvent

The event `ApacheSolrForTypo3\Solrfal\Event\Repository\BeforeFileQueueItemHasBeenRemovedEvent` is dispatched before the file-index-queue-item will be removed from queue. This event allows third party extensions to react on properties provided by this event.

#### 4.2.2 AfterFileQueueItemHasBeenRemovedEvent

The event `ApacheSolrForTypo3\Solrfal\Event\Repository\AfterFileQueueItemHasBeenRemovedEvent` is dispatched after a single file-index-queue-item has been removed. **Note:** You can not fetch those items from database anymore, because they are really deleted. This event allows third party extensions to react on properties provided by this event.

#### 4.2.3 BeforeMultipleFileQueueItemsHaveBeenRemovedEvent

The event `ApacheSolrForTypo3\Solrfal\Event\Repository\BeforeMultipleFileQueueItemsHaveBeenRemovedEvent` is dispatched before the list of file-index-queue-items is removed from index queue. This event allows third party extensions to react on properties provided by this event.

#### 4.2.4 AfterMultipleFileQueueItemsHaveBeenRemovedEvent

The event `ApacheSolrForTypo3\Solrfal\Event\Repository\AfterMultipleFileQueueItemsHaveBeenRemovedEvent` is dispatched after the list of file-index-queue-items is removed from index queue. **Note:** You can not fetch those items from database anymore, because they are really deleted. This event allows third party extensions to react on properties provided by this event.

## RELEASES

### 5.1 Version 2.1

#### 5.1.1 Release 2.1.3

- Fixed typo in `__RecordContext` configuration example

#### 5.1.2 Release 2.1.2

- Added hook in `FileAttachmentResolver` (`FileAttachmentResolverAspectInterface`) that can be used to modify the result of “`detectFilesInField`”
- Fix index queue initialization, since the site root wasn’t considered entries of all sites were deleted or files of other site roots got detected
- `ContextFactory` provided context record field name instead of expected index configuration name. Fixed this by using the right database column
- Fixed Bug that space in “`plugin.tx_solr.enableFileIndexing.pageContext.fileExtensions`”, was not handled properly

#### 5.1.3 Release 2.1.1

- Set dependency of `EXT:solr` to 3.1.1 because 3.1.0 contains an invalid `composer.json` file without a version number

### 5.2 Version 3.1

#### 5.2.1 Release 3.1.0

- Use configuration object from `EXT:solr` without array access operators
- Use correct var to load storage configuration in `StorageContextDetector`
- Missing namespace change in `RecordContextDetector`
- Add configuration options for valid and exclude folders in storage context
- Respect boolean clause limit and site
- Skip storage permission evaluation while running the indexing task “File Index Queue Worker”
- Use translated page for file reference information

### 5.3 Version 3.3

#### 5.3.1 Release 3.3.0

- Fix handling of access restrictions in record context if `fe_group` is not defined in enable columns

## 5.4 Version 3.2

### 5.4.1 Release 3.2.0

- Added possibility to configure webRoot or set it to PATH\_site to allow having absRefPrefix auto in the TypoScript configuration
- Added improved observation of record, tt\_content and sys\_file\_reference updates
- Add field variantId to the solr fields to allow variants for duplicates in EXT:solr

## 5.5 Version 4.1

### 5.5.1 Release 4.1.0

- Fixed file detection in pageContext, detected file uids have to be stored and considered in all indexing runs of a single page
- Fixed broken tests and initialize TSFE in pageContext since this is required in TYPO3 8 when the PageRepository is used
- Add CE textmedia to the list of content element types supported in the pageContext by default
- Add cache to consistency aspect to avoid multiple processing of the same record
- Followup of split of Site into Site and SiteRepository in EXT:solr
- Adapt Queue to implement method getStatisticsBySite to show number of files on file queue initialization (Deprecates Queue::getItemsCountBySite and ItemRepository::countBySiteAndIndexConfigurationName).
- Trigger no detector when a page or content element is updated outside a site root
- Adapt tests for TYPO3 8 LTS
- Add support for new link syntax(t3://file) support in FileAttachmentResolver
- Use the RootPageResolver instead of the Util functions
- To allow grouping accross different systems a systemHash was added to the variantId, see <https://github.com/TYPO3-Solr/ext-solr/pull/929>
- Wrapped CliEnvironment to be used only in cli-mode, see <https://github.com/TYPO3-Solr/ext-solr/pull/936>
- Documentation is now a part from Git Repo.

## 5.6 Version 4.2

### 5.6.1 Release 4.2.0

- Make “siteExclusiveRecordTables” configurable: By now the tables “pages”, “pages\_language\_overlay” and “tt\_content” were treated as “site exclusive records”. A change on these record only triggers the detectors of the current site because checking references to other sites is not needed. Now this setting is configurable and the table “sys\_file\_reference” was added to the default tables since we also know from this table, that is only related to the current site.
- Make compatible with EXT:solr 7.0.0

## 5.7 Version 5.0

### 5.7.1 Release 5.0.0

- Feature: Add backend module to see the file indexing queue status
- Bugfix: PageContextDetector::addDetectedFilesToPage should return empty array, when page context is disabled

- Feature: Refactor connection handling to differ between read and write connections
- Feature: Add hook in PageContextDetector (\$GLOBALS['TYPO3\_CONF\_VARS']['EXTCONF']['solrfal']['PageContextDetector']) to force added files in pageContext
- Bugfix: Prevent interruption of indexer when file is missing
- Feature: Allow to requeue from EXT:solr index inspector
- Feature: Added command “solrfal:resetQueueError” to reset errored items in the queue
- Bugfix: Prevent interruption of indexer with invalid connection
- Bugfix: findAllOutStandingMergeIdSets() does not recognize errored items
- Task: Remove MigrationsIntroduceIndexingConfiguration
- Task: Migrate queries to doctrine dbal

## 5.8 Version 3.0

### 5.8.1 Release 3.0.0

- Set EXT:solr dependency to 4.0.0 since TypoScript configuration object is not available until version 4
- Removed usage of deprecated method to get the table name (“getTableToIndexByIndexingConfigurationName()”) and switched to TypoScript configuration object
- Fix file detection in record context to ensure file detection in translated records
- Added validity check to the file attachment resolver, to ensure that hidden references and missing files are not added to the index queue
- Registered new signal to remove files marked as missing
- Updates of sys\_file records are respected
- Add file extension configuration option to record context
- Added check for non existing index queue items on updates to storage context detector
- Fix incomplete definition of index configuration name in storage context

## 5.9 Version 6.0

### 5.9.1 Release 6.0.0

- Feature: EXT:solr 9 follow up, differ between reading and writing connections
- Bugfix: Catch indexing exceptions
- Bugfix: Catch solarium HttpException
- Feature: EXT:solr 9 follow up, use new location of Site class
- Feature: Allow to limit the indexing on a certain siteId (for solrconsole)
- Feature: Implement ItemRepository::findBy for solrconsole
- Bugfix: Reindex files for PageContext after changing sys\_file\_metadata
- Feature: Add method “deleteBy” to ItemRepository for solrconsole

## 5.10 Version 4.0

### 5.10.1 Release 4.0.0

- Document merging to avoid duplicates of multiple files

- Performance optimization in AbstractRecordDetector
- Fixed TS path in default template plugin.tx\_solr.index.queue.\_\_TABLENAME\_\_.fields.url
- Performance fix in backend: When content or pages are updated, now only the detectors for the affected site are triggered instead of all. This gives a huge performance boost on systems with many sites.

## 5.11 Version 7.0

### 5.11.1 Release 7.0.0

- Feature: EXT:solr 10 followups (Adapt functional tests to use the side handling)
- Bugfix: Consider site on file detection in recordq!Context

## 5.12 Version 8.0

### 5.12.1 Release 8.0.0

- Feature: EXT:solr 11 followups: - Fix deprecations - Adapt tests to run them against TYPO3 10.4 LTS - Define EXT:Solr 11 as dependency in composer

## 5.13 Version 10.0

### 5.13.1 Release 10.0.0

**Note:** As we're planning to simplify the extension and dependency handling we're harmonize the extension version with TYPO3. So the next version of EXT:solrfal is 10.0.0 and compatible with TYPO3 10 LTS and EXT:solr v 11.1.x.

New in this release:

[TASK!!!] Clean up of legacy notation `<link file:2016` syntax detection

- removes old style TYPO3 file links syntax `<link file:2016 ...` replaces in all fixtures to `t3://file?uid=syntax`

**Warning:** No support for `<link file:2016` anymore, please use `t3://file?uid=2016` instead.

Fixes: #72

[BUGFIX] RecordContextDetector uses always default language

The RecordContextDetector always queued the files in default language parallel to the correct language entry. The file for default language MUST NOT be queued, if translation for record exists.

Fixes: #129

[TASK] Solarium 6 follow-ups

Adjust method calls according to Solarium method signature.

Fixes: 123

[TASK] Drop TYPO3 9.5 support : via `Util::getIsTYPO3VersionBelow10()`

- drops TYPO3 9.5 support in PHP code
- adjusts travis version matrix

- changes EXT:Solr dependency to dev-master
- changes dependencies TYPO3 9.5 in `ext_emconf.php` und `composer.json`

Followup of: [TYPO3-Solr/ext-solr#2889](#)

[BUGFIX] Fix type hinting issues

This pull Request fixes following issues:

- ... `ConsistencyAspect::issueCommandOnDetectors()` must be of the type `int`, string given ...
- ... `ReferenceIndexEntry::getRecord()` must be of the type `array`, null returned ...
- ... Argument 4 passed to `RequeueItemHandler::postProcessIndexQueueUpdateItem()` must be of the type `integer`, null given

Fixes: #108

[BUGFIX] Fix access evaluation in page context

Combinations of page and content element access settings are not evaluated correctly, the indexing currently relies on the access settings set in the file index queue. Unfortunately due to the functioning of the page indexing this information may be faulty, especially if there are several permission combinations.

This commit fixes this issue by evaluation the page and content element permissions in the page context itself, instead of using the permissions from indexed page document.

Additionally a new TypoScript configuration is added, to allow to configure the considered page content permissions:

```
plugin.tx_solr.index.enableFileIndexing {
    pageContext.contentEnableFields.accessGroup = fe_group
}
```

Fixes: #119

[BUGFIX] Access settings in PageContext ignored

EXT:solrfal versions (8.0.0+) did not respect the access settings of pages and/or content elements in PageContext. Required code parts and settings to consider page access settings were deactivated, see PageContext and default TypoScript.

Access settings of pages and content elements are respected again.

Fixes: #98

[BUGFIX] Ignore invalid `sys_refindex` entries

Fixes errors caused by invalid `sys_refindex` by ignoring invalid `sys_refindex` entries. This is achieved by setting a nullable return type `getRecord` and handling invalid references in the detector.

Fixes: #115

[TASK] Change configuration files to TYPO3 file extensions

This Pull-Request change the file extensions from TypoScript files. Previously TypoScript files use the file extension `txt` which is replaced with `typoscript`.

Hints within code and documentations are updated according to these changes.

All changes them self should no effect current installations, since the old files are still in place and include the new ones.

---

**Note:** There are no changes to the TypoScript itself.

---

#### #### Changes:

- Move TypoScript configuration Move TypoScript configuration files from file extension 'txt' to 'typoscript'. Old files now import the new files in order to avoid a breaking change.
- Documentation update

Change includes in examples from text file into TypoScript file.

Fixes: #109

[BUGFIX] print\_r() expects bool, but int used in ConsistencyAspect

See [TYPO3-Solr/ext-solr#2753](#)

Fixes: #99

[BUGFIX] Do not fail on PageContextDetector if file is not in database

Previously EXT:solrfal failed on files, which were not present in sys\_file or file system. This change prevents EXT:solrfal to fail on this scenario.

## 5.14 Version 11.0

### 5.14.1 Release 11.0.1

#standwithukraine #nowar

This is a maintenance release for TYPO3 11.5, containing:

- [Bugfix 204] add parentheses to function call backport by @thomashohn #206
- [Bugfix 202] undefined array key backport by @thomashohn #207
- [BUGFIX::BP:11.0] Fix handling of invalid sites by @dkd-friedrich #201
- [BUGFIX] cast to int by exact comparison operator by @dkd-kaehm #214
- [TASK:BP:11.0] Support queue type setting by @dkd-friedrich #199
- [BUGFIX:BP:11.0] Fix environment for overlays by @dkd-friedrich #210
- [BUGFIX:BP:11.0] Fix observation of translations in RecordContext by @dkd-friedrich #217
- [BUGFIX:BP:11.0] Respect configured languages in RecordContext by @dkd-friedrich #228

### 5.14.2 Release 11.0.0

EXT:solrfal 11.0 is the new release for TYPO3 11.5, our major versions are now matching the version of TYPO3. At this point we would like to express our thanks to all contribution partners and especially Georg Ringer for the work on the TYPO3 11 LTS compatibility.

Note: Delayed update handling introduced with EXT:solr 11.2 is not yet supported by solrfal. Solrfal will consider the introduced UpdateEvents in an upcoming version and with this adaption also the delayed update handling.

#standwithukraine #nowar

New in this release:

[TASK] Improve exception handling (#190)

It might occur that the storage page of an updated record couldn't be found, e.g. as already deleted or an inconsistent database. To ensure system won't be blocked and the index is up-to-date handling of this exceptions in the ConsistencyAspect is improved. To simplify exception handling in general and also failure analysis, solrfal specific exceptions and individual error codes are introduced.



[CLEANUP] Improve/simplify valid filePublicUrl

Simplifies and improves the determination of the correct path for dumpFile eID script and non-public storages. By adapting the url generation in the DocumentFactory it can be simplified and also allows a separate usage of the DocumentFactory.

[BUGFIX] Fix file detection in PageContext (#170)

Since signal preGeneratePublicUrl has been removed in TYPO3 11 LTS no files in PageContext can be detected. This is fixed by using the existing GeneratePublicUrlForResourceEventListener to detect the files.

[BUGFIX] Undefined array key “\_FILES” in StorageContextDetector line 167 (#172)

Starting the indexer on PHP 8+ with fileadmin lead to unwanted warnings.

[BUGFIX] Fix notices during indexing (#174)

[TASK] Fix DocumentFactoryTest (#160)

Adapts and fixes the DocumentFactoryTest case handleProtectedFilesInPageContext, as there were changes in the basic tests of solr and solrfal.

[BUGFIX] Ensure valid filePublicUrl (#164)

Determining the correct path for dumpFile eID script for non-public storages might fail if indexing in CLI context. This issue is fixed by using the site and language information stored in the solrfal contexts and build the url on that basis.

[TASK:11.5] Sync EXT:solr and EXT:solrfal APIs

This change contains a sync to EXT:solr release-11.5.x API changes.

[TASK] Apply TYPO3 Coding Standards on ext\_\*.php and Tests/Unit

TYPO3 Coding Standards introduced and required fixes carried out

[BUGFIX] Fix progress bar styling

[TASK] Remove not used ext\_update file

[TASK] Simplify controller’s code

[TASK] Use new namespace of PageRepository

[TASK] Replace usage of TYPO3\_REQUESTTYPE\_CLI

[TASK] Replace usage of TextExtractorRegistry::getInstance()

[BUGFIX] Properly fake TSFE

[TASK] Move extension icon

[BUGFIX] Fix calls of getSolrConfigurationFromPageId

The order of arguments changed, however 2nd argument must be the language

[BUGFIX] Correct link to index queue module

[BUGFIX] Make the backend module work again

[TASK] Migrate signals of core to events (#159)

Migrate signals of core to events

[TASK] use TYPO3\_VERSION and make it mandatory in composer scripts

## 5.15 Version 12.0

### 5.15.1 Release 12.0.0

EXT:solrfal 12.0 is the new release for TYPO3 12.4 and EXT:solr 12.0.

New in this release:

!!![TASK] Replace signal-slot with event dispatcher

The previously available Extbase Signals have been removed from EXT:solrfal in favor of PSR-14 Events.

- The signal `\ApacheSolrForTypo3\Solrfal\Indexing\DocumentFactory::emitAddedSolrFileInfo` has been replaced by `\ApacheSolrForTypo3\Solrfal\Event\Indexing\AfterFileInfoHasBeenAddedTo`
- The signal `\ApacheSolrForTypo3\Solrfal\Indexing\DocumentFactory::emitFileMetaDataRetrie` has been replaced by `\ApacheSolrForTypo3\Solrfal\Event\Indexing\AfterFileMetaDataHasBeenRet`
- The signal `\ApacheSolrForTypo3\Solrfal\Indexing\Indexer::addGroupDocumentsToIndex::emitIn` has been replaced by `\ApacheSolrForTypo3\Solrfal\Event\Indexing\AfterSingleFileDocumentOfIt`
- The signal `\ApacheSolrForTypo3\Solrfal\Queue\ItemRepository::emitItemRemovedFromQueue` has been replaced by `\ApacheSolrForTypo3\Solrfal\Event\Repository\AfterFileQueueItemHasBeen`
- The signal `\ApacheSolrForTypo3\Solrfal\Queue\ItemRepository::emitMultipleItemsRemovedFrom` has been replaced by `\ApacheSolrForTypo3\Solrfal\Event\Repository\AfterMultipleFileQueueIte`
- The signal `\ApacheSolrForTypo3\Solrfal\Queue\ItemRepository::emitBeforeItemRemovedFromQ` has been replaced by `\ApacheSolrForTypo3\Solrfal\Event\Repository\BeforeFileQueueItemHasBee`
- The signal `\ApacheSolrForTypo3\Solrfal\Queue\ItemRepository::emitBeforeMultipleItemsRemo` has been replaced by `\ApacheSolrForTypo3\Solrfal\Event\Repository\BeforeMultipleFileQueueIt`

All changes:

- [TASK] Prepare requirements for TYPO3 12 [d2a7f49 on @2023-06-23](#) (thanks to Rafael Kähm)
- [TASK] Integrate PHPStan in EXT:solrfal :: initial [d644010 on @2023-07-13](#) (thanks to Rafael Kähm)
- [BUGFIX] Set PSR-4 Namespaces properly [e891164 on @2023-07-01](#) (thanks to Rafael Kähm)
- [TASK] add EXT:tika to require-dev, since PHPStan requires it [d102af3 on @2023-07-13](#) (thanks to Rafael Kähm)
- [TASK] Reuse EXT:solr cibuild.sh script [98d63a9 on @2023-07-17](#) (thanks to Rafael Kähm)
- [TASK] Apply PHPStan level 3 and typo3-codings standards :: basics [44c7bd2 on @2023-07-01](#) (thanks to Rafael Kähm)
- [TASK] Apply PHPStan level 3 and typo3-codings standards :: basics 2 [162218d on @2023-07-17](#) (thanks to Rafael Kähm)
- [TASK] Migrate fixtures from XML to CSV :: renaming part [a2ffc80 on @2023-07-17](#) (thanks to Rafael Kähm)
- [TASK] Make EXT:solrfal be module working [f30faa9 on @2023-08-04](#) (thanks to Rafael Kähm)
- [TASK] Migrate fixtures from XML to CSV :: formatting part [fadaa43 on @2023-07-17](#) (thanks to Rafael Kähm)
- [FIX] Unit/Integration tests :: basics [82a67a3 on @2023-07-21](#) (thanks to Rafael Kähm)

- [BUGFIX] Add `implements LoggerAwareInterface` to classes using `LoggerAwareTrait` [268c611 on @2023-07-28](#) (thanks to Rafael Kähm)
- !!![TASK:T12] Replace signal-slot with event dispatcher [20cbc52 on @2023-07-21](#) (thanks to Rafael Kähm)
- !!![TASK] Refactor `pagelrecordstorage` context :: integrates usage of core context [da84929 on @2023-08-06](#) (thanks to Rafael Kähm)
- [TASK] Raise PHPStan to level 4 and fix its issues [bfe722a on @2023-08-11](#) (thanks to Rafael Kähm)
- [TASK:T12] Make new TCA type 'file' functional [f07912b on @2023-08-11](#) (thanks to Rafael Kähm)
- [TASK] Raise PHPStan to level 6 and fix its issues [2efcbb3 on @2023-08-11](#) (thanks to Rafael Kähm)
- [TASK] Raise PHPStan to level 7 and fix its issues [7ecdc56 on @2023-08-14](#) (thanks to Rafael Kähm)
- [TASK] Refactor `postProcessIndexQueueInitialization` hook to `AfterIndexQueueHasBeenInitializedEvent` listener [178821d on @2023-08-14](#) (thanks to Rafael Kähm)
- [BUGFIX] Use `AfterItemHasBeenIndexedEvent` instead of its renamed version `AfterIndexItemEvent` [68cc885 on @2023-08-14](#) (thanks to Rafael Kähm)
- [BUGFIX] Regressions of : [TASK] Raise PHPStan to level 6 and fix its issues [111acc2 on @2023-08-14](#) (thanks to Rafael Kähm)
- [TASK] Refactor `postProcessIndexQueueUpdateItem` hook to `AfterIndexQueueItemHasBeenMarkedForReindexing` listener [d4840c7 on @2023-08-14](#) (thanks to Rafael Kähm)
- [CI] Use the newest Composer 2.5.x [cfde511 on @2023-08-15](#) (thanks to Rafael Kähm)
- [TASK] Prepare releases for EXT:solrfal 12.0.0 [b78868a on @2023-08-15](#) (thanks to Rafael Kähm)
- [TASK] Streamline EXT:solrfal events to same naming [451b63d on @2023-08-15](#) (thanks to Rafael Kähm)
- [TASK] Include release 12 in release notes [e581327 on @2023-08-21](#) (thanks to Markus Friedrich)
- [BUGFIX] Actions replace `LOCAL_VOLUME_NAME` with `SOLR_VOLUME_NAME` [a696be7 on @2023-09-22](#) (thanks to Rafael Kähm)
- [TASK] Fix PHP-CS for `single_line_empty_body` rule [ce5b689 on @2023-09-22](#) (thanks to Rafael Kähm)
- [TASK] Sync with EXT:solr dev-main 2023.09.21 [d00e3f7 on @2023-09-21](#) (thanks to Rafael Kähm)
- [TASK] Enable page context detection on `tt_content` [9f96e3e on @2023-09-22](#) (thanks to Rafael Kähm)
- [BUGFIX] Sometimes the `ctype` seems to be empty [10ab740 on @2023-09-25](#) (thanks to Georg Ringer)
- [BUGFIX] `pages_language_overlay` table was still used for translated page attachments [1a57201 on @2023-10-06](#) (thanks to Rafael Kähm)
- [TASK] Add missing `@throws` on `RecordContextDetector` methods [6e34200 on @2023-10-13](#) (thanks to Rafael Kähm)
- [TASK] use `Throwable` instead of `Exception` and apply EXT:solr\* inheritance in exceptions [fa4772b on @2023-10-13](#) (thanks to Rafael Kähm)
- [TASK] Refresh README.md [e214c91 on @2023-10-13](#) (thanks to Rafael Kähm)

### 5.15.2 Contributors

- Dmitry Dulepov
- Georg Ringer
- Lars Tode
- Markus Friedrich
- Rafael Kähm

Thanks to everyone who helped in creating this release!

Also a big thanks to our partners that have joined the Apache Solr EB für TYPO3 12 LTS (Feature) program:

- .hausformat GmbH
- b13 GmbH
- clickstorm GmbH
- Connecta AG
- cron IT GmbH
- CS2 AG
- digit.ly GmbH
- FTI Touristik GmbH
- Hochschule Furtwangen
- in2code GmbH
- internezzo ag
- jweiland.net
- Kassenärztliche Vereinigung Rheinland-Pfalz
- KONVERTO AG
- Leuchtfeuer Digital Marketing GmbH
- Lingner Consulting New Media GmbH
- LOUIS INTERNET GmbH
- mehrwert intermediale kommunikation GmbH
- plan2net GmbH
- queo GmbH
- Québec.ca gouv.qc.ca
- Snowflake Productions GmbH
- Stadtverwaltung Villingen-Schwenningen
- Stämpfli AG
- unternehmen online GmbH & Co. KG
- VisionConnect GmbH
- WIND INTERNET

### 5.15.3 How to Get Involved

There are many ways to get involved with Apache Solr for TYPO3:

- Submit bug reports and feature requests on [GitHub](<https://github.com/TYPO3-Solr/ext-solr>)
- Ask or help or answer questions in our [Slack channel](<https://typo3.slack.com/messages/ext-solr/>)
- Provide patches through Pull Request or review and comment on existing [Pull Requests](<https://github.com/TYPO3-Solr/ext-solr/pulls>)
- Go to [www.typo3-solr.com](<http://www.typo3-solr.com>) or call [dkd](<http://www.dkd.de>) to sponsor the ongoing development of Apache Solr for TYPO3

Support us by becoming an EB partner:

<http://www.typo3-solr.com/en/contact/>

or call:

+49 (0)69 - 2475218 0