

# Package ‘R.filesets’

April 9, 2018

**Version** 2.12.1

**Depends** R (>= 3.2.0), R.oo (>= 1.21.0)

**Imports** digest (>= 0.6.15), future.apply (>= 0.1.0), R.methodsS3 (>= 1.7.1), R.utils (>= 2.6.0), R.cache (>= 0.13.0), stats

**Suggests** future (>= 1.7.0), gtools (>= 3.5.0)

**SuggestsNote** Recommended: gtools

**Title** Easy Handling of and Access to Files Organized in Structured Directories

**Author** Henrik Bengtsson [aut, cre, cph]

**Maintainer** Henrik Bengtsson <henrikb@braju.com>

**Description** A file set refers to a set of files located in one or more directories on the file system. This package provides classes and methods to locate, setup, subset, navigate and iterate such sets. The API is designed such that these classes can be extended via inheritance to provide a richer API for special file formats. Moreover, a specific name format is defined such that filenames and directories can be considered to have full names which consists of a name followed by comma-separated tags. This adds additional flexibility to identify file sets and individual files. NOTE: This package's API should be considered to be in an beta stage. Its main purpose is currently to support the aroma.\* packages, where it is one of the main core components; if you decide to build on top of this package, please contact the author first.

**License** LGPL (>= 2.1)

**URL** <https://github.com/HenrikBengtsson/R.filesets>,  
<http://www.aroma-project.org/>

**BugReports** <https://github.com/HenrikBengtsson/R.filesets/issues>

**LazyLoad** TRUE

**biocViews** Infrastructure, DataImport, DataRepresentation

**RoxygenNote** 6.0.1

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2018-04-09 04:30:39 UTC

**R topics documented:**

R.filesets-package . . . . .	2
ChecksumFile . . . . .	3
ChecksumFileSet . . . . .	5
ColumnNamesInterface . . . . .	6
FullNameInterface . . . . .	7
GenericDataFile . . . . .	10
GenericDataFileSet . . . . .	12
GenericDataFileSetList . . . . .	14
GenericTabularFile . . . . .	16
GenericTabularFileSet . . . . .	18
loadRDS . . . . .	20
loadToEnv . . . . .	21
RDataFile . . . . .	21
RDataFileSet . . . . .	23
RdsFile . . . . .	24
RdsFileSet . . . . .	26
readDataFrame . . . . .	27
TabularTextFile . . . . .	29
TabularTextFileSet . . . . .	32
<b>Index</b>	<b>35</b>

---

R.filesets-package      *Package R.filesets*

---

**Description**

A file set refers to a set of files located in one or more directories on the file system. This package provides classes and methods to locate, setup, subset, navigate and iterate such sets. The API is designed such that these classes can be extended via inheritance to provide a richer API for special file formats. Moreover, a specific name format is defined such that filenames and directories can be considered to have full names which consists of a name followed by comma-separated tags. This adds additional flexibility to identify file sets and individual files. NOTE: This package's API should be considered to be in an beta stage. Its main purpose is currently to support the aroma.\* packages, where it is one of the main core components; if you decide to build on top of this package, please contact the author first.

This package should be considered to be in an alpha or beta phase. You should expect the API to be changing over time.

**Installation**

To install this package, call `install.packages("R.filesets")`.

**To get started**

To get started, see:

1. [GenericDataFileSet](#)
2. [TabularTextFile](#)

**How to cite this package**

Please cite references [1] when using this package.

**License**

The releases of this package is licensed under LGPL version 2.1 or newer.

The development code of the packages is under a private licence (where applicable) and patches sent to the author fall under the latter license, but will be, if incorporated, released under the "release" license above.

**Author(s)**

Henrik Bengtsson

**References**

[1] H. Bengtsson, *The R.oo package - Object-Oriented Programming with References Using Standard R Code*, In Kurt Hornik, Friedrich Leisch and Achim Zeileis, editors, Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), March 20-22, Vienna, Austria. <http://www.ci.tuwien.ac.at/Conferences/DSC-2003/Proceedings/>

---

ChecksumFile

*The ChecksumFile class*

---

**Description**

Package: R.filesets

**Class ChecksumFile**

**Object**

~~|

~~+--FullNameInterface

~~~~~|

~~~~~+--GenericDataFile

~~~~~|

~~~~~+--ChecksumFile

**Directly known subclasses:**

public abstract static class **ChecksumFile**  
 extends *GenericDataFile*

A ChecksumFile is an object referring to a file that contains a checksum for a corresponding "main" file.

**Usage**

```
ChecksumFile(...)
```

**Arguments**

... Arguments passed to *GenericDataFile*.

**Fields and Methods****Methods:**

|                           |   |
|---------------------------|---|
| <code>create</code>       | -   |
| <code>getChecksum</code>  | -   |
| <code>isOld</code>        | -   |
| <code>readChecksum</code> | Reads the checksum value.                               |
| <code>validate</code>     | Asserts that the checksum matches the checksum of file. |

**Methods inherited from GenericDataFile:**

`as.character`, `clone`, `compareChecksum`, `copyTo`, `equals`, `fromFile`, `getAttribute`, `getAttributes`, `getChecksum`, `getChecksumFile`, `getCreatedOn`, `getDefaultFullName`, `getExtension`, `getExtensionPattern`, `getFileSize`, `getFileType`, `getFilename`, `getFilenameExtension`, `getLastAccessedOn`, `getLastModifiedOn`, `getOutputExtension`, `getPath`, `getPathname`, `gunzip`, `gzip`, `hasBeenModified`, `is.na`, `isFile`, `isGzipped`, `linkTo`, `readChecksum`, `renameTo`, `renameToUpperCaseExt`, `setAttribute`, `setAttributes`, `setAttributesBy`, `setAttributesByTags`, `setExtensionPattern`, `testAttributes`, `validate`, `validateChecksum`, `writeChecksum`

**Methods inherited from FullNameInterface:**

`appendFullNameTranslator`, `appendFullNameTranslatorByNULL`, `appendFullNameTranslatorByTabularTextFile`, `appendFullNameTranslatorByTabularTextFileSet`, `appendFullNameTranslatorBycharacter`, `appendFullNameTranslatorBydata.frame`, `appendFullNameTranslatorByfunction`, `appendFullNameTranslatorBylist`, `clearFullNameTranslator`, `clearListOfFullNameTranslators`, `getDefaultFullName`, `getFullName`, `getFullNameTranslator`, `getListOfFullNameTranslators`, `getName`, `getTags`, `hasTag`, `hasTags`, `resetFullName`, `setFullName`, `setFullNameTranslator`, `setListOfFullNameTranslators`, `setName`, `setTags`, `updateFullName`

**Methods inherited from Object:**

`$`, `$<-`, `[[`, `[[<-`, `as.character`, `attach`, `attachLocally`, `clearCache`, `clearLookupCache`, `clone`, `detach`,

equals, extend, finalize, getEnvironment, getFieldModifier, getFieldModifiers, getFields, getInstantiationTime, getStaticInstance, hasField, hashCode, ll, load, names, objectSize, print, save

### Author(s)

Henrik Bengtsson

---

ChecksumFileSet      *The ChecksumFileSet class*

---

### Description

Package: R.filesets

#### Class ChecksumFileSet

Object

~~|

~~+--FullNameInterface

~~~~~|

~~~~~+--GenericDataFileSet

~~~~~|

~~~~~+--ChecksumFileSet

#### Directly known subclasses:

public static class **ChecksumFileSet**  
 extends [GenericDataFileSet](#)

An ChecksumFileSet object represents a set of [ChecksumFiles](#).

### Usage

```
ChecksumFileSet(...)
```

### Arguments

...                    Arguments passed to [GenericDataFileSet](#).

### Fields and Methods

#### Methods:

```
readChecksums -
validate      -
```

**Methods inherited from GenericDataFileSet:**

[, [[, anyDuplicated, anyNA, append, appendFiles, appendFullNamesTranslator, appendFullNamesTranslatorByNULL, appendFullNamesTranslatorByTabularTextFile, appendFullNamesTranslatorByTabularTextFileSet, appendFullNamesTranslatorBydata.frame, appendFullNamesTranslatorByfunction, appendFullNamesTranslatorBylist, as.character, as.list, byName, byPath, c, clearCache, clearFullNamesTranslator, clone, copyTo, dsApply, dsApplyInPairs, duplicated, equals, extract, findByName, findDuplicated, getChecksum, getChecksumFileSet, getChecksumObjects, getDefaultFullName, getFile, getFileClass, getFileSize, getFiles, getFullNames, getNames, getOneFile, getPath, getPathnames, getSubdirs, gunzip, gzip, hasFile, indexOf, is.na, names, nbrOfFiles, rep, resetFullNames, setFullNamesTranslator, sortBy, unique, update2, updateFullName, updateFullNames, validate

**Methods inherited from FullNameInterface:**

appendFullNameTranslator, appendFullNameTranslatorByNULL, appendFullNameTranslatorByTabularTextFile, appendFullNameTranslatorByTabularTextFileSet, appendFullNameTranslatorBycharacter, appendFullNameTranslatorBydata.frame, appendFullNameTranslatorByfunction, appendFullNameTranslatorBylist, clearFullNameTranslator, clearListOfFullNameTranslators, getDefaultFullName, getFullName, getFullNameTranslator, getListOfFullNameTranslators, getName, getTags, hasTag, hasTags, resetFullName, setFullName, setFullNameTranslator, setListOfFullNameTranslators, setName, setTags, updateFullName

**Methods inherited from Object:**

\$, \$<-, [[, [[<-, as.character, attach, attachLocally, clearCache, clearLookupCache, clone, detach, equals, extend, finalize, getEnvironment, getFieldModifier, getFieldModifiers, getFields, getInstantiationTime, getStaticInstance, hasField, hashCode, ll, load, names, objectSize, print, save

**Author(s)**

Henrik Bengtsson

---

ColumnNamesInterface *The ColumnNamesInterface class interface*

---

**Description**

Package: R.filesets

**Class ColumnNamesInterface****Interface**

```
~~|
~~+--ColumnNamesInterface
```

**Directly known subclasses:**

*GenericTabularFile, TabularTextFile*

```
public abstract class ColumnNamesInterface
  extends Interface
```

**Usage**

```
ColumnNamesInterface(...)
```

**Arguments**

```
...          Not used.
```

**Fields and Methods****Methods:**

|   |                             |
|---|-----------------------------|
| <code>clearColumnNamesTranslator</code> | -                           |
| <code>getColumnNames</code>             | Gets the column names.      |
| <code>nbrOfColumns</code>               | Gets the number of columns. |
| <code>setColumnNames</code>             | Sets the column names.      |
| <code>setColumnNamesTranslator</code>   | -                           |

**Methods inherited from Interface:**

extend, print, uses

**Author(s)**

Henrik Bengtsson

---

FullNameInterface      *The FullNameInterface class interface*

---

**Description**

Package: R.filesets

**Class FullNameInterface****Interface**

```
~~|
~~+--FullNameInterface
```

**Directly known subclasses:**

*ChecksumFile*, *ChecksumFileSet*, *GenericDataFile*, *GenericDataFileSet*, *GenericDataFileSetList*, *GenericTabularFile*, *GenericTabularFileSet*, *RDataFile*, *RDataFileSet*, *RdsFile*, *RdsFileSet*, *TabularTextFile*, *TabularTextFileSet*

```
public abstract class FullNameInterface
extends Interface
```

**Usage**

```
FullNameInterface(...)
```

**Arguments**

```
...          Not used.
```

**Details**

The full name consists of a name followed by optional comma-separated tags. For instance, the full name of foo, a.2, b has name foo with tags a.2 and b.

**Fields and Methods****Methods:**

|                                       |   |
|---------------------------------------|---|
| <code>appendFullNameTranslator</code> | -   |
| <code>clearFullNameTranslator</code>  | -   |
| <code>getFullName</code>              | Gets the full name.   |
| <code>getName</code>                  | Gets the name.  |
| <code>getTags</code>                  | Gets the tags.  |
| <code>hasTag</code>                   | -   |
| <code>hasTags</code>                  | Checks whether the fullname contains a given set of tag(s). |
| <code>setFullName</code>              | Sets the full name.   |
| <code>setFullNameTranslator</code>    | -   |
| <code>setName</code>                  | Sets the name part of the fullname.                         |
| <code>setTags</code>                  | Sets the tags.  |

**Methods inherited from Interface:**

```
extend, print, uses
```

**Author(s)**

```
Henrik Bengtsson
```

**Examples**

```
# Setup a file set
path <- system.file("R", package="R.filesets")
ds <- GenericDataFileSet$byPath(path)

# -----
# Data set
# -----
cat("Path of data set:\n")
print(getPath(ds))

cat("Fullname of data set:\n")
```

```

print(getFullName(ds))

# -----
# Data files
# -----
cat("Pathnames:\n")
print(getPathnames(ds))

cat("Filenames:\n")
print(sapply(ds, getFilename))

cat("Default fullnames:\n")
print(getFullNames(ds))

cat("Extensions:\n")
print(sapply(ds, getExtension))

# -----
# Translation of data file names
# -----
# Translate fullnames to lower case
setFullNamesTranslator(ds, function(names, ...) tolower(names))
cat("Lower-case fullnames:\n")
print(getFullNames(ds))

# Append a translator that reverse the order of the letters
revStr <- function(names, ...) {
  names <- strsplit(names, split="", fixed=TRUE)
  names <- lapply(names, FUN=rev)
  names <- sapply(names, FUN=paste, collapse="")
  names
}
appendFullNamesTranslator(ds, revStr)
cat("Reversed lower-case fullnames:\n")
fn3 <- getFullNames(ds)
print(fn3)

# Alternative for setting up a sequence of translators
setFullNamesTranslator(ds, list(function(names, ...) tolower(names), revStr))
cat("Reversed lower-case fullnames:\n")
fn3b <- getFullNames(ds)
print(fn3b)
stopifnot(identical(fn3b, fn3))

# Reset
clearFullNamesTranslator(ds)
cat("Default fullnames (after resetting):\n")
print(getFullNames(ds))

```

GenericDataFile

*The abstract GenericDataFile class***Description**

Package: R.filesets

**Class GenericDataFile**

Object

~~|

~~+---FullNameInterface

~~~~~|

~~~~~+---GenericDataFile

**Directly known subclasses:***ChecksumFile, GenericTabularFile, RDataFile, RdsFile, TabularTextFile*public abstract static class **GenericDataFile**extends *FullNameInterface*

A GenericDataFile is an object referring to a data file on a file system. Note that this class is abstract and can not be instantiated, but instead you have to use one of the subclasses or the generic `*fromFile()` method.

**Usage**

```
GenericDataFile(filename=NULL, path=NULL, mustExist=!is.na(filename), ...,
  .onUnknownArgs=c("error", "warning", "ignore"))
```

**Arguments**

|                |  |
|----------------|--|
| filename       | The filename of the file.  |
| path           | An optional path to the file.  |
| mustExist      | If <code>TRUE</code> , an exception is thrown if the file does not exist, otherwise not.           |
| ...            | Not used.  |
| .onUnknownArgs | A <code>character</code> string specifying what should occur if there are unknown arguments in ... |

**Fields and Methods****Methods:**

|                              |   |
|------------------------------|---|
| <code>compareChecksum</code> | Compares the file checksum with the value of the checksum file. |
| <code>equals</code>          | Checks if a file equals another.                                |

|                               |  |
|-------------------------------|--|
| <code>getChecksum</code>      | Gets the checksum of a file.   |
| <code>getChecksumFile</code>  | -  |
| <code>getExtension</code>     | Gets the filename extension.   |
| <code>getFileSize</code>      | Gets the size of a file.   |
| <code>getFileType</code>      | Gets the file type of a file.  |
| <code>getFilename</code>      | Gets the filename of the file.                                       |
| <code>getPath</code>          | Gets the path (directory) of the file.                               |
| <code>getPathname</code>      | Gets the pathname of the file.                                       |
| <code>is.na</code>            | -  |
| <code>isFile</code>           | Checks if this is an existing file.                                  |
| <code>validateChecksum</code> | Asserts that the file checksum matches the one of the checksum file. |
| <code>writeChecksum</code>    | Write the file checksum to a checksum file.                          |

#### Methods inherited from FullNameInterface:

`appendFullNameTranslator`, `appendFullNameTranslatorByNULL`, `appendFullNameTranslatorByTabularTextFile`, `appendFullNameTranslatorByTabularTextFileSet`, `appendFullNameTranslatorBycharacter`, `appendFullNameTranslatorBydata.frame`, `appendFullNameTranslatorByfunction`, `appendFullNameTranslatorBylist`, `clearFullNameTranslator`, `clearListOfFullNameTranslators`, `getDefaultFullName`, `getFullName`, `getFullNameTranslator`, `getListOfFullNameTranslators`, `getName`, `getTags`, `hasTag`, `hasTags`, `resetFullName`, `setFullName`, `setFullNameTranslator`, `setListOfFullNameTranslators`, `setName`, `setTags`, `updateFullName`

#### Methods inherited from Object:

`$`, `$<-`, `[]`, `[[<-`, `as.character`, `attach`, `attachLocally`, `clearCache`, `clearLookupCache`, `clone`, `detach`, `equals`, `extend`, `finalize`, `getEnvironment`, `getFieldModifier`, `getFieldModifiers`, `getFields`, `getInstantiationTime`, `getStaticInstance`, `hasField`, `hashCode`, `ll`, `load`, `names`, `objectSize`, `print`, `save`

### Filename convention

The filename of an `GenericDataFile` is structured as follows:

- filename: "sample001,a,b,c.CEL" (this follows the R convention, but not the Unix convention)
- fullname: "sample001,a,b,c"
- name: "sample001"
- tags: c("a", "b", "c")
- extension: "CEL"

### Author(s)

Henrik Bengtsson

### See Also

An object of this class is typically part of an [GenericDataFileSet](#).

---

GenericDataFileSet      *The GenericDataFileSet class*

---

## Description

Package: R.filesets

**Class GenericDataFileSet**

Object

~~|

~~+---FullNameInterface

~~~~~|

~~~~~+---GenericDataFileSet

**Directly known subclasses:**

[ChecksumFileSet](#), [GenericTabularFileSet](#), [RDataFileSet](#), [RdsFileSet](#), [TabularTextFileSet](#)

public static class **GenericDataFileSet**

extends *FullNameInterface*

A GenericDataFileSet object represents a set of [GenericDataFiles](#).

## Usage

```
GenericDataFileSet(files=NULL, tags="*", depth=NULL, ...,
  .onUnknownArgs=c("error", "warning", "ignore"))
```

## Arguments

|                |   |
|----------------|---|
| files          | A <a href="#">list</a> of <a href="#">GenericDataFile</a> :s or a <a href="#">GenericDataFileSet</a> .  |
| tags           | A <a href="#">character vector</a> of tags to be used for this file set. The string "*" indicates that it should be replaced by the tags part of the file set pathname. |
| depth          | An non-negative <a href="#">integer</a> .   |
| ...            | Not used.   |
| .onUnknownArgs | A <a href="#">character</a> string specifying what should occur if there are unknown arguments in ...   |

## Fields and Methods

**Methods:**

|               |   |
|---------------|---|
| anyDuplicated | - |
| anyNA         | - |
| append        | - |

```

appendFiles      -
as.list          -
byName           -
byPath           -
duplicated       -
equals           -
extract          -
getChecksum      -
getChecksumFileSet -
getDefaultFullName -
getFile          -
getFileClass     -
getFileSize      -
getFullNames     -
getNames         -
getOneFile       -
getPath          -
getPathnames     -
gunzip           -
gzip             -
hasFile          -
indexOf          -
is.na            -
sortBy          -
unique           -
validate         -

```

**Methods inherited from FullNameInterface:**

appendFullNameTranslator, appendFullNameTranslatorByNULL, appendFullNameTranslatorByTabularTextFile, appendFullNameTranslatorByTabularTextFileSet, appendFullNameTranslatorBycharacter, appendFullNameTranslatorBydata.frame, appendFullNameTranslatorByfunction, appendFullNameTranslatorBylist, clearFullNameTranslator, clearListOfFullNameTranslators, getDefaultFullName, getFullName, getFullNameTranslator, getListOfFullNameTranslators, getName, getTags, hasTag, hasTags, resetFullName, setFullName, setFullNameTranslator, setListOfFullNameTranslators, setName, setTags, updateFullName

**Methods inherited from Object:**

\$, \$<-, [, [(<-, as.character, attach, attachLocally, clearCache, clearLookupCache, clone, detach, equals, extend, finalize, getEnvironment, getFieldModifier, getFieldModifiers, getFields, getInstantiationTime, getStaticInstance, hasField, hashCode, ll, load, names, objectSize, print, save

**Author(s)**

Henrik Bengtsson

**Examples**

```
# Setup a file set
```

```

path <- system.file(package="R.filesets")
ds <- GenericDataFileSet$byPath(path)

# -----
# Data set
# -----
cat("Path of data set:\n")
print(getPath(ds))

cat("Fullname of data set:\n")
print(getFullName(ds))

# -----
# Data files
# -----
cat("Pathnames:\n")
print(getPathnames(ds))

cat("Filenames:\n")
print(sapply(ds, getFilename))

cat("Extensions:\n")
print(sapply(ds, getExtension))

# -----
# Subsetting
# -----
n <- length(ds)
ds2 <- extract(ds, 1:n)
print(ds2)

ds3 <- extract(ds, n:1)
print(ds3)

stopifnot(identical(rev(getPathnames(ds3)), getPathnames(ds2)))

idxs <- c(1,2,NA,n,NA)
ds4 <- extract(ds, idxs, onMissing="NA")
print(ds4)
print(getFullNames(ds4))
print(getFiles(ds4))

stopifnot(identical(is.na(idxs), unname(is.na(getPathnames(ds4))))))

```

---

GenericDataFileSetList

*The GenericDataFileSetList class*

---

**Description**

Package: R.filesets

**Class GenericDataFileSetList****Object**

```

~~|
~~+---FullNameInterface
~~~~~|
~~~~~+---GenericDataFileSetList

```

**Directly known subclasses:**

```

public static class GenericDataFileSetList
extends FullNameInterface

```

A GenericDataFileSetList object represents a list of [GenericDataFileSets](#).

**Usage**

```

GenericDataFileSetList(dsList=list(), tags="*", ..., allowDuplicates=TRUE,
  .setClass="GenericDataFileSet")

```

**Arguments**

|                 |   |
|-----------------|---|
| dsList          | A single or a <a href="#">list</a> of <a href="#">GenericDataFileSet</a> :s.                                      |
| tags            | A <a href="#">character vector</a> of tags.   |
| ...             | Not used.   |
| allowDuplicates | If <a href="#">FALSE</a> , files with duplicated names are not allowed and an exception is thrown, otherwise not. |
| .setClass       | A <a href="#">character</a> string specifying a name of the class that each data set must be an instance of.      |

**Fields and Methods****Methods:**

```

as -
as.GenericDataFileSetList -
as.list -
getFileList -
getFullNames -
getNames -
getSet -
getSets -

```

|           |   |
|-----------|---|
| indexOf   | - |
| length    | - |
| nbrOfSets | - |

**Methods inherited from FullNameInterface:**

appendFullNameTranslator, appendFullNameTranslatorByNULL, appendFullNameTranslatorByTabularTextFile, appendFullNameTranslatorByTabularTextFileSet, appendFullNameTranslatorBycharacter, appendFullNameTranslatorBydata.frame, appendFullNameTranslatorByfunction, appendFullNameTranslatorBylist, clearFullNameTranslator, clearListOfFullNameTranslators, getDefaultFullName, getFullName, getFullNameTranslator, getListOfFullNameTranslators, getName, getTags, hasTag, hasTags, resetFullName, setFullName, setFullNameTranslator, setListOfFullNameTranslators, setName, setTags, updateFullName

**Methods inherited from Object:**

\$, \$<-, [[, [[<-, as.character, attach, attachLocally, clearCache, clearLookupCache, clone, detach, equals, extend, finalize, getEnvironment, getFieldModifier, getFieldModifiers, getFields, getInstantiationTime, getStaticInstance, hasField, hashCode, ll, load, names, objectSize, print, save

**Author(s)**

Henrik Bengtsson

**Examples**

```
# Setup a file set
path1 <- system.file(package="R.filesets")
ds1 <- GenericDataFileSet$byPath(path1)

path2 <- system.file(package="R.utils")
ds2 <- GenericDataFileSet$byPath(path2)

dsl <- GenericDataFileSetList(list(ds1, ds2), tags=c("*", "CustomTag"))
print(dsl)

df <- as.data.frame(dsl)
print(df)

print(df["DESCRIPTION", "R.filesets"])
```

**Description**

Package: R.filesets

**Class GenericTabularFile****Object**

```

~~|
~~+--FullNameInterface
~~~~~|
~~~~~+--GenericDataFile
~~~~~|
~~~~~+--ColumnNamesInterface
~~~~~|
~~~~~+--GenericTabularFile

```

**Directly known subclasses:**

[TabularTextFile](#)

```

public abstract static class GenericTabularFile
extends ColumnNamesInterface

```

A TabularTextFile is an object referring to a tabular text file on a file system containing data in a tabular format. Methods for reading all or a subset of the tabular data exist.

**Usage**

```
GenericTabularFile(..., .verify=TRUE, verbose=FALSE)
```

**Arguments**

... Arguments passed to [GenericDataFile](#).

.verify, verbose (Internal only) If **TRUE**, the file is verified while the object is instantiated by the constructor. The verbose argument is passed to the verifier function.

**Fields and Methods****Methods:**

```

dim -
extractMatrix -
head -
nbrOfColumns -
nbrOfRows -
readColumns -
readDataFrame -
tail -

```

writeColumnsToFiles -

**Methods inherited from ColumnNamesInterface:**

appendColumnNamesTranslator, appendColumnNamesTranslatorByNULL, appendColumnNamesTranslatorBycharacter, appendColumnNamesTranslatorByfunction, appendColumnNamesTranslatorBylist, clearColumnNamesTranslator, clearListOfColumnNamesTranslators, getColumnNames, getColumnNamesTranslator, getDefaultColumnNames, getListOfColumnNamesTranslators, nbrOfColumns, setColumnNames, setColumnNamesTranslator, setListOfColumnNamesTranslators, updateColumnNames

**Methods inherited from GenericDataFile:**

as.character, clone, compareChecksum, copyTo, equals, fromFile, getAttribute, getAttributes, getChecksum, getChecksumFile, getCreatedOn, getDefaultFullName, getExtension, getExtensionPattern, getFileSize, getFileType, getFilename, getFilenameExtension, getLastAccessedOn, getLastModifiedOn, getOutputExtension, getPath, getPathname, gunzip, gzip, hasBeenModified, is.na, isFile, isGzipped, linkTo, readChecksum, renameTo, renameToUpperCaseExt, setAttribute, setAttributes, setAttributesBy, setAttributesByTags, setExtensionPattern, testAttributes, validate, validateChecksum, writeChecksum

**Methods inherited from FullNameInterface:**

appendFullNameTranslator, appendFullNameTranslatorByNULL, appendFullNameTranslatorByTabularTextFile, appendFullNameTranslatorByTabularTextFileSet, appendFullNameTranslatorBycharacter, appendFullNameTranslatorBydata.frame, appendFullNameTranslatorByfunction, appendFullNameTranslatorBylist, clearFullNameTranslator, clearListOfFullNameTranslators, getDefaultFullName, getFullName, getFullNameTranslator, getListOfFullNameTranslators, getName, getTags, hasTag, hasTags, resetFullName, setFullName, setFullNameTranslator, setListOfFullNameTranslators, setName, setTags, updateFullName

**Methods inherited from Object:**

\$, \$<-, [[, [[<-, as.character, attach, attachLocally, clearCache, clearLookupCache, clone, detach, equals, extend, finalize, getEnvironment, getFieldModifier, getFieldModifiers, getFields, getInstantiationTime, getStaticInstance, hasField, hashCode, ll, load, names, objectSize, print, save

**Author(s)**

Henrik Bengtsson

**See Also**

An object of this class is typically part of an [GenericTabularFileSet](#).

**Description**

Package: R.filesets

**Class GenericTabularFileSet****Object**

```

~~|
~~+---FullNameInterface
~~~~~|
~~~~~+---GenericDataFileSet
~~~~~|
~~~~~+---GenericTabularFileSet

```

**Directly known subclasses:**

[TabularTextFileSet](#)

```

public static class GenericTabularFileSet
  extends GenericDataFileSet

```

An `GenericTabularFileSet` object represents a set of [GenericTabularFiles](#).

**Usage**

```
GenericTabularFileSet(...)
```

**Arguments**

... Arguments passed to [GenericDataFileSet](#).

**Fields and Methods****Methods:**

`extractMatrix` -

**Methods inherited from GenericDataFileSet:**

[, [[, anyDuplicated, anyNA, append, appendFiles, appendFullNamesTranslator, appendFullNamesTranslatorByNULL, appendFullNamesTranslatorByTabularTextFile, appendFullNamesTranslatorByTabularTextFileSet, appendFullNamesTranslatorBydata.frame, appendFullNamesTranslatorByfunction, appendFullNamesTranslatorBylist, as.character, as.list, byName, byPath, c, clearCache, clearFullNamesTranslator, clone, copyTo, dsApply, dsApplyInPairs, duplicated, equals, extract, findByName, findDuplicated, getChecksum, getChecksumFileSet, getChecksumObjects, getDefaultFullName, getFile, getFileClass, getFileSize, getFiles, getFullNames, getNames, getOneFile, getPath, getPathnames, getSubdirs, gunzip, gzip, hasFile, indexOf, is.na, names, nbrOfFiles, rep, resetFullNames, setFullNamesTranslator, sortBy, unique, update2, updateFullName, updateFullNames, validate

**Methods inherited from FullNameInterface:**

appendFullNameTranslator, appendFullNameTranslatorByNULL, appendFullNameTranslatorByTabularTextFile, appendFullNameTranslatorByTabularTextFileSet, appendFullNameTranslatorBycharacter, appendFullNameTranslatorBydata.frame, appendFullNameTranslatorByfunction, appendFullNameTranslatorBylist, clearFullNameTranslator, clearListOfFullNameTranslators, getDefaultFullName, getFullName, getFullNameTranslator, getListOfFullNameTranslators, getName, getTags, hasTag, hasTags, resetFullName, setFullName, setFullNameTranslator, setListOfFullNameTranslators, setName, setTags, updateFullName

**Methods inherited from Object:**

\$, \$<-, [[, [[<-, as.character, attach, attachLocally, clearCache, clearLookupCache, clone, detach, equals, extend, finalize, getEnvironment, getFieldModifier, getFieldModifiers, getFields, getInstantiationTime, getStaticInstance, hasField, hashCode, ll, load, names, objectSize, print, save

**Author(s)**

Henrik Bengtsson

---

loadRDS

*Reads data from a RDS file*

---

**Description**

Reads data from a RDS file.

**Usage**

```
## Default S3 method:
loadRDS(file, ...)
## S3 method for class 'RdsFile'
loadRDS(file, ...)
```

**Arguments**

`file` A [character](#) string, a [connection](#), or an [RdsFile](#) specifying a RDS file/connection to be read.

`...` Additional arguments passed to [readRDS\(\)](#).

**Value**

Returns an R object.

**Author(s)**

Henrik Bengtsson

**See Also**

[readRDS\(\)](#).

---

|           |                                   |
|-----------|-----------------------------------|
| loadToEnv | <i>Reads data from a RDS file</i> |
|-----------|-----------------------------------|

---

**Description**

Reads data from a RDS file.

**Usage**

```
## S3 method for class 'RDataFile'
loadToEnv(file, ...)
```

**Arguments**

|      |   |
|------|---|
| file | A <a href="#">character</a> string, a <a href="#">connection</a> , or an <a href="#">RDataFile</a> specifying an RData file to be read. |
| ...  | Additional arguments passed to <a href="#">loadToEnv</a> .  |

**Value**

Returns an [environment](#).

**Author(s)**

Henrik Bengtsson

**See Also**

[loadToEnv](#).

---

|           |                            |
|-----------|----------------------------|
| RDataFile | <i>The RDataFile class</i> |
|-----------|----------------------------|

---

**Description**

Package: R.filesets

**Class RDataFile****Object**

```
~|
~+--FullNameInterface
~~~~~|
~~~~~+--GenericDataFile
~~~~~|
~~~~~+--RDataFile
```

**Directly known subclasses:**

public abstract static class **RDataFile**  
 extends *GenericDataFile*

An RDataFile represents a binary file containing R objects saved using the `save()` function.

**Usage**

```
RDataFile(...)
```

**Arguments**

... Arguments passed to *GenericDataFile*.

**Fields and Methods****Methods:**

```
loadObject -  
loadToEnv -
```

**Methods inherited from GenericDataFile:**

as.character, clone, compareChecksum, copyTo, equals, fromFile, getAttribute, getAttributes, getChecksum, getChecksumFile, getCreatedOn, getDefaultFullName, getExtension, getExtensionPattern, getFileSize, getFileType, getFilename, getFilenameExtension, getLastAccessedOn, getLastModifiedOn, getOutputExtension, getPath, getPathname, gunzip, gzip, hasBeenModified, is.na, isFile, isGzipped, linkTo, readChecksum, renameTo, renameToUpperCaseExt, setAttribute, setAttributes, setAttributesBy, setAttributesByTags, setExtensionPattern, testAttributes, validate, validateChecksum, writeChecksum

**Methods inherited from FullNameInterface:**

appendFullNameTranslator, appendFullNameTranslatorByNULL, appendFullNameTranslatorByTabularTextFile, appendFullNameTranslatorByTabularTextFileSet, appendFullNameTranslatorBycharacter, appendFullNameTranslatorBydata.frame, appendFullNameTranslatorByfunction, appendFullNameTranslatorBylist, clearFullNameTranslator, clearListOfFullNameTranslators, getDefaultFullName, getFullName, getFullNameTranslator, getListOfFullNameTranslators, getName, getTags, hasTag, hasTags, resetFullName, setFullName, setFullNameTranslator, setListOfFullNameTranslators, setName, setTags, updateFullName

**Methods inherited from Object:**

\$, \$<-, [[, [[<-, as.character, attach, attachLocally, clearCache, clearLookupCache, clone, detach, equals, extend, finalize, getEnvironment, getFieldModifier, getFieldModifiers, getFields, getInstantiationTime, getStaticInstance, hasField, hashCode, ll, load, names, objectSize, print, save

**Author(s)**

Henrik Bengtsson

**See Also**An object of this class is typically part of an [RDataFileSet](#).

---

RDataFileSet*The RDataFileSet class*

---

**Description**

Package: R.filesets

**Class RDataFileSet****Object**

```

~~|
~~+--FullNameInterface
~~~~~|
~~~~~+--GenericDataFileSet
~~~~~|
~~~~~+--RDataFileSet

```

**Directly known subclasses:**

```

public static class RDataFileSet
  extends GenericDataFileSet

```

An RDataFileSet object represents a set of [RDataFile](#)s.**Usage**

RDataFileSet(...)

**Arguments**... Arguments passed to [GenericDataFileSet](#).**Fields and Methods****Methods:**

byPath -

**Methods inherited from GenericDataFileSet:**

[, [[, anyDuplicated, anyNA, append, appendFiles, appendFullNamesTranslator, appendFullNamesTranslatorByNULL, appendFullNamesTranslatorByTabularTextFile, appendFullNamesTranslatorByTabularTextFileSet, appendFullNamesTranslatorBydata.frame, appendFullNamesTranslatorByfunction, appendFullNamesTranslatorBylist, as.character, as.list, byName, byPath, c, clearCache, clearFullNamesTranslator, clone, copyTo, dsApply, dsApplyInPairs, duplicated, equals, extract, findByName, findDuplicated, getChecksum, getChecksumFileSet, getChecksumObjects, getDefaultFullName, getFile, getFileClass, getFileSize, getFiles, getFullNames, getNames, getOneFile, getPath, getPathnames, getSubdirs, gunzip, gzip, hasFile, indexOf, is.na, names, nbrOfFiles, rep, resetFullNames, setFullNamesTranslator, sortBy, unique, update2, updateFullName, updateFullNames, validate

**Methods inherited from FullNameInterface:**

appendFullNameTranslator, appendFullNameTranslatorByNULL, appendFullNameTranslatorByTabularTextFile, appendFullNameTranslatorByTabularTextFileSet, appendFullNameTranslatorBycharacter, appendFullNameTranslatorBydata.frame, appendFullNameTranslatorByfunction, appendFullNameTranslatorBylist, clearFullNameTranslator, clearListOfFullNameTranslators, getDefaultFullName, getFullName, getFullNameTranslator, getListOfFullNameTranslators, getName, getTags, hasTag, hasTags, resetFullName, setFullName, setFullNameTranslator, setListOfFullNameTranslators, setName, setTags, updateFullName

**Methods inherited from Object:**

\$, \$<-, [[, [[<-, as.character, attach, attachLocally, clearCache, clearLookupCache, clone, detach, equals, extend, finalize, getEnvironment, getFieldModifier, getFieldModifiers, getFields, getInstantiationTime, getStaticInstance, hasField, hashCode, ll, load, names, objectSize, print, save

**Author(s)**

Henrik Bengtsson

---

RdsFile

*The RdsFile class*

---

**Description**

Package: R.filesets

**Class RdsFile****Object**

```

~~|
~~+--FullNameInterface
~~~~~|
~~~~~+--GenericDataFile
~~~~~|
~~~~~+--RdsFile

```

**Directly known subclasses:**

```
public abstract static class RdsFile
  extends GenericDataFile
```

An RdsFile represents a binary file containing an R object saved using the `saveRDS()` function.

### Usage

```
RdsFile(...)
```

### Arguments

```
... Arguments passed to GenericDataFile.
```

### Fields and Methods

#### Methods:

```
loadObject -
loadRDS -
```

#### Methods inherited from *GenericDataFile*:

as.character, clone, compareChecksum, copyTo, equals, fromFile, getAttribute, getAttributes, getChecksum, getChecksumFile, getCreatedOn, getDefaultFullName, getExtension, getExtensionPattern, getFileSize, getFileType, getFilename, getFilenameExtension, getLastAccessedOn, getLastModifiedOn, getOutputExtension, getPath, getPathname, gunzip, gzip, hasBeenModified, is.na, isFile, isGzipped, linkTo, readChecksum, renameTo, renameToUpperCaseExt, setAttribute, setAttributes, setAttributesBy, setAttributesByTags, setExtensionPattern, testAttributes, validate, validateChecksum, writeChecksum

#### Methods inherited from *FullNameInterface*:

appendFullNameTranslator, appendFullNameTranslatorByNULL, appendFullNameTranslatorByTabularTextFile, appendFullNameTranslatorByTabularTextFileSet, appendFullNameTranslatorBycharacter, appendFullNameTranslatorBydata.frame, appendFullNameTranslatorByfunction, appendFullNameTranslatorBylist, clearFullNameTranslator, clearListOfFullNameTranslators, getDefaultFullName, getFullName, getFullNameTranslator, getListOfFullNameTranslators, getName, getTags, hasTag, hasTags, resetFullName, setFullName, setFullNameTranslator, setListOfFullNameTranslators, setName, setTags, updateFullName

#### Methods inherited from *Object*:

\$, \$<-, [[, [[<-, as.character, attach, attachLocally, clearCache, clearLookupCache, clone, detach, equals, extend, finalize, getEnvironment, getFieldModifier, getFieldModifiers, getFields, getInstantiationTime, getStaticInstance, hasField, hashCode, ll, load, names, objectSize, print, save

### Author(s)

Henrik Bengtsson

**See Also**

An object of this class is typically part of an [RdsFileSet](#).

---

RdsFileSet

*The RdsFileSet class*


---

**Description**

Package: R.filesets

**Class RdsFileSet**

Object

```

~~|
~~+--FullNameInterface
~~~~~|
~~~~~+--GenericDataFileSet
~~~~~|
~~~~~+--RdsFileSet

```

**Directly known subclasses:**

```

public static class RdsFileSet
extends GenericDataFileSet

```

An RdsFileSet object represents a set of [RdsFile](#):s.

**Usage**

```
RdsFileSet(...)
```

**Arguments**

... Arguments passed to [GenericDataFileSet](#).

**Fields and Methods****Methods:**

byPath -

**Methods inherited from GenericDataFileSet:**

[, [[, anyDuplicated, anyNA, append, appendFiles, appendFullNamesTranslator, appendFullNames-

TranslatorByNULL, appendFullNamesTranslatorByTabularTextFile, appendFullNamesTranslatorByTabularTextFileSet, appendFullNamesTranslatorBydata.frame, appendFullNamesTranslatorByfunction, appendFullNamesTranslatorBylist, as.character, as.list, byName, byPath, c, clearCache, clearFullNamesTranslator, clone, copyTo, dsApply, dsApplyInPairs, duplicated, equals, extract, findByName, findDuplicated, getChecksum, getChecksumFileSet, getChecksumObjects, getDefaultFullName, getFile, getFileClass, getFileSize, getFiles, getFullNames, getNames, getOneFile, getPath, getPathnames, getSubdirs, gunzip, gzip, hasFile, indexOf, is.na, names, nbrOfFiles, rep, resetFullNames, setFullNamesTranslator, sortBy, unique, update2, updateFullName, updateFullNames, validate

**Methods inherited from FullNameInterface:**

appendFullNameTranslator, appendFullNameTranslatorByNULL, appendFullNameTranslatorByTabularTextFile, appendFullNameTranslatorByTabularTextFileSet, appendFullNameTranslatorBycharacter, appendFullNameTranslatorBydata.frame, appendFullNameTranslatorByfunction, appendFullNameTranslatorBylist, clearFullNameTranslator, clearListOfFullNameTranslators, getDefaultFullName, getFullName, getFullNameTranslator, getListOfFullNameTranslators, getName, getTags, hasTag, hasTags, resetFullName, setFullName, setFullNameTranslator, setListOfFullNameTranslators, setName, setTags, updateFullName

**Methods inherited from Object:**

\$, \$<-, [[, [[<-, as.character, attach, attachLocally, clearCache, clearLookupCache, clone, detach, equals, extend, finalize, getEnvironment, getFieldModifier, getFieldModifiers, getFields, getInstantiationTime, getStaticInstance, hasField, hashCode, ll, load, names, objectSize, print, save

**Author(s)**

Henrik Bengtsson

---

|               |                                       |
|---------------|---------------------------------------|
| readDataFrame | <i>Reads data from a tabular file</i> |
|---------------|---------------------------------------|

---

**Description**

Reads data from a tabular file or a set of such files.

**Usage**

```
## Default S3 method:
readDataFrame(filename, path=NULL, ...)
```

**Arguments**

filename, path [character vector](#) specifying one or more files to be read.  
 ... Additional arguments passed to either (i) [readDataFrame](#) for class [TabularTextFile](#), or (ii) [readDataFrame](#) for class [TabularTextFileSet](#), depending on whether one or multiple files are read.

**Details**

When reading multiple files at once, first each file is read into a `data.frame`, and then these `data.frames` are (by default) merged into one `data.frame` using `rbind()`. This requires that the same set of columns are read for each file. Which columns to read can be controlled by specifying their names in argument `colClasses`. To change how the `data.frames` are merged, use argument `combineBy`. For more information, follow the help on the above to `readDataFrame()` help links.

**Value**

Returns a `data.frame`.

**Author(s)**

Henrik Bengtsson

**See Also**

`read.table`. For further details, see classes `TabularTextFile` and `TabularTextFileSet`.

**Examples**

```
path <- system.file("exData/dataSetA,original", package="R.filesets")

# - - - - -
# Example: Standard tab-delimited file with header comments
# - - - - -
pathname <- file.path(path, "fileA,20100112.dat")

# Read all data
df <- readDataFrame(pathname)
print(df)

# Read columns 'x', 'y', and 'char'
df <- readDataFrame(pathname, colClasses=c("(x|y)"="integer", "char"="character"))
print(df)

# - - - - -
# Example: Tab-delimited file with header comments but
#           also two garbage at the very beginning
# - - - - -
pathname <- file.path(path, "fileA,20130116.datx")

# Explicitly skip the two rows
df <- readDataFrame(pathname, skip=2)
print(df)

# Skip until the first data row
df <- readDataFrame(pathname, skip="^x")
print(df)
```

```

# -----
# Example: Tab-delimited file without column header
# -----
path <- system.file("exData/dataSetB", package="R.filesets")
pathname <- file.path(path, "fileF,noHeader.dat")

# Incorrectly assuming column header
df <- readDataframe(pathname)
print(df)

# No column header
df <- readDataframe(pathname, header=FALSE)
print(df)

```

---

TabularTextFile

*The TabularTextFile class*


---

## Description

Package: R.filesets

**Class TabularTextFile**

### Object

```

~~|
~~+--FullNameInterface
~~~~~|
~~~~~+--GenericDataFile
~~~~~|
~~~~~+--ColumnNamesInterface
~~~~~|
~~~~~+--GenericTabularFile
~~~~~|
~~~~~+--TabularTextFile

```

### Directly known subclasses:

```

public abstract static class TabularTextFile
extends GenericTabularFile

```

A TabularTextFile is an object referring to a tabular text file on a file system containing data in a tabular format. Methods for reading all or a subset of the tabular data exist.

**Usage**

```
TabularTextFile(..., sep=c("\t", ","), quote="\\"", fill=FALSE, skip=0L, columnNames=NA,
  commentChar="#", .verify=TRUE, verbose=FALSE)
```

**Arguments**

... Arguments passed to [GenericTabularFile](#).

sep A [character](#) specifying the symbol used to separate the cell entries. If more than one symbol is specified, it will try to select the correct one by peeking into the file.

quote A [character](#) specifying the quote symbol used, if any.

fill As in [read.table](#).

skip As in [read.table](#).

columnNames A [logical](#) or a [character vector](#). If `TRUE`, then column names are inferred from the file. If a [character vector](#), then the column names are given by this argument.

commentChar A single [character](#) specifying which symbol should be used for comments, cf. [read.table](#).

.verify, verbose (Internal only) If `TRUE`, the file is verified while the object is instantiated by the constructor. The verbose argument is passed to the verifier function.

**Fields and Methods****Methods:**

```
getHeader      -
nbrOfLines     -
nbrOfRows      -
readDataFrame  -
readLines     -
```

**Methods inherited from GenericTabularFile:**

[, as.character, dim, extractMatrix, head, nbrOfColumns, nbrOfRows, readColumns, readDataFrame, tail, writeColumnsToFiles

**Methods inherited from ColumnNamesInterface:**

appendColumnNamesTranslator, appendColumnNamesTranslatorByNULL, appendColumnNamesTranslatorBycharacter, appendColumnNamesTranslatorByfunction, appendColumnNamesTranslatorBylist, clearColumnNamesTranslator, clearListOfColumnNamesTranslators, getColumnNames, getColumnNamesTranslator, getDefaultColumnNames, getListOfColumnNamesTranslators, nbrOfColumns, setColumnNames, setColumnNamesTranslator, setListOfColumnNamesTranslators, updateColumnNames

**Methods inherited from GenericDataFile:**

as.character, clone, compareChecksum, copyTo, equals, fromFile, getAttribute, getAttributes, getCheck-

sum, getChecksumFile, getCreatedOn, getDefaultFullName, getExtension, getExtensionPattern, getFileSize, getFileType, getFilename, getFilenameExtension, getLastAccessedOn, getLastModifiedOn, getOutputExtension, getPath, getPathname, gunzip, gzip, hasBeenModified, is.na, isFile, isGzipped, linkTo, readChecksum, renameTo, renameToUpperCaseExt, setAttribute, setAttributes, setAttributesBy, setAttributesByTags, setExtensionPattern, testAttributes, validate, validateChecksum, writeChecksum

**Methods inherited from FullNameInterface:**

appendFullNameTranslator, appendFullNameTranslatorByNULL, appendFullNameTranslatorByTabularTextFile, appendFullNameTranslatorByTabularTextFileSet, appendFullNameTranslatorBycharacter, appendFullNameTranslatorBydata.frame, appendFullNameTranslatorByfunction, appendFullNameTranslatorBylist, clearFullNameTranslator, clearListOfFullNameTranslators, getDefaultFullName, getFullName, getFullNameTranslator, getListOfFullNameTranslators, getName, getTags, hasTag, hasTags, resetFullName, setFullName, setFullNameTranslator, setListOfFullNameTranslators, setName, setTags, updateFullName

**Methods inherited from Object:**

\$, \$<-, [[, [[<-, as.character, attach, attachLocally, clearCache, clearLookupCache, clone, detach, equals, extend, finalize, getEnvironment, getFieldModifier, getFieldModifiers, getFields, getInstantiationTime, getStaticInstance, hasField, hashCode, ll, load, names, objectSize, print, save

**Author(s)**

Henrik Bengtsson

**See Also**

An object of this class is typically part of an [TabularTextFileSet](#).

**Examples**

```
path <- system.file("exData/dataSetA,original", package="R.filesets")

db <- TabularTextFile("fileA,20100112.dat", path=path)
print(db)

# Read all data
data <- readDataframe(db)
print(data)

# Read columns 'x', 'y', and 'char'
data <- readDataframe(db, colClasses=c("(x|y)"="integer", "char"="character"))
print(data)

# Translate column names on the fly
db <- setColumnNamesTranslator(db, function(names, ...) toupper(names))
data <- readDataframe(db, colClasses=c("(X|Y)"="integer", "CHAR"="character"))
print(data)
```

---

 TabularTextFileSet      *The TabularTextFileSet class*


---

**Description**

Package: R.filesets

**Class TabularTextFileSet**

Object

~~|

~~+--FullNameInterface

~~~~~|

~~~~~+--GenericDataFileSet

~~~~~|

~~~~~+--GenericTabularFileSet

~~~~~|

~~~~~+--TabularTextFileSet

**Directly known subclasses:**public static class **TabularTextFileSet**extends [GenericTabularFileSet](#)An TabularTextFileSet object represents a set of [TabularTextFiles](#).**Usage**

TabularTextFileSet(...)

**Arguments**...                    Arguments passed to [GenericTabularFileSet](#).**Fields and Methods****Methods:**

readDataFrame -

**Methods inherited from GenericTabularFileSet:**

extractMatrix

**Methods inherited from GenericDataFileSet:**

[, [[, anyDuplicated, anyNA, append, appendFiles, appendFullNamesTranslator, appendFullNamesTranslatorByNULL, appendFullNamesTranslatorByTabularTextFile, appendFullNamesTranslatorByTabularTextFileSet, appendFullNamesTranslatorBydata.frame, appendFullNamesTranslatorByfunction, appendFullNamesTranslatorBylist, as.character, as.list, byName, byPath, c, clearCache, clearFullNamesTranslator, clone, copyTo, dsApply, dsApplyInPairs, duplicated, equals, extract, findByName, findDuplicated, getChecksum, getChecksumFileSet, getChecksumObjects, getDefaultFullName, getFile, getFileClass, getFileSize, getFiles, getFullNames, getNames, getOneFile, getPath, getPathnames, getSubdirs, gunzip, gzip, hasFile, indexOf, is.na, names, nbrOfFiles, rep, resetFullNames, setFullNamesTranslator, sortBy, unique, update2, updateFullName, updateFullNames, validate

**Methods inherited from FullNameInterface:**

appendFullNameTranslator, appendFullNameTranslatorByNULL, appendFullNameTranslatorByTabularTextFile, appendFullNameTranslatorByTabularTextFileSet, appendFullNameTranslatorBycharacter, appendFullNameTranslatorBydata.frame, appendFullNameTranslatorByfunction, appendFullNameTranslatorBylist, clearFullNameTranslator, clearListOfFullNameTranslators, getDefaultFullName, getFullName, getFullNameTranslator, getListOfFullNameTranslators, getName, getTags, hasTag, hasTags, resetFullName, setFullName, setFullNameTranslator, setListOfFullNameTranslators, setName, setTags, updateFullName

**Methods inherited from Object:**

\$, \$<-, [[, [[<-, as.character, attach, attachLocally, clearCache, clearLookupCache, clone, detach, equals, extend, finalize, getEnvironment, getFieldModifier, getFieldModifiers, getFields, getInstantiationTime, getStaticInstance, hasField, hashCode, ll, load, names, objectSize, print, save

**Author(s)**

Henrik Bengtsson

**Examples**

```
# Setup a file set consisting of all *.dat tab-delimited files
# in a particular directory
path <- system.file("exData/dataSetA,original", package="R.filesets")
ds <- TabularTextFileSet$byPath(path, pattern=".dat$")
print(ds)

# Read column 'y' and a subset of the rows from each of the
# tab-delimited files and combine into a matrix
rows <- c(3:5, 8, 2)
data <- extractMatrix(ds, column="y", colClass="integer", rows=rows)
print(data)

# -----
# See also help("readDataFrame.TabularTextFileSet")
# -----

# -----
# ADVANCED: Translation of fullnames
```

```
# - - - - -
fnts <- TabularTextFileSet$byPath(getPath(ds), pattern=",fullnames[.]txt$")
appendFullNamesTranslator(ds, as.list(fnts))

cat("Default fullnames:\n")
print(head(getFullNames(ds, translate=FALSE)))
cat("Translated fullnames:\n")
print(head(getFullNames(ds)))

cat("Default fullnames:\n")
print(getFullNames(ds, translate=FALSE))
cat("Translated fullnames:\n")
print(getFullNames(ds))
```

# Index

## \*Topic **classes**

- ChecksumFile, 3
- ChecksumFileSet, 5
- ColumnNamesInterface, 6
- FullNameInterface, 7
- GenericDataFile, 10
- GenericDataFileSet, 12
- GenericDataFileSetList, 14
- GenericTabularFile, 16
- GenericTabularFileSet, 18
- RDataFile, 21
- RDataFileSet, 23
- RdsFile, 24
- RdsFileSet, 26
- TabularTextFile, 29
- TabularTextFileSet, 32

## \*Topic **methods**

- loadRDS, 20
- loadToEnv, 21

## \*Topic **package**

- R.filesets-package, 2

\*fromFile, 10

byPath.RDataFileSet (RDataFileSet), 23

byPath.RdsFileSet (RdsFileSet), 26

character, 10, 12, 15, 20, 21, 27, 30

ChecksumFile, 3, 5, 7, 10

ChecksumFileSet, 5, 7, 12

ColumnNamesInterface, 6, 17, 29

compareChecksum, 10

connection, 20, 21

data.frame, 28

environment, 21

equals, 10

FALSE, 15

FullNameInterface, 3, 5, 7, 10, 12, 15, 17,  
19, 21, 23, 24, 26, 29, 32

GenericDataFile, 3, 4, 7, 10, 12, 17, 21, 22,  
24, 25, 29

GenericDataFileSet, 3, 5, 7, 11, 12, 12, 15,  
19, 23, 26, 32

GenericDataFileSetList, 7, 14

GenericTabularFile, 6, 7, 10, 16, 19, 29, 30

GenericTabularFileSet, 7, 12, 18, 18, 32

getChecksum, 11

getChecksum, ChecksumFile-method  
(ChecksumFile), 3

getChecksum.ChecksumFile  
(ChecksumFile), 3

getColumnNames, 7

getExtension, 11

getFilename, 11

getFileSize, 11

getFileType, 11

getFullName, 8

getName, 8

getPath, 11

getPathname, 11

getTags, 8

hasTags, 8

integer, 12

Interface, 6, 7

isFile, 11

list, 12, 15

loadObject (RDataFile), 21

loadObject.RdsFile (RdsFile), 24

loadRDS, 20

loadToEnv, 21, 21

logical, 30

numberOfColumns, 7

Object, 3, 5, 10, 12, 15, 17, 19, 21, 23, 24, 26,  
29, 32

R.filesets (R.filesets-package), 2  
R.filesets-package, 2  
rbind, 28  
RDataFile, 7, 10, 21, 21, 23  
RDataFileSet, 7, 12, 23, 23  
RdsFile, 7, 10, 20, 24, 26  
RdsFileSet, 7, 12, 26, 26  
read.table, 28, 30  
readChecksum, 4  
readDataFrame, 27, 27  
readRDS, 20  
  
save, 22  
saveRDS, 25  
setColumnNames, 7  
setFullName, 8  
setName, 8  
setTags, 8  
  
TabularTextFile, 3, 6, 7, 10, 17, 27, 28, 29,  
32  
TabularTextFileSet, 7, 12, 19, 27, 28, 31, 32  
TRUE, 10, 17, 30  
  
validate, 4  
validateChecksum, 11  
vector, 12, 15, 27, 30  
  
writeChecksum, 11