

Package ‘R.filesets’

January 12, 2012

Version 1.1.4

Date 2011-11-19

Title Easy handling of and access to files organized in structured directories

Author Henrik Bengtsson <henrikb@braju.com>

Maintainer Henrik Bengtsson <henrikb@braju.com>

Depends R (>= 2.6.0), R.methodsS3 (>= 1.2.1), R.oo (>= 1.8.3), R.utils (>= 1.9.4), digest (>= 0.5.1)

Suggests gtools (>= 2.6.2)

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 suchs sets. The API is designed such that these classes can be subsetted 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 fullnames 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 <http://www.aroma-project.org/>

PatchURL <http://www.braju.com/R/>

LazyLoad TRUE

biocViews Infrastructure, DataImport, DataRepresentation

Repository CRAN

Date/Publication 2012-01-12 09:24:11

R topics documented:

R.filesets-package	2
FullNameInterface	3
GenericDataFile	5
GenericDataFileSet	7
GenericDataFileSetList	10
GenericTabularFile	12
GenericTabularFileSet	14
TabularTextFile	15
TabularTextFileSet	18

Index	21
--------------	-----------

R.filesets-package	<i>Package R.filesets</i>
--------------------	---------------------------

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 subsetting 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 fullnames 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.

Requirements

This package requires external package.

Installation and updates

To install this package, see instructions at <http://www.braju.com/R/>.

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/>

FullNameInterface *The FullNameInterface class interface*

Description

Package: R.filesets

Class FullNameInterface

Interface

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

Directly known subclasses:

[GenericDataFile](#), [GenericDataFileSet](#), [GenericDataFileSetList](#), [GenericTabularFile](#), [GenericTabularFileSet](#), [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:**

appendFullNameTranslatorByTabularTextFile	-
appendFullNameTranslatorByTabularTextFileSet	-
clearFullNameTranslator	-
getFullName	Gets the full name.
getName	Gets the name.
getTags	Gets the tags.
hasTag	-
hasTags	Checks whether the fullname contains a given set of tag(s).
setFullName	Sets the full name.
setFullNameTranslator	-
setName	Sets the name part of the fullname.
setTags	Sets the tags.

Methods inherited from Interface:

extend, print, uses

Author(s)

Henrik Bengtsson (<http://www.braju.com/R/>)

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:

GenericTabularFile, *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=TRUE, ..., .onUnknownArgs=c("error", "warning", "

```

Arguments

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

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>fromFile</code>	Defines a <code>GenericDataFile</code> from a pathname.
<code>getChecksum</code>	Gets the checksum of a file.
<code>getExtension</code>	Gets the filename extension.
<code>getFilename</code>	Gets the filename of the file.
<code>getFileSize</code>	Gets the size of a file.
<code>getFileType</code>	Gets the file type of a file.
<code>getPath</code>	Gets the path (directory) of the file.
<code>getPathname</code>	Gets the pathname of the file.
<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.

`writeChecksum` Write the file checksum to a checksum file.

Methods inherited from FullNameInterface:

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

Methods inherited from Object:

`$`, `$<-`, `[[`, `[[<-`, `as.character`, `attach`, `attachLocally`, `clearCache`, `clearLookupCache`, `clone`, `detach`, `equals`, `extend`, `finalize`, `gc`, `getEnvironment`, `getFieldModifier`, `getFieldModifiers`, `getFields`, `getInstantiationTime`, `getStaticInstance`, `hasField`, `hashCode`, `ll`, `load`, `objectSize`, `print`, `registerFinalizer`, `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"
- `tagsc`("a", "b", "c")
- `extension`"CEL"

Author(s)

Henrik Bengtsson (<http://www.braju.com/R/>)

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:

[GenericTabularFileSet](#), [TabularTextFileSet](#)

```

public static class GenericDataFileSet
extends FullNameInterface

```

A `GenericDataFileSet` object represents a set of `GenericDataFiles`.

Usage

```
GenericDataFileSet(files=NULL, tags="*", depth=NULL, alias=NULL, ..., .onUnknownArgs=c("error", "warn"
```

Arguments

<code>files</code>	A list of <code>GenericDataFile</code> :s.
<code>tags</code>	A character vector 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.
<code>depth</code>	An non-negative integer .
<code>alias</code>	A character string specifying a name alias overriding the name inferred from the pathname.
<code>...</code>	Not used.
<code>.onUnknownArgs</code>	A character string specifying what should occur if there are unknown arguments in <code>...</code>

Fields and Methods

Methods:

<code>append</code>	Appends one data set to an existing one.
<code>appendFiles</code>	Appends a list of files to a data set.
<code>as.list</code>	Returns the files of the file set.
<code>byName</code>	Locates and sets up a file set by its name.
<code>byPath</code>	Defines a <code>GenericDataFileSet</code> by searching for files in a directory.
<code>clearCache</code>	-
<code>equals</code>	-
<code>extract</code>	Extract a subset of the file set.
<code>getAlias</code>	Gets the alias of the file set.
<code>getChecksum</code>	Generates a deterministic checksum for an R object.
<code>getDefaultFullName</code>	Gets the default full name of the file set.
<code>getFile</code>	Get a particular file of the file set.
<code>getFileClass</code>	-
<code>getFileSize</code>	-

<code>getFullNames</code>	-
<code>getNames</code>	Gets the names (or fullnames) of the files in the file set.
<code>getPath</code>	Gets the path (directory) of the file set.
<code>getPathnames</code>	Gets the pathnames of the files in the file set.
<code>hasFile</code>	-
<code>indexOf</code>	Finds indices of a subset of files.
<code>lapply</code>	Applies a function to each of the data files.
<code>nbrOfFiles</code>	-
<code>sapply</code>	-
<code>seq</code>	Gets an vector of file indices.
<code>setAlias</code>	Sets the alias of the file set.
<code>sortBy</code>	Sorts the set.
<code>validate</code>	-

Methods inherited from FullNameInterface:

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

Methods inherited from Object:

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

Author(s)

Henrik Bengtsson (<http://www.braju.com/R/>)

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="GenericDataFile
```

Arguments

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

Fields and Methods

Methods:

| | |
|---------------------------|---|
| as.character | - |
| as | - |
| as.GenericDataFileSetList | - |
| as.list | - |
| clone | - |
| getDefaultFullName | - |
| getFileList | - |
| getFullNames | - |
| getNames | - |
| getSet | - |
| getSets | - |
| indexOf | - |
| length | - |
| nbrOfFiles | - |
| nbrOfSets | - |
| setTags | - |

Methods inherited from [FullNameInterface](#):

appendFullNameTranslator, appendFullNameTranslatorBycharacter, appendFullNameTranslatorBydata.frame, appendFullNameTranslatorByfunction, appendFullNameTranslatorBylist, appendFullNameTranslatorByNULL, appendFullNameTranslatorByTabularTextFile, appendFullNameTranslatorByTabularTextFileSet, clearFullNameTranslator, clearListOfFullNameTranslators, getDefault-

FullName, getFullName, getFullNameTranslator, getListOfFullNameTranslators, getName, getTags, hasTag, hasTags, setFullName, setFullNameTranslator, setListOfFullNameTranslators, setName, setTags, updateFullName

Methods inherited from Object:

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

Author(s)

Henrik Bengtsson (<http://www.braju.com/R/>)

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"])
```

GenericTabularFile *The abstract GenericTabularFile class*

Description

Package: R.filesets

Class GenericTabularFile

Object

```
~~|
~~+--FullNameInterface
~~~~~|
~~~~~+--GenericDataFile
~~~~~|
~~~~~+--GenericTabularFile
```

Directly known subclasses:*TabularTextFile*

public abstract static class **GenericTabularFile**
 extends *GenericDataFile*

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:**

| | |
|-------------------------|------------------------------------------------|
| as.character | - |
| dim | Gets the dimension of data table. |
| extractMatrix | Reads one of the columns. |
| getColumnNames | Gets the column names. |
| getColumnNameTranslator | - |
| nbrOfColumns | Gets the number of columns. |
| nbrOfRows | Gets the number of data rows. |
| readColumns | - |
| readDataFrame | Reads a subset of the columns as a data frame. |
| setColumnNameTranslator | - |

Methods inherited from *GenericDataFile*:

compareChecksum, copyTo, equals, fromFile, getAlias, getAttribute, getAttributes, getChecksum, getCreatedOn, getDefaultFullName, getExtension, getExtensionPattern, getFilename, getFilenameExtension, getFileSize, getFileType, getLastAccessedOn, getLastModifiedOn, getOutputExtension, getPath, getPathname, gunzip, gzip, hasBeenModified, isFile, readChecksum, renameTo, renameToUpperCaseExt, setAlias, setAttribute, setAttributes, setAttributesBy, setAttributesByTags, setExtensionPattern, testAttributes, validateChecksum, writeChecksum

Methods inherited from *FullNameInterface*:

appendFullNameTranslator, appendFullNameTranslatorBycharacter, appendFullNameTranslatorBydata.frame, appendFullNameTranslatorByfunction, appendFullNameTranslatorBylist, appendFullNameTranslatorByNULL, appendFullNameTranslatorByTabularTextFile, appendFullNameTrans-

latorByTabularTextFileSet, clearFullNameTranslator, clearListOfFullNameTranslators, getDefaultFullName, getFullName, getFullNameTranslator, getListOfFullNameTranslators, getName, getTags, hasTag, hasTags, setFullName, setFullNameTranslator, setListOfFullNameTranslators, setName, setTags, updateFullName

Methods inherited from Object:

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

Author(s)

Henrik Bengtsson (<http://www.braju.com/R/>)

See Also

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

GenericTabularFileSet *The GenericTabularFileSet class*

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:

append, appendFiles, appendFullNamesTranslator, appendFullNamesTranslatorBydata.frame, appendFullNamesTranslatorByfunction, appendFullNamesTranslatorBylist, appendFullNamesTranslatorByNULL, appendFullNamesTranslatorByTabularTextFile, appendFullNamesTranslatorByTabularTextFileSet, as.list, byName, byPath, clearCache, clearFullNamesTranslator, copyTo, equals, extract, findByName, getAlias, getChecksum, getChecksumObjects, getDefaultFullName, getFile, getFileClass, getFileSize, getFullNames, getNames, getPath, getPathnames, getSubdirs, hasFile, indexOf, lapply, nbrOfFiles, sapply, seq, setAlias, setFullNamesTranslator, sortBy, update2, updateFullName, updateFullNames, validate

Methods inherited from FullNameInterface:

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

Methods inherited from Object:

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

Author(s)

Henrik Bengtsson (<http://www.braju.com/R/>)

TabularTextFile

The TabularTextFile class

Description

Package: R.filesets

Class TabularTextFile

```

Object
~~|
~~+---FullNameInterface
~~~~~|
~~~~~+---GenericDataFile
~~~~~|
~~~~~+---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=0, columnNames=TRUE, commentChar=

```

Arguments

| | |
|------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ... | Arguments passed to <code>GenericTabularFile</code> . |
| sep | A <code>character</code> 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 <code>character</code> specifying the quote symbol used, if any. |
| fill | As in <code>read.table</code> . |
| skip | As in <code>read.table</code> . |
| columnNames | A <code>logical</code> or a <code>character vector</code> . If <code>TRUE</code> , then column names are inferred from the file. If a <code>character vector</code> , then the column names are given by this argument. |
| commentChar | A single <code>character</code> specifying which symbol should be used for comments, cf. <code>read.table</code> . |
| .verify, verbose | (Internal only) If <code>TRUE</code> , 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:

```

[
-
```

| | |
|----------------------------------|----------------------------------------------------------------------|
| <code>as.character</code> | - |
| <code>getColumnNames</code> | Gets the column names. |
| <code>getCommentChar</code> | - |
| <code>getHeader</code> | Gets the file header. |
| <code>hasColumnHeader</code> | Checks if there are column names in the header. |
| <code>nrOfLines</code> | Counts the number of lines in the data file. |
| <code>nrOfRows</code> | Counts the number of data rows. |
| <code>readColumnNames</code> | - |
| <code>readDataFrame</code> | Reads the tabular data as a data frame. |
| <code>readLines</code> | Reads the lines of the data file as strings. |
| <code>setCommentChar</code> | - |
| <code>writeColumnsToFiles</code> | Read each column from a data file and exports it to a separate file. |

Methods inherited from GenericTabularFile:

`as.character`, `dim`, `extractMatrix`, `getColumnNames`, `getColumnNameTranslator`, `nrOfColumns`, `nrOfRows`, `readColumns`, `readDataFrame`, `setColumnNameTranslator`, `translateColumnNames`

Methods inherited from GenericDataFile:

`compareChecksum`, `copyTo`, `equals`, `fromFile`, `getAlias`, `getAttribute`, `getAttributes`, `getChecksum`, `getCreatedOn`, `getDefaultFullName`, `getExtension`, `getExtensionPattern`, `getFilename`, `getFilenameExtension`, `getFileSize`, `getFileType`, `getLastAccessedOn`, `getLastModifiedOn`, `getOutputExtension`, `getPath`, `getPathname`, `gunzip`, `gzip`, `hasBeenModified`, `isFile`, `readChecksum`, `renameTo`, `renameToUpperCaseExt`, `setAlias`, `setAttribute`, `setAttributes`, `setAttributesBy`, `setAttributesByTags`, `setExtensionPattern`, `testAttributes`, `validateChecksum`, `writeChecksum`

Methods inherited from FullNameInterface:

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

Methods inherited from Object:

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

Author(s)

Henrik Bengtsson (<http://www.braju.com/R/>)

See Also

An object of this class is typically part of an `TabularTextFileSet`.

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:

No methods defined.

Methods inherited from [GenericTabularFileSet](#):

extractMatrix

Methods inherited from [GenericDataFileSet](#):

append, appendFiles, appendFullNamesTranslator, appendFullNamesTranslatorBydata.frame, appendFullNamesTranslatorByfunction, appendFullNamesTranslatorBylist, appendFullNamesTranslatorByNULL, appendFullNamesTranslatorByTabularTextFile, appendFullNamesTranslatorByTabularTextFileSet, as.list, byName, byPath, clearCache, clearFullNamesTranslator, copyTo, equals,

extract, findByName, getAlias, getChecksum, getChecksumObjects, getDefaultFullName, getFile, getFileClass, getFileSize, getFullNames, getNames, getPath, getPathnames, getSubdirs, hasFile, indexOf, lapply, nbrOfFiles, sapply, seq, setAlias, setFullNamesTranslator, sortBy, update2, updateFullName, updateFullNames, validate

Methods inherited from FullNameInterface:

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

Methods inherited from Object:

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

Author(s)

Henrik Bengtsson (<http://www.braju.com/R/>)

Examples

```
# Setup a file set
path <- system.file("exData", package="R.filesets")

ds <- TabularTextFileSet$byName("dataSetA,original", pattern=".[.]dat$", paths=path)
print(ds)

dfList <- lapply(ds, readDataframe)
print(dfList)

rows <- c(3:5, 8, 2);
dfList <- lapply(ds, readDataframe, rows=rows)
print(dfList)

data <- extractMatrix(ds, column="y", colClass="integer", rows=rows)
print(data)

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

verbose && cat(verbose, "Default fullnames:");
verbose && print(verbose, head(getFullNames(ds, translate=FALSE)));
verbose && cat(verbose, "Translated fullnames:");
verbose && print(verbose, head(getFullNames(ds)));

cat("Default fullnames:\n")
```

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

Index

*Topic **classes**

- FullNameInterface, 3
- GenericDataFile, 5
- GenericDataFileSet, 7
- GenericDataFileSetList, 10
- GenericTabularFile, 12
- GenericTabularFileSet, 14
- TabularTextFile, 15
- TabularTextFileSet, 18

*Topic **package**

- R.filesets-package, 2

*fromFile, 6

append, 8
appendFiles, 8
as.list, 8

byName, 8
byPath, 8

character, 6, 8, 11, 16
compareChecksum, 6

dim, 13

equals, 6
extract, 8
extractMatrix, 13

FALSE, 11
fromFile, 6
FullNameInterface, 3, 6, 8, 10–12, 14, 16, 18

GenericDataFile, 3, 5, 8, 12, 13, 16
GenericDataFileSet, 2, 3, 7, 7, 11, 14, 15, 18
GenericDataFileSetList, 3, 10
GenericTabularFile, 3, 6, 12, 14, 16
GenericTabularFileSet, 3, 8, 14, 14, 18
getAlias, 8
getChecksum, 6, 8
getColumnNames, 13, 17

getDefaultFullName, 8
getExtension, 6
getFile, 8
getFilename, 6
getFileSize, 6
getFileType, 6
getFullName, 4
getHeader, 17
getName, 4
getNames, 9
getPath, 6, 9
getPathname, 6
getPathnames, 9
getTags, 4

hasColumnHeader, 17
hasTags, 4

indexOf, 9
integer, 8
Interface, 3
isFile, 6

lapply, 9
list, 8, 11
logical, 16

nbrOfColumns, 13
nbrOfLines, 17
nbrOfRows, 13, 17

Object, 6, 7, 10, 12, 14, 16, 18

R.filesets (R.filesets-package), 2
R.filesets-package, 2
read.table, 16
readDataFrame, 13, 17
readLines, 17
seq, 9
setAlias, 9

setFullName, 4
setName, 4
setTags, 4
sortBy, 9

TabularTextFile, 2, 3, 6, 13, 15, 18

TabularTextFileSet, 3, 8, 14, 17, 18

TRUE, 6, 13, 16

validateChecksum, 6

vector, 8, 11, 16

writeChecksum, 7

writeColumnsToFiles, 17