Package ‘rpx’

April 15, 2017

Type Package

Title R Interface to the ProteomeXchange Repository

Version 1.10.2

Author Laurent Gatto

Maintainer Laurent Gatto <`lg390@cam.ac.uk`>

Description This package implements an interface to proteomics
data submitted to the ProteomeXchange consortium.

Depends methods

Imports XML, RCurl, utils

Suggests MSnbase, Biostrings, BiocStyle, testthat, knitr

License GPL-2

URL `https://github.com/lgatto/rpx`

BugReports `https://github.com/lgatto/rpx/issues`

VignetteBuilder knitr

biocViews Proteomics, MassSpectrometry, DataImport, ThirdPartyClient

RoxygenNote 5.0.1

NeedsCompilation no

R topics documented:

`pxannounced` *Return recent PX announcements*

Description

Queries the PX rss feed file for the latest PX dataset announcements.

Usage

```
pxannounced()
```

Value

A `data.frame` with announcements data set identifiers, publication dates and annoucement messages.

Author(s)

Laurent Gatto

Examples

```
pxannounced()
```

`PXDataset-class` *Class* `"PXDataset"`

Description

An S4 class to store and access information about ProteomeXchange (PX) data sets. Schema versions 1.0, 1.1 and 1.2 are supported and are documented on the PX code repository: https://code.google.com/p/proteomex

Objects from the Class

Objects can be created with the constructor `PXDataset`.

Slots

`id`: Object of class `"character"` storing the datasets unique identifier. Can be accessed with `pxid()`.

`formatVersion`: Object of class `"character"` storing the version of the ProteomeXchange schema.

`Data`: Object of class `"XMLNode"` storing the ProteomeXchange description as XML node tree.

Methods

- **pxfiles** `signature(object = "PXDataset")`: return a `character` of all available files.
- **pxget** `signature(object = "PXDataset", list, force = FALSE, ...)`: downloads the files from the ProteomeXchange repository. If `list` is missing, the file to be downloaded can be selected from a menu. If `list = "all"`, all files are downloaded. The file names, as returned by `pxfiles` can also be used. Alternatively, a `logical` or `numeric` indices can be used. All files will be downloaded in the working directory. Unless `force` is set to `TRUE`, files are not downloaded if already present in the working directory. Additional parameters can be passed to `download.file` via `...`. Invisibly returns the names of the downloaded files.
- **pxid** `signature(object = "PXDataset")`: returns the unique ProteomeXchange identifier.
- **pxref** `signature(object = "PXDataset")`: returns the reference associcated with the `object`.
- **pxtax** `signature(object = "PXDataset")`: returns the scientic taxonomic name of `object`.
- **pxurl** `signature(object = "PXDataset")`: returns the base url on the ProteomeXchange server where `pxfiles(object)` reside.
- **show** `signature(object = "PXDataset")`: textual representation of `object`.

Author(s)

Laurent Gatto <lg390@cam.ac.uk>

References

Vizcaino J.A. et al. 'ProteomeXchange: globally co-ordinated proteomics data submission and dissemination', Nature Biotechnology 2014, 32, 223 – 226, doi:10.1038/nbt.2839.

Source repository for the ProteomeXchange project: https://code.google.com/p/proteomexchange/

Examples

```
px <- PXDataset("PXD000001")
px
pxtax(px)
pxurl(px)
pxref(px)
pxfiles(px)
fnm <- pxget(px, "PXD000001_mztab.txt")
library("MSnbase")
readMzTabData(fnm, "PEP")
unlink("PXD000001_mztab.txt")
```

`pxnodes` *Return the nodes of a* `PXDataset`

Description

Returns the node names of the underliyng XML content of an `PXDataset` object, available in the `Data` slot. This function is meant to be used if additional parsing of the XML structure is needed.

Usage

```
pxnodes(pxdata, name, all = FALSE)
```

Arguments

`pxdata`	An instance of class `PXDataset`.
`name`	The name of a node.
`all`	Should node from all levels be returned. Default is `FALSE`.

Value

A `character` with XML node names.

Author(s)

Laurent Gatto

Index