

dcBuildAnno

December 22, 2014

dcBuildAnno

Function to build an object of the S4 class Anno from input files

Description

dcBuildAnno is supposed to build an object of the S4 class [Anno](#), given input files. These input files include 1) a file containing domain information, 2) a file containing term information, and 3) a file containing associations between domains and terms.

Usage

```
dcBuildAnno(domain_info.file, term_info.file, association.file,  
output.file = "Anno.RData")
```

Arguments

domain_info.file

an input file containing domain information. For example, a file containing InterPro domains (InterPro) can be found in <http://dcbioinformatics.org/data/InterPro/InterPro.txt>. As seen in this example, the input file must contain the header (in the first row), and entries in the first column intend to be domain ID (and must be unique). Note: the file should use the tab delimiter as the field separator between columns

term_info.file

an input file containing term information. For example, a file containing Gene Ontology (GO) terms can be found in <http://dcbioinformatics.org/data/InterPro/GO.txt>. As seen in this example, the input file must contain the header (in the first row) and four columns: 1st column for term ID (must be unique), 2nd column for term name, 3rd column for term namespace, and 4th column for term distance. These four columns must be provided, but the content for the last column can be arbitrary (if it is hard to prepare). Note: the file should use the tab delimiter as the field separator between columns

association.file

an input file containing associations between domains and terms. For example, a file containing associations between InterPro domains and GO Molecular Function (GOMF) terms can be found in <http://dcbioinformatics.org/data/InterPro/Domain2GOMF.txt>. As seen in this example, the input file must contain the header (in the first row) and two columns: 1st column for domain ID (corresponding to the first column in 'domain_info.file'), 2nd column

for term ID (corresponding to the first column in 'term_info.file'). If there are additional columns, these columns will be ignored. Note: the file should use the tab delimiter as the field separator between columns

`output.file` an output file used to save the built object as an RData-formatted file. If NULL, this file will be saved into "Anno.RData" in the current working local directory

Value

Any use-specified variable that is given on the right side of the assignment sign '`<-`', which contains the built Anno object. Also, an RData file specified in "output.file" is saved in the local directory.

Note

If there are no use-specified variable that is given on the right side of the assignment sign '`<-`', then no object will be loaded onto the working environment.

See Also

[Anno](#)

Examples

```
# build an "Anno" object that contains SCOP domain superfamilies (sf) annotated by GOBP terms
InterPro2GOMF <-
dcBuildAnno(domain_info.file="http://dgc.org/r-forge.r-project.org/data/InterPro/InterPro.txt",
term_info.file="http://dgc.org/r-forge.r-project.org/data/InterPro/GO.txt",
association.file="http://dgc.org/r-forge.r-project.org/data/InterPro/Domain2GOMF.txt",
output.file="InterPro2GOMF.RData")
InterPro2GOMF
```