

Package ‘provSummarizeR’

September 10, 2019

Title Summarizes Provenance Related to Inputs and Outputs of a Script
or Console Commands

Version 1.2

Date 2019-09-09

Copyright President and Fellows of Harvard College, Trustees of Mount
Holyoke College

Description Reads the provenance collected by the 'rdt' or
'rdtLite' packages, or other tools providing compatible PROV JSON output
created by the execution of a script, and provides a human-readable
summary identifying the input and output files, the script used
(if any), errors and warnings produced, and the environment in
which it was executed. It can also optionally package all the
files into a zip file. The exact format of the JSON created by
'rdt' and 'rdtLite' is described in
<<https://github.com/End-to-end-provenance/ExtendedProvJson>>.
More information about 'rdtLite' and associated tools is available
at <<https://github.com/End-to-end-provenance/>> and Barbara Lerner,
Emery Boose, and Luis Perez (2018), Using Introspection to Collect
Provenance in R, Informatics, <doi: 10.3390/informatics5010012>.

Depends R (>= 3.5)

License GPL-3 | file LICENSE

Encoding UTF-8

LazyData true

Imports dplyr, provParseR (>= 0.2)

Suggests rdtLite, rdt, testthat

Additional_repositories <https://end-to-end-provenance.github.io/drat/>

URL <https://github.com/End-to-end-provenance>

BugReports <https://github.com/End-to-end-provenance/provSummarizeR/issues>

RoxygenNote 6.1.1

NeedsCompilation no

Author Barbara Lerner [cre],
Emery Boose [aut]

Maintainer Barbara Lerner <blerner@ntholyoke.edu>

Repository CRAN

Date/Publication 2019-09-09 23:30:12 UTC

R topics documented:

prov.summarize	2
Index	4

prov.summarize	<i>Provenance summarization functions</i>
----------------	---

Description

prov.summarize uses the provenance from the last execution of prov.run and outputs a text summary to the R console

prov.summarize.file reads a JSON file that contains provenance and outputs a text summary to the R console

prov.summarize.run executes a script, collects provenance, and outputs a text summary to the console.

Usage

```
prov.summarize(save = FALSE, create.zip = FALSE)
```

```
prov.summarize.file prov.file, save = FALSE, create.zip = FALSE)
```

```
prov.summarize.run(r.script, save = FALSE, create.zip = FALSE, ...)
```

Arguments

save	if true saves the summary to the file prov-summary.txt in the provenance directory
create.zip	if true all of the provenance data will be packaged up into a zip file stored in the current working directory.
prov.file	the path to the file containing provenance
r.script	the name of a file containing an R script
...	extra parameters are passed to the provenance collector. See rdt's prov.run function or rdtLite's prov.run function for details.

Details

These functions use provenance collected using the rdtLite or rdt packages.

For provenance collected from executing a script file, the summary identifies:

- The name of the script file executed
- Environmental information identifying when the script was executed, the version of R, the computing system, the tool and version used to collect the provenance, the location of the provenance file, and the hash algorithm used to hash data files.
- A list of libraries loaded and their versions
- The names of any scripts sourced
- The names of files input or output, the file timestamp, and its hashvalue
- Any URLs loaded and the time loaded
- Any errors or warnings along with the line on which they occurred, if known.
- Any messages sent to standard output along with the line on which they occurred, if known.

For provenance collected from a console session, only the environment and library information appears in the summary.

Creating a zip file depends on a zip executable being on the search path. By default, it looks for a program named zip. To use a program with a different name, set the value of the R_ZIPCMD environment variable. This code has been tested with Unix zip and with 7-zip on Windows.

Examples

```
## Not run: prov.summarize ()
testdata <- system.file("testdata", "prov.json", package = "provSummarizeR")
prov.summarize.file (testdata)
## Not run:
testdata <- system.file("testscripts", "console.R", package = "provSummarizeR")
prov.summarize.run (testdata)
## End(Not run)
```

Index

prov.summarize, [2](#)