

Package: infoveillance (via r-universe)

October 15, 2026

Version 0.0.1

Title 'infoveillance' Package

Date 2026-09-13

Description Utilities to collect web-based data to track public health trends, behaviors, and disease outbreaks in real time.

License BSD_3_clause + file LICENSE

URL <https://www.rtemis.org>,
<https://github.com/rtemis-org/infoveillance>

BugReports <https://github.com/rtemis-org/infoveillance/issues>

ByteCompile yes

Depends R (>= 4.1.0)

Imports curl, data.table, jsonlite, methods, rtemis.core (>= 0.4.6),
S7, stats, utils

Suggests testthat (>= 3.0.0)

Additional_repositories <https://rtemis-org.r-universe.dev>

Encoding UTF-8

Roxygen list(markdown = TRUE)

Config/testthat/edition 3

Config/testthat/parallel true

Config/roxygen2/version 8.1.0

Config/pak/sysreqs libssl-dev

Repository <https://rtemis-org.r-universe.dev>

Date/Publication 2026-09-14 02:09:34 UTC

RemoteUrl <https://github.com/rtemis-org/infoveillance>

RemoteRef HEAD

RemoteSha 5d87a6121a30acb8317a158c5d6a731c5328ab4e

Contents

inforeillance-package	2
query_gdelt	2
Index	6

inforeillance-package	inforeillance: <i>'inforeillance' Package</i>
-----------------------	--

Description

Collect web-based data to track public health trends, behaviors, and disease outbreaks in real time

Author(s)

Maintainer: E.D. Gennatas <gennatas@gmail.com> (**ORCID**) [copyright holder]

Authors:

- E.D. Gennatas <gennatas@gmail.com> (**ORCID**) [copyright holder]

See Also

Useful links:

- <https://www.rtemis.org>
- <https://github.com/rtemis-org/inforeillance>
- Report bugs at <https://github.com/rtemis-org/inforeillance/issues>

query_gdelt	<i>Query GDELT for news articles matching keywords</i>
-------------	--

Description

Search the GDELT Global Knowledge Graph’s DOC 2.0 API for online news coverage matching one or more keywords, and return either the matching articles or an aggregate timeline of coverage volume, tone, language, or source country. Coverage begins on 2017-01-01 and is updated every 15 minutes; a query with no explicit window searches the last three months.

Usage

```

query_gdelt(
  keywords,
  match = c("all", "any"),
  mode = c("artlist", "timelinevol", "timelinevolraw", "timelinetone", "timelinelang",
    "timelinesourcecountry", "tonechart"),
  timespan = NULL,
  start_datetime = NULL,
  end_datetime = NULL,
  source_lang = NULL,
  source_country = NULL,
  domain = NULL,
  theme = NULL,
  tone = NULL,
  max_records = 250L,
  sort = c("datedesc", "dateasc", "tonedesc", "toneasc", "hybridrel"),
  timeline_smooth = NULL,
  timeout = 60,
  retries = 5L,
  verbosity = 1L
)

```

Arguments

keywords	Character: One or more keywords or phrases to search for.
match	Character: "all" requires every keyword; "any" requires at least one.
mode	Character: What to return. "artlist" returns matching articles; "timelinevol" the share of all monitored coverage matching the query over time; "timelinevolraw" the raw article counts together with the total monitored (norm); "timelinetone" average tone over time; "timelinelang" and "timelinesourcecountry" volume over time split by language or source country; "tonechart" a histogram of article tone.
timespan	Character or NULL: Search window ending now, e.g. "1w". Not used together with start_datetime/end_datetime.
start_datetime	Date, POSIXct, or character or NULL: Start of the search window, in UTC. A character value must be "YYYYMMDDHHMMSS".
end_datetime	Date, POSIXct, or character or NULL: End of the search window, in UTC. A Date end covers the whole day.
source_lang	Character or NULL: Language(s) of the source article, as a name ("english") or ISO 639 code ("eng").
source_country	Character or NULL: Country/ies the source outlet is based in, as a name ("unitedstates") or FIPS code ("US").
domain	Character or NULL: Source domain(s), e.g. "nytimes.com".
theme	Character or NULL: GDELT Global Knowledge Graph theme(s), e.g. "TAX_DISEASE_CORONAVIRUS".
tone	Character or NULL: Tone filter, a comparison such as ">5" (positive) or "<-5" (negative).

max_records	Integer: Maximum number of articles to return in "artlist" mode, at most 250L. Ignored by other modes.
sort	Character: Order of articles in "artlist" mode.
timeline_smooth	Integer or NULL: Moving-average window (in time steps, up to 30L) applied to timeline modes.
timeout	Numeric: Seconds to wait for a response.
retries	Integer: Maximum number of times to retry a request GDELT rejects as too frequent. The pause before each retry doubles from five seconds up to thirty.
verbosity	Integer: Verbosity level.

Details

Each keyword is one search term; a keyword containing whitespace is searched as an exact phrase. With `match = "all"` every keyword must appear in an article; with `match = "any"` at least one must. GDELT requires each term to be at least three characters and rejects terms that are too common to search on. Filters are combined with the keywords: an article must be in one of `source_lang`, from one of `source_country`, published on one of `domain`, tagged with one of `theme`, and have a tone within `tone`.

`timespan` and `start_datetime/end_datetime` are two ways to set the search window; supply one or the other. `timespan` is a number followed by a unit: `min` (minutes), `h` (hours), `d` (days), `w` (weeks), or `m` (months), e.g. `"24h"` or `"2w"`.

GDELT accepts at most one request every five seconds per client. `query_gdelt()` makes one request per call and waits as needed so that repeated calls in the same session respect that limit. GDELT also answers with the same "too many requests" rejection when its servers are busy, regardless of how requests are spaced, so a rejected request is retried up to `retries` times with an increasing pause between attempts.

Value

`data.table`. In "artlist" mode, one row per article with columns `url`, `url_mobile`, `title`, `seendate` (POSIXct, UTC), `socialimage`, `domain`, `language`, and `sourcecountry`. In timeline modes, one row per time step with columns `series`, `date` (POSIXct, UTC), `value`, and, for "timelineevolraw", `norm`. In "tonechart" mode, one row per tone bin with columns `bin`, `count`, and `toparts` (a list of `data.frames` of sample articles). A query with no matches returns an empty `data.table` with the same columns. The query string sent to GDELT and the full request URL are attached as attributes `"query"` and `"url"`.

Author(s)

EDG

Examples

```
## Not run:
# Articles from the last week mentioning both terms
query_gdelt(c("measles", "outbreak"), timespan = "1w")
```

```
# Daily coverage volume mentioning either phrase, English sources only
query_gdelt(
  c("bird flu", "avian influenza"),
  match = "any",
  mode = "timelinevol",
  start_datetime = as.Date("2025-01-01"),
  end_datetime = as.Date("2025-03-31"),
  source_lang = "english"
)

## End(Not run)
```

Index

`infoveillance (infoveillance-package)`, [2](#)

`infoveillance-package`, [2](#)

`query_gdelt`, [2](#)