Package: poddr (via r-universe)

November 1, 2026

Title Collect Metadata for Selected Podcasts

Version 0.3.3

Description Collecting all the data, but just for The Incomparable, Relay.fm and ATP.

License MIT + file LICENSE

URL https://github.com/jemus42/poddr, https://jemus42.github.io/poddr/

BugReports https://github.com/jemus42/poddr/issues

Depends R (>= 4.1.0)

Imports checkmate, cli, dplyr, fs, glue, here, hms, httr2, lubridate, purrr, readr, rlang (>= 0.1.2), robotstxt, rvest, stringr, tibble, tidyr, xml2

Suggests testthat (>= 3.0.0), vcr (>= 2.0.0), withr

Config/testthat/edition 3

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 8.0.0

Config/pak/sysreqs cmake make libicu-dev libuv1-dev libxml2-dev libssl-dev libx11-dev

Repository https://jemus42.r-universe.dev

Date/Publication 2026-09-02 11:58:18 UTC

RemoteUrl https://github.com/jemus42/poddr

RemoteRef HEAD

RemoteSha e7201719c767c2d704b82d01ec41a6ca45a0a3fb

Contents

atp_get_episodes 2
atp_parse_page 3
cache_podcast_data 3
gather_people 4

incomparable_get_episodes 5
incomparable_get_shows 5
incomparable_get_subcategories 6
incomparable_parse_archive 7
incomparable_parse_episode 7
incomparable_parse_stats 8
label_n 9
parse_duration 9
relay_get_episodes 10
relay_get_shows 10
relay_parse_feed 11
update_cached_data 12

Index **13**

`atp_get_episodes` *Retrieve ATP episodes*

Description

Retrieve ATP episodes

Usage

```
atp_get_episodes(page_limit = NULL, cache = TRUE)
```

Arguments

`page_limit`	Number of pages to scrape, from newest to oldest episode. Page 1 contains the 5 most recent episodes, and subsequent pages contain 50 episodes per page. Pass `NULL` (default) to get all pages.
`cache`	(`logical(1)`) Toggle the httr2 HTTP cache. Default `TRUE`. Disk writes are not performed by this function; call `cache_podcast_data()` explicitly if you want RDS/CSV artefacts.

Value

A tibble.

Examples

```
## Not run:
atp_new <- atp_get_episodes(page_limit = 1)
atp_full <- atp_get_episodes()

## End(Not run)
```

atp_parse_page *Parse a single ATP page*

Description

Parse a single ATP page

Usage

```
atp_parse_page(page)
```

Arguments

page	Scraped page object (xml_document).

Value

A tibble.

Examples

```
## Not run:
html <- poddr:::poddr_get("https://atp.fm", as = "html", query = list(page = 1))
atp_parse_page(html)

## End(Not run)
```

cache_podcast_data *Cache episode data to disk*

Description

Writes a tibble to RDS (and optionally CSV) in dir. Default dir is resolved with here::here() so the path is anchored to the project root rather than the current working directory.

Usage

```
cache_podcast_data(
  x,
  dir = here::here("data_cache"),
  filename = NULL,
  csv = TRUE
)
```

Arguments

x	Object to cache.
dir	Directory to save data to. Default: here::here("data_cache").
filename	Optional filename sans extension; defaults to deparse(substitute(x)).
csv	If TRUE (default), also saves a CSV file with the same base name.

Value

Invisibly returns the path(s) written, or NULL for empty input.

Examples

```
## Not run:
atp <- atp_get_episodes(page_limit = 1)
cache_podcast_data(atp, csv = FALSE)

## End(Not run)
```

---

gather_people *Gather episode datasets by people*

Description

A thin wrapper around tidyr::pivot_longer() and tidyr::separate_rows().

Usage

```
gather_people(episodes)
```

Arguments

episodes	A tibble containing host and guest columns, with names separated by ;.

Value

A tibble with new columns "role" and "person", one row per person.

Examples

```
## Not run:
incomparable <- incomparable_get_episodes(incomparable_get_shows())
incomparable_wide <- gather_people(incomparable)

## End(Not run)
```

incomparable_get_episodes

Retrieve all episodes for The Incomparable shows

Description

Retrieve all episodes for The Incomparable shows

Usage

```
incomparable_get_episodes(incomparable_shows, cache = TRUE)
```

Arguments

incomparable_shows	Dataset of shows as returned by `incomparable_get_shows()`.
cache	(`logical(1)`) Toggle the httr2 HTTP cache. Default `TRUE`.

Value

A tibble.

Examples

```
## Not run:
shows <- incomparable_get_shows()
incomparable_get_episodes(shows)

## End(Not run)
```

incomparable_get_shows

Retrieve all The Incomparable shows

Description

Retrieve all The Incomparable shows

Usage

```
incomparable_get_shows(cache = TRUE)
```

Arguments

cache	(`logical(1)`) Toggle the httr2 HTTP cache. Default `TRUE`.

Value

A tibble with columns show, stats_url, archive_url, status.

Examples

```
## Not run:
incomparable_get_shows()

## End(Not run)
```

`incomparable_get_subcategories`
Extract subcategory index for given show

Description

Extract subcategory index for given show

Usage

```
incomparable_get_subcategories(
  archive_url = "https://www.theincomparable.com/gameshow/archive/",
  cache = TRUE
)
```

Arguments

archive_url	E.g. "https://www.theincomparable.com/theincomparable/archive/".
cache	(logical(1)) Toggle the httr2 HTTP cache. Default TRUE.

Value

A tibble with subcategory links and category names.

Examples

```
## Not run:
incomparable_get_subcategories("https://www.theincomparable.com/gameshow/archive/")

## End(Not run)
```

incomparable_parse_archive

Parse a show's archive page on The Incomparable website

Description

Parse a show's archive page on The Incomparable website

Usage

```
incomparable_parse_archive(archive_url, cache = TRUE)
```

Arguments

archive_url	E.g. "https://www.theincomparable.com/theincomparable/archive/".
cache	(logical(1)) Toggle the httr2 HTTP cache. Default TRUE.

Value

A tibble.

Examples

```
## Not run:
incomparable_parse_archive("https://www.theincomparable.com/gameshow/archive/")

## End(Not run)
```

incomparable_parse_episode

Parse a single Incomparable episode page

Description

Recovers summary (and topic when present) for episodes that aren't on the archive page yet. The archive page is re-rendered on a slower cadence than stats.txt updates, so the newest episode of an active show is typically missing from the archive for hours to weeks. incomparable_get_episodes() calls this automatically for any episode in stats.txt that the archive doesn't list.

Usage

```
incomparable_parse_episode(episode_url, cache = TRUE)
```

Arguments

episode_url	The per-episode URL, e.g. "https://www.theincomparable.com/sophomorelit/190/".
cache	(logical(1)) Toggle the httr2 HTTP cache. Default TRUE.

Value

A one-row tibble with columns `summary` and `topic` (either may be `NA_character_` if the page doesn't expose them).

Examples

```
## Not run:
incomparable_parse_episode("https://www.theincomparable.com/sophomorelit/190/")

## End(Not run)
```

`incomparable_parse_stats`

Parse The Incomparable stats.txt files

Description

Parse The Incomparable stats.txt files

Usage

```
incomparable_parse_stats(stats_url, cache = TRUE)
```

Arguments

`stats_url`	URL to the `stats.txt`.
`cache`	(`logical(1)`) Toggle the httr2 HTTP cache. Default `TRUE`.

Value

A tibble.

Examples

```
## Not run:
incomparable_parse_stats("https://www.theincomparable.com/salvage/stats.txt")

## End(Not run)
```

label_n *Convenience function to display N*

Description

Convenience function to display N

Usage

```
label_n(x, brackets = FALSE)
```

Arguments

x	Data or singular value.
brackets	Set TRUE to enclose result in ( ).

Value

A character of length 1.

Examples

```
label_n(100)
label_n(tibble::tibble(x = 1:10, y = 1:10), brackets = TRUE)
```

parse_duration *Converting HH:MM:SS or MM:SS to* hms

Description

Converting HH:MM:SS or MM:SS to hms

Usage

```
parse_duration(x)
```

Arguments

x	A duration

Value

A numeric of durations in hms::hms().

Note

Only needed to parse durations in The Incomparable stats.txt files.

Examples

```
parse_duration("32:12")
parse_duration("32:12:04")
```

`relay_get_episodes` *Retrieve all episodes for relay.fm shows*

Description

Retrieve all episodes for relay.fm shows

Usage

```
relay_get_episodes(relay_shows, cache = TRUE)
```

Arguments

`relay_shows`	A tibble of shows, from `relay_get_shows()`.
`cache`	(`logical(1)`) Toggle the httr2 HTTP cache. Default `TRUE`.

Value

A tibble.

Examples

```
## Not run:
relay_shows <- relay_get_shows()
relay <- relay_get_episodes(relay_shows)

## End(Not run)
```

`relay_get_shows` *Retrieve all relay.fm shows*

Description

Parses the show overview page and returns a tibble of show names with corresponding feed URLs, which in turn can then be passed to `relay_parse_feed()` individually.

Usage

```
relay_get_shows(cache = TRUE)
```

Arguments

`cache`	(`logical(1)`) Toggle the httr2 HTTP cache. Default `TRUE`.

Value

A tibble with one row for each show.

Examples

```
## Not run:
relay_get_shows()

## End(Not run)
```

`relay_parse_feed` *Parse a relay.fm show feed*

Description

Parses a single feed and returns its content as a tibble.

Usage

```
relay_parse_feed(url, cache = TRUE)
```

Arguments

`url`	A show's feed URL, e.g. `"https://www.relay.fm/ungeniused/feed"`. Use `relay_get_shows()` to retrieve feed URLs.
`cache`	(`logical(1)`) Toggle the httr2 HTTP cache. Default `TRUE`.

Value

A tibble.

Examples

```
## Not run:
relay_parse_feed(url = "https://www.relay.fm/ungeniused/feed")

## End(Not run)
```

`update_cached_data` *Fetch all sources and cache results to disk*

Description

Convenience entry point used by the scheduled GitHub Action. Calls each fetch orchestrator and writes its output via `cache_podcast_data()`. Targets users typically don't want this — call the individual `*_get_episodes()` functions instead.

Usage

```
update_cached_data(dir = here::here("data_cache"))
```

Arguments

`dir`	Directory to save data to. Default: `here::here("data_cache")`.

Value

Invisibly returns the list of paths written.

Examples

```
## Not run:
update_cached_data()

## End(Not run)
```

Index

atp_get_episodes, 2
atp_parse_page, 3

cache_podcast_data, 3
cache_podcast_data(), *2*, *12*

gather_people, 4

here::here(), *3*

incomparable_get_episodes, 5
incomparable_get_shows, 5
incomparable_get_subcategories, 6
incomparable_parse_archive, 7
incomparable_parse_episode, 7
incomparable_parse_stats, 8

label_n, 9

parse_duration, 9

relay_get_episodes, 10
relay_get_shows, 10
relay_parse_feed, 11

update_cached_data, 12