Communication with official Hacker News API

Ryszard Szymański, Andrzej Nowikowski, Zuzanna Magierska,
Rafał Muszyński, Piotr Janus

November 17, 2019

Warsaw University of Technology

Hacker News

Hacker News — a social news website focusing on computer science. The Hacker News API gives access to information about:

- stories (latest, best, top, new),
- items,
- comments,
- profiles.

Purpose

The goal of this project is to give an opportunity to R users to process and analyze data from Hacker News website with a convenient to use tool.

Example use cases of the package
(https://github.com/szymanskir/hackeRnews#examples):

- identify buzzwords in job offers of Hacker News,
- check what's trending on Hacker News,
- sentiment analysis on two best stories from Hacker News.

The Hacker News API is constructed in such a way that a single item is retrieved with a single request. In order to improve performance the **hackeRnews** package uses the `future.apply` package and fetches all of the requested items in parallel.

hackeRnews CHEAT SHEET

An R wrapper for the official Hacker News API

Profile

get_updates()
Hacker News updated profiles
get_user_by_username(username**)**
Get Hacker News user

{ username
Username of the user,
that should be retrieved.

Latest stories

get_latest_ask_stories(max_items**)**
Hacker News latest ask stories
get_latest_ask_stories_ids()
Hacker News latest ask stories ids
get_latest_job_stories(max_items**)**
Hacker News latest job stories
get_latest_job_stories_ids()
Hacker News latest job stories ids
get_latest_show_stories(max_items**)**
Hacker News latest show stories
get_latest_show_stories_ids()
Hacker News latest show stories ids

Best stories

get_best_stories(max_items**)**
Hacker News best stories
get_best_stories_ids()
Hacker News best stories ids

Top stories

get_top_stories(max_items**)**
Hacker News top stories
get_top_stories_ids()
Hacker News top stories ids

New stories

get_new_stories(max_items**)**
Hacker News newest stories
get_new_stories_ids()
Hacker News best stories ids

Items

get_item_by_id(id**)**
Get Hacker News by Id

{ id
Id of the item.
that should be retrieved.

get_items_by_ids(ids**)**
Get Hacker News by ids

{ ids
List of ids

get_max_item_id()
Hacker News item largest id
get_comments(item**)**
Hacker News nested comments

{ item
Item whose children (comments)
will be retrieved

{ max_items
Maximum number of items to retrieve.
If max_items = null, returns all available.

Install the development version from GitHub with **devtools::install_github("szymanskir/hackeRnews")**

Plans for the future

Our plans are:

- submitting to CRAN,
- support for streaming live data,
- productize cheatsheet generator.

Thank you