
Source: <https://my.remarkbox.com/90cb90bc-94fa-11e8-bf30-040140774501>

Snapshot: 2026-05-03T08:31:12Z

Generator: Remarkbox 763cacb

This is a subthread snapshot. The living document lives at the source URI above — it may have been edited, extended, or replied-to since.



Scan for living
source

I like your article. I got your crawler working on Ubuntu 16.04 just fine and I'm inserting posts into mysql just fine. It runs every 5 minutes and grabs 3000 posts (about 2100 or so at the time of this posting are not dupes). Mysql holds the post_id so it knows if it has seen the post_id before. I'd like to grab comments too. Would be nice if your example has a bulk method of grabbing comments as well as your post example. Love that it's in PHP. I do a bunch of stuff in command-line PHP because it's quick and dirty and with PHP you don't need to npm or pip install anything - it's all just there! :)

CodePlea — Aug 01, 2018 09:55 am

The code for comments is exactly the same. You only need to change the t3_ part to t1_.

I agree with you on PHP. It's just so convenient to setup and maintain long-term.

Source: <https://my.remarkbox.com/90cb90bc-94fa-11e8-bf30-040140774501>

Snapshot: 2026-05-03T08:31:12Z

Generator: Remarkbox 763cacb