GitHub API search/code limits to 1000 items when using pagination

● GitHub Ecosystem   ● REST API

carldea   2 🖉   Feb 17

Dear GitHub Community,

When I perform a repo search like the following:
`GET /search/repositories`
`language:python stars:>=1000 &sort=stars&order=desc&page=10&per_page=100`

I am get something like the following:

```
{"total_count":2001, ...
```

But when I change the `page=11` I get the following:

```
{"message":"Only the first 1000 search results are available","documentation
```

Is there a way to overcome this limitation?

Thank you,
Carl

3 ♡   🔗

🔗 How to obtain all existing topics and relations info   1

created	last reply	16	1.6k	4	11	8		
Feb 17	23d	replies	views	users	likes	links		⌄

byrneh   Feb 17

Hi **@carldea**
The documentation and experience says "the GitHub Search API provides **up to 1,000 results for each search**" and you are unable page further.
Reference: about-the-search-api 47

1 ♡   🔗

carldea   1 🖉   Feb 17

I guess the answer is 'no' no way to overcome this limitation.
My mistake: The title of post is about **search/code** and the description is about the **search/repositories** API.

When I did some further digging around I found out that I could narrow results by using a date range.
The bigger question or ask is what about **search/code** API? There isn't a way to really filter those queries with date ranges. Any ideas?

♡   🔗

byrneh   Feb 18

Hi **@carldea** , as yo have already explored your only option seems to be to narrow the results using search qualifiers in any combination that can limit the result set to under 1000.
Search/code potentially being a little more challenging if you cannot limit query scope to an Organization/User.

There are other considerations-for-code-search 20 to be aware of where some things that will not be returned.
If you still need to search a result set greater than 1,000 you could try querying between greater and less than size value (not tested, so I don't know it works), then repeating for different size ranges, all of which would be to the 1,000 search limitation

1 ♡   🔗

Skip to main content

carldea   1 🖉   Feb 19

Any way to enhance this query?

When narrowing results by size in my opinion it is basically a guessing game. While I understand you don't want an API to be over taxes, I still feel there could've been a compromise. To make it flexible. GraphQL api doesn't seem to support search in code option.

Question: Is/ when would this api exist?
I'll still give it try, but I think the docs should have better examples.

Thank you,
Carl

1 ♡ ⧉

byrneh   Feb 19

I don't disagree, I am just a community user sharing the limitations I was aware related to your enquiry, and ensure you was aware of what you were and were not getting returned in the results as well as the 1,000 search results limitations.

Yon can submit a feedback request for any enhancements
You can make a contribution (correction or enhancement) to the Docs by following the Make a Contribution at the foot of each Docs page.

Good luck with your endeavours

1 Reply ⌄   ♡ ⧉

nosferican   1 ✎   Feb 19

For finding all repositories that meet some criteria you should limit the results based on when the repositories were created. That is a stable criterion supported. I suggest using the GraphQL version for making it a lot more feasible. What I have found to be quite useful is to use a preconditioner such as making the intervals large for earlier dates (e.g., 2018) and then smaller for more recent periods (e.g., 2017). You can do daily queries as well. Identify how many counts there are in each period and then merge those with fewer than 1,000 and split the ones that have higher than 1,000. If close enough, you can probably split those based on the count (e.g., if 3000 results, split the interval in four equal parts and most likely each will have fewer than 1,000). After you have all the intervals, collect the repositories. See this example Manual · GHOST 6

In terms of searching the code, you could search by user for instance or user/repo in case some users give more than 1,000 results.

1 Reply ⌄   2 ♡ ⧉

carldea   ⤵ 👤 byrneh   Feb 19

I thought this forum area is where feedback is communicated and monitored by the GitHub API team.

I find the documentation decent, however it would be nice if there was an OpenAPI (aka Swagger UI) allowing users to test REST APIs. GraphQL seems very nice, but again so much you have to know up front before getting started on your first query.

Maybe if the many of the "community users" ask enough they'll consider it being heard (or not).

According to StackOverflow it seems this has been asked in the GraphQL APIs and no one seems to want to respond.

2 Answers | Active | Oldest | **Votes** |

▲
21   `Type: CODE` is not supported yet. There is no way you can search the code using graphql right
▼   now.

        Your understanding is right. Just that you are missing one piece. The search you are doing is
✔   happening against the `type: REPOSITORY` . if you replace your search with

        `search(query: "beef", type: REPOSITORY, first: 10) {`

⟲   you will get all the repos having beef in their name.

                    GraphQL search/code not supported

        Share Improve this answer Follow   edited Oct 1 '17 at 9:57   answered Sep 20 '17 at 9:08
                                                                              👤 sreenivas
                                                                              1,374 ● 1 ● 10 ● 25

        11   two years later and `type: CODE` is still not supported, we still need to use the v3 API for this.. – Lloyd
              Sep 5 '19 at 11:43

        1   Make that three years later – GotDibbs Sep 12 '20 at 13:12

        Add a comment

Carl

♡ ⧉

byrneh Feb 19

> **carldea:**
>
> Swagger UI

@carldea There is mention in the REST API docs of an OpenAPI description - GitHub Docs 9 , in case that helps.

For GraphQL other than a single REST API like query I struggled at first and found myself having to the learn the basics of GraphQL to develop anything more complex. There is also the GraphQL Explorer they provide which helps running and testing some queries. I am currently using POSTMAN for REST and GRAPHQL development and testing when needed, but any similar tooling is also useful.

1 Reply ⌄   1 ♡   🔗

---

carldea Feb 19

> **nosferican:**
>
> rch by user for instance or user/re

@nosferican José,
I think I will try what you have suggested. Like you said, "is a stable criterion…" as far as I know it's (at least predictable). I've not tried the GraphQL API's search/repositories equivalent.
But yes, I like the idea of using user/repo to search the code.

Thank you,
Carl

                                                            ♡   🔗

---

carldea ↱ 👤 byrneh   Feb 19

@byrneh Hugh,

Swagger UI / Open API
I did see (Open API), but not sure if a swagger-ui is different. Such as the pet store example here:
https://petstore.swagger.io 3

GraphQL Explorer
We're in the same boat, with GraphQL, while very powerful and flexible it's just hard to visualize without some UML or domain models to understand relationships, datatypes, and meanings of attributes. It's basically lot's of trial and error. As for me it's more error than trial 😉

Thank you,
Carl

                                                            ♡   🔗

---

byrneh Feb 20

@carldea Carl,
been using GraphQL for a while and still revert to the good old trial and error a little, but it does get better (less bad) with use you will be please to know :).

With GitHub REST and GraphQL both has the use cases, and some things can only be done in one or the other. GraphQL is satisfying in the cases where you can do 1 GraphQL query vs. code and multiple calls with the equivalent REST options.

                                                  1 ♡   🔗

---

nosferican Feb 23

A recent example I shared was checking whether users were bots or not. I used the GraphQL endpoint to process 50,000 accounts per request. That's a pretty nice factor going from 3.3M requests using the REST endpoint to 66 using the GraphQL one.

                                                  1 ♡   🔗

---

carldea Feb 25

@nosferican , Do you have a sample (example) of your GraphQL search to share?

nosferican   Feb 25

Here is an example 26 . I just verified my code and I was actually doing 5,000 per request.

The first line is the query and the third line is the variables. You can copy those into the explorer 7 to see it in action. I construct those programmatically.

22 DAYS LATER

carldea   Mar 19

How did you construct and get those cursors, next. Can you explain each section?
The first line doesn't make sense to me. Does GraphQL understand TypeScript's spread operator?

6 MONTHS LATER

zjcnb   23d

Hello, this question are resolve?

↩ **Reply**

Suggested Topics

Topic	Replies	Views	Activity
How to compare two commits using GH Api? ◖ REST API	0	137	Feb 17
How can I get "Linked issues" information from REST API? ◖ REST API	7	323	Jun 24
☑ Github's API is not returning Status in it's response header ◖ REST API	1	180	Feb 9
Add or update team permission to org repo not working ◖ REST API	0	226	Jan 21
API Searching user repositories and filtering by organization ◖ REST API	0	135	Jun 22

Want to read more? Browse other topics in ◖ **REST API or view latest topics.**