Search more than 5,411,518 Repositories

Search

ProTip: For an advanced search, use some of our prefixes.
Search

Search more than 5,411,518 Repositories

ProTip: For an advanced search, use some of our prefixes.
Search January 2013

Search more than 5,411,518 Repositories

ProTip: For an advanced search, use some of our prefixes.
Planning
We knew search sucked
Start by talking about it
NOTE: This stuff is already underway, but I'd like to document some of my ideas on how all this stuff will come together.

Documents in Gist and GitHub are probably one of two types: prose or code. In order to get the most from our search engine, we should run documents through some kind of classifier to categorize them into one of these two kinds. Prose is pretty simple to index, code is not so easy.

When you search for code, you are generally searching for something specific, and so we need to keep more information than we would with an English document. We also need to do ranking in a different way.

Consider the following snippet of a document:

```
EventMachine::run
```

The punctuation here is relevant and we’d like to keep it around so that a perfect match search for this snippet will rank highly. Therefore we keep the colons during tokenization:

```
EventMachine :: run
```

We also want to match "Event" and "Machine" separately as well as the lower case versions of all these items. So in the end we get the following tokens:
If you have a spare moment, go to github.com/search.

These search pages need some lovin’ in a bad, bad way.
Two-person teams
Two-person teams

Frontend
Okay, let’s build something
git checkout -b search-next
Really simple branching
Cut a branch from master
(rarely from non-master branch)
Push whenever; let others help
No need for complicated forks
git commit -m "Design Search"
git commit -m ":lipstick:"
jonrohan authored 4 days ago
We don’t rebase or do any fancy merge strategies
Small, incremental commits
Open a Pull Request
Search Page New Design

I'm updating the design for the code search results.

[X] I couldn't tell private/public/fork/mirror in the repo results
[X] I might need a user object for the user results, to render the follow button on the user results
[X] For the code results my original idea included the owner's gravatar smaller to the side. (design
[ ] For the code results I need what line the code begins on, right now it's always 1 which is incor
[X] For the "Advanced Search" page I wanted to build something similar to [Google’s Advanced Search
[ ] Have the command bar in the search box also. There's a bug where we can't do this yet. #6434 If
[X] Pjax the search results tabs
[ ] Search prefixes help popup
[ ] Make it fast!!!
Sometimes we open pulls before code
Sometimes we open pulls after code
Sometimes pulls aren’t for merging
130 commits  2 months  10 contributors
Let your pulls evolve
This pull request ships the new search page to Gen POP.

- [x] blog post needs eyeballs https://github.com/blog/1381-a-whole-new-code-search
- [x] @ mention branch needs to merge
- [x] Update advanced search to include @ mention
- [x] elastic search upgrades by @wfarr github/puppet#1061

@kneath @rtomayko @TwP @dgraham
@kneath @rtomayko @TwP @dgraham
Low-overhead way to ask for help
Non-intrusive way to ask for help
Feel good about a ship by mentioning the relevant team.
Tests are kinda critical
Every commit we push triggers three builds:
ruby 1.8
ruby 1.9
github enterprise
Make it hard to ignore failures
Shipping
Limit your exposure
We have ~35 features staff-shipped
My GitHub looks different from yours:
Ranges from four-line config changes to entire feature redesigns
Coworkers are the only real beta testers
def search_enabled?
  current_user.staff?
end
def search_enabled?
  true
end
STAFF-SHIPPING

No risky deploys; it’s already production-tested
Search’s design was staff-shipped for two months.
Search's backend was staff-shipped for four months.
Gist 2.0 was staff-shipped for a year
More real-world usage means fewer surprises
Shipping is scary; reduce surprises
Deploying is scary; reduce surprises
Deploying to production involves 80 servers and hundreds of Unicorn processes and magic and it’s a wonder this stuff works
DEPLOYMENT

We want it easy enough for a designer to deploy.
Hubot talks to an API that knows how to deploy every app.
hubot deploy github/ship-search
Deploy a branch; rollback by deploying master
DEPLOYMENT

hubot deploy github/master
Chat’s important to us
Search didn’t need standup meetings
Chat is asynchronous and non-blocking
“Chatops”
Transparent way of handling ops
Search deployed, check perf
/graph me -1h substr(es-*__github.load)
CHATOPS
load on es-storage15 is higher than the other storage nodes
let's let it go a bit- could just be due to relocating shards causing iowait
Working over SSH is non-collaborative
Makes it obvious what needs automation and tooling
POST-SHIP

Keep shipping
I decided to blog on our blog about our new blog.

Design. The blog looks fresh no matter where you read it — be it on your desktop or on your phone.
Keep fixing
A User email@example.com 4 days ago

I want to find examples of people using the threaded_reactor function from the nose library (a testing tool for Python).

Searching for "threaded_reactor" gives about 100 hits, but almost all the hits return copies of the same library.

Would it be possible to optionally display just the filenames in the search results?

Many thanks for all the great services you provide!

System changed inbox to Technical 4 days ago

steveward added label search 4 days ago

Related Discussions
No related discussions found.
Don’t assume someone else will fix ‘em
Thanks.