

Track geospatial & non-geospatial data edits with GitHub.

- 1 Two basic commands for working with APIs
- 2 What GitHub Actions are
- 3 How to setup »scrapers« for GeoNode
- 4 GeoNode Api endpoints

1 Commands you need to know!

cURL

Is a http Client for your command line

<https://curl.se>

»curl is used in command lines or scripts to transfer data. It is also used in cars, television sets, routers, printers, audio equipment, mobile phones, tablets, settop boxes, media players and is the internet transfer backbone for thousands of software applications affecting *billions of humans* daily.«



-zsh

UTF-8

tonischoenbuchner@Tonis-MBP ~ % curl https://master.demo.geonode.org/api/

```
{"base": {"list_endpoint": "/api/base/", "schema": "/api/base/schema/"}, "categories": {"list_endpoint": "/api/categories/", "schema": "/api/categories/schema/"}, "documents": {"list_endpoint": "/api/documents/", "schema": "/api/documents/schema/"}, "featured": {"list_endpoint": "/api/featured/", "schema": "/api/featured/schema/"}, "group_profile": {"list_endpoint": "/api/group_profile/", "schema": "/api/group_profile/schema/"}, "groupcategory": {"list_endpoint": "/api/groupcategory/", "schema": "/api/groupcategory/schema/"}, "groups": {"list_endpoint": "/api/groups/", "schema": "/api/groups/schema/"}, "keywords": {"list_endpoint": "/api/keywords/", "schema": "/api/keywords/schema/"}, "layers": {"list_endpoint": "/api/layers/", "schema": "/api/layers/schema/"}, "maps": {"list_endpoint": "/api/maps/", "schema": "/api/maps/schema/"}, "owners": {"list_endpoint": "/api/owners/", "schema": "/api/owners/schema/"}, "profiles": {"list_endpoint": "/api/profiles/", "schema": "/api/profiles/schema/"}, "regions": {"list_endpoint": "/api/regions/", "schema": "/api/regions/schema/"}, "styles": {"list_endpoint": "/api/styles/", "schema": "/api/styles/schema/"}, "thesaurus/keywords": {"list_endpoint": "/api/thesaurus/keywords/", "schema": "/api/thesaurus/keywords/schema/"}}
```

tonischoenbuchner@Tonis-MBP ~ %

1 Commands you need to know!

Jq

is a command-line JSON processor
<https://stedolan.github.io/jq/>

|

(pipe) can be used to *forward* output
to another command



-zsh

UTF-8

```
tonischoenbuchner@Tonis-MBP ~ % curl https://master.demo.geonode.org/api/ | jq .
```

% Total	% Received	% Xferd	Average Speed	Time	Time	Time	Current
			Dload	Upload	Total	Spent	Speed
100	1278	100	1278	0	0	6908	0
					--:--:--	--:--:--	--:--:--
							6945

```
{
  "base": {
    "list_endpoint": "/api/base/",
    "schema": "/api/base/schema/"
  },
  "categories": {
    "list_endpoint": "/api/categories/",
    "schema": "/api/categories/schema/"
  },
  "documents": {
    "list_endpoint": "/api/documents/",
    "schema": "/api/documents/schema/"
  },
  "featured": {
```

1 The last Command you need to know!

```
Curl https://master.demo.geonode.org/api |  
jq . > somefile.json
```


2 Github Actions

... is a module from and for GitHub to automate things.

Actions · GeoNode/geonode

github.com/GeoNode/geonode/actions?query=workflow%3ABackport

Search or jump to...

Pull requestsIssuesMarketplaceExplore

1

+

GeoNode / geonode

Unwatch

116

Star

947

Fork

847

<> Code

! Issues25

🔗 Pull requests1

▶ Actions

📁 Projects1

📖 Wiki

🛡 Security

📈 Insights

⚙ Settings

Workflows

New workflow

All workflows

🔗 Backport

Backport

🔍 workflow:Backport

2,373 results

	Event	Status	Branch	Actor
✓ [Backport 3.x] Bump tqdm from 4.51.0 to 4.53.0 Backport #2384: Pull request #6677 closed by afabiani	backport-6662-to-3.x	2 days ago 17s	...	
✓ [Backport 3.x] Bump docker from 4.3.1 to 4.4.0 Backport #2383: Pull request #6676 closed by afabiani	backport-6664-to-3.x	2 days ago 18s	...	
✓ [Backport 3.x] Bump requests from 2.24.0 to 2.25.0 Backport #2382: Pull request #6674 closed by afabiani	backport-6650-to-3.x	2 days ago 17s	...	
✓ [Backport 3.x] Bump urllib3 from 1.25.11 to 1.26.2 Backport #2381: Pull request #6672 closed by afabiani	backport-6647-to-3.x	2 days ago 17s	...	
✓ [Backport 3.x] Bump tqdm from 4.51.0 to 4.53.0 Backport #2380: Pull request #6677 labeled by cla-bot bot	backport-6662-to-3.x	2 days ago 17s	...	

<https://docs.github.com/en/free-pro-team@latest/actions>

A Github Action can be used for:

- + **downloading** content from GeoNode
- + checking and **formatting** the output
- + and if the content has changed
automatically add it to our **git history**

This process is called **git scraping**.

<https://simonwillison.net/2020/Oct/9/git-scraping/> 🙌

Actions · t-book/geonode-scraper

github.com/t-book/geonode-scraper/actions/new

Choose a workflow template

Build, test, and deploy your code. Make code reviews, branch management, and issue triaging work the way you want. Select a workflow template to get started.

Skip this and [set up a workflow yourself](#) →

Workflows made for your repository Suggested

Simple workflow

By GitHub Actions

Start with a file with the minimum necessary structure.

Set up this workflow

echo Hello, world!
echo Add other actions to build,
echo test, and deploy your project.

actions/starter-workflows

Deploy your code with these popular services

Deploy Node.js to Azure Web App

By Microsoft Azure

Build a Node.js project and deploy it to an Azure Web App.

Deploy to Amazon ECS

By Amazon Web Services

Deploy a container to an Amazon ECS service powered by AWS Fargate or Amazon EC2.

Build and Deploy to GKE

By Google Cloud

Build a docker container, publish it to Google Container Registry, and deploy to GKE.

New File

github.com/t-book/geonode-scraper/new/main?filename=.github%2Fworkflows%2Fblank.yml&workflow_template=blank

Search or jump to...

Pull requests

Issues

Marketplace

Explore

t-book / geonode-scraper

Unwatch 1

Star 0

Fork 0

<> Code

Issues

Pull requests

Actions

Projects

Wiki

Security

Insights

Settings

geonode-scraper / github / workflows / blank.yml

Cancel

Start commit

<> Edit new file

Preview

Spaces 2 No wrap

```
1 # This is a basic workflow to help you get started with Actions
2
3 name: CI
4
5 # Controls when the action will run.
6 on:
7   # Triggers the workflow on push or pull request events but only for the main branch
8   push:
9     branches: [ main ]
10  pull_request:
11    branches: [ main ]
12
13 # Allows you to run this workflow manually from the Actions tab
14 workflow_dispatch:
15
16 # A workflow run is made up of one or more jobs that can run sequentially or in parallel
17 jobs:
18   # This workflow contains a single job called "build"
19   build:
20     # The type of runner that the job will run on
21     runs-on: ubuntu-latest
```

Use Control + Space to trigger autocomplete in most situations.

Marketplace

Documentation

Search Marketplace for Actions

Featured Actions

Cache

By actions

1.5k

Cache artifacts like dependencies and build outputs to improve workflow execution time

Close Stale Issues

By actions

292

Close issues and pull requests with no recent activity

Setup .NET Core SDK

By actions

232

Set up a specific version of the .NET Core CLI in the PATH and set up authentication to a private NuGet repository

name: GeoNode Profile scraper

on:

workflow_dispatch:

schedule:

- cron: '6,26,46 * * * *'

jobs:

scheduled:

runs-on: ubuntu-latest

steps:

- name: Check out this repo

uses: actions/checkout@v2

- name: Fetch latest data

run: |-

curl "https://master.demo.geonode.org/api/profiles/?

abstract__icontains=Toni&f_method=or&limit=5&offset=0&purpose__icontains=Toni&username__icontains=Toni" | jq 'del(.requested_time)' > profiles.json

- name: Commit and push if it changed

run: |-

git config user.name "Automated"

git config user.email "actions@users.noreply.github.com"

git pull

git add -A

timestamp=\$(date -u)

git commit -m "Latest data: \${timestamp}" || exit 0

git push

name: GeoNode Profile scraper

on:

workflow_dispatch:

schedule:

- cron: '6,26,46 * * * *'

jobs:

scheduled:

runs-on: ubuntu-latest

steps:

- name: Check out this repo

uses: actions/checkout@v2

- name: Fetch latest data

run: |-

curl "https://master.demo.geonode.org/api/profiles/?

abstract__icontains=Toni&f_method=or&limit=5&offset=0&purpose__icontains=Toni&username__icontains=Toni" | jq 'del(.requested_time)' > profiles.json

- name: Commit and push if it changed

run: |-

git config user.name "Automated"

git config user.email "actions@users.noreply.github.com"

git pull

git add -A

timestamp=\$(date -u)

git commit -m "Latest data: \${timestamp}" || exit 0

git push

name: GeoNode Profile scraper

on:

workflow_dispatch:

schedule:

- cron: '6,26,46 * * * *'

jobs:

scheduled:

runs-on: ubuntu-latest

steps:

- name: Check out this repo

uses: actions/checkout@v2

- name: Fetch latest data

run: |-

curl "https://master.demo.geonode.org/api/profiles/?

abstract__icontains=Toni&f_method=or&limit=5&offset=0&purpose__icontains=Toni&username__icontains=Toni" | jq 'del(.requested_time)' > profiles.json

- name: Commit and push if it changed

run: |-

git config user.name "Automated"

git config user.email "actions@users.noreply.github.com"

git pull

git add -A

timestamp=\$(date -u)

git commit -m "Latest data: \${timestamp}" || exit 0

git push

name: GeoNode Profile scraper

on:

workflow_dispatch:

schedule:

- cron: '6,26,46 * * * *'

jobs:

scheduled:

runs-on: ubuntu-latest

steps:

- name: Check out this repo

uses: actions/checkout@v2

- name: Fetch latest data

run: |-

curl "https://master.demo.geonode.org/api/profiles/?

abstract__icontains=Toni&f_method=or&limit=5&offset=0&purpose__icontains=Toni&username__icontains=Toni" | jq 'del(.requested_time)' > profiles.json

- name: Commit and push if it changed

run: |-

git config user.name "Automated"

git config user.email "actions@users.noreply.github.com"

git pull

git add -A

timestamp=\$(date -u)

git commit -m "Latest data: \${timestamp}" || exit 0

git push

name: GeoNode Profile scraper

on:

workflow_dispatch:

schedule:

- cron: '6,26,46 * * * *'

jobs:

scheduled:

runs-on: ubuntu-latest

steps:

- name: Check out this repo

uses: actions/checkout@v2

- name: Fetch latest data

run: |-

curl "https://master.demo.geonode.org/api/profiles/?

abstract__icontains=Toni&f_method=or&limit=5&offset=0&purpose__icontains=Toni&username__icontains=Toni" | jq 'del(.requested_time)' > profiles.json

- name: Commit and push if it changed

run: |-

git config user.name "Automated"

git config user.email "actions@users.noreply.github.com"

git pull

git add -A

timestamp=\$(date -u)

git commit -m "Latest data: \${timestamp}" || exit 0

git push

name: GeoNode Profile scraper

on:

workflow_dispatch:

schedule:

- cron: '6,26,46 * * * *'

jobs:

scheduled:

runs-on: ubuntu-latest

steps:

- name: Check out this repo

uses: actions/checkout@v2

- name: Fetch latest data

run: |-

curl "https://master.demo.geonode.org/api/profiles/?

abstract__icontains=Toni&f_method=or&limit=5&offset=0&purpose__icontains=Toni&username__icontains=Toni" | jq 'del(.requested_time)' > profiles.json

- name: Commit and push if it changed

run: |-

git config user.name "Automated"

git config user.email "actions@users.noreply.github.com"

git pull

git add -A

timestamp=\$(date -u)

git commit -m "Latest data: \${timestamp}" || exit 0

git push

t-book/geonode-scraper

github.com/t-book/geonode-scraper

Search or jump to...

Pull requestsIssuesMarketplaceExplore

Unwatch

1

Star0

Fork0

<> Code

Issues

Pull requests

Actions

Projects

Wiki

Security

Insights

Settings

main1 branch0 tags

Go to fileAdd fileCode

Automated

Latest data: Sun Nov 29 12:52:33 UTC 2020

67d47192 minutes ago309 commits

.github/workflows	Update layer-csv-scraper.yml	2 days ago
groups.json	Latest data: Sun Nov 29 12:26:37 UTC 2020	28 minutes ago
profiles.json	Latest data: Fri Nov 27 11:20:00 UTC 2020	2 days ago
swamp.csv	Latest data: Fri Nov 27 07:50:46 UTC 2020	2 days ago
swamp.geojson	Latest data: Sun Nov 29 12:52:33 UTC 2020	2 minutes ago

Help people interested in this repository understand your project by adding a README.

Add a README

About

No description, website, or topics provided.

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

© 2020 GitHub, Inc.

TermsPrivacyCookie PreferencesSecurityStatusHelpContact GitHubPricingAPITrainingBlogAbout

Latest data: Fri Nov 27 11:20:00 UTC 2020

main

Automated committed 2 days ago

1 parent [9a2f939](#) commit [d2ef9bb1751c321abe2c1cad7f2b04de1a9222c0](#)

Showing 1 changed file with 2 additions and 2 deletions.

UnifiedSplit

4 profiles.json

@@ -19,7 +19,7 @@

19"delivery": null,

20"documents_count": 0,

21"email": "",

22- "fax": null,

23"first_name": "Toni",

24"id": 1001,

25"language": "en",

@@ -33,7 +33,7 @@

33"resource_uri": "/api/profiles/1001/",

34"timezone": "Europe/Berlin",

35"username": "toni",

36- "voice": null,

37"zipcode": null

38},

39{

19"delivery": null,

20"documents_count": 0,

21"email": "",

22+ "fax": "000",

23"first_name": "Toni",

24"id": 1001,

25"language": "en",

33"resource_uri": "/api/profiles/1001/",

34"timezone": "Europe/Berlin",

35"username": "toni",

36+ "voice": "000",

37"zipcode": null

38},

39{

0 comments on commit [d2ef9bb](#)

Lock conversation

swamp — master.demo.geonode.org

master.demo.geonode.org/layers/geonode_master_data:geonode:swamp

GeoNode

Data

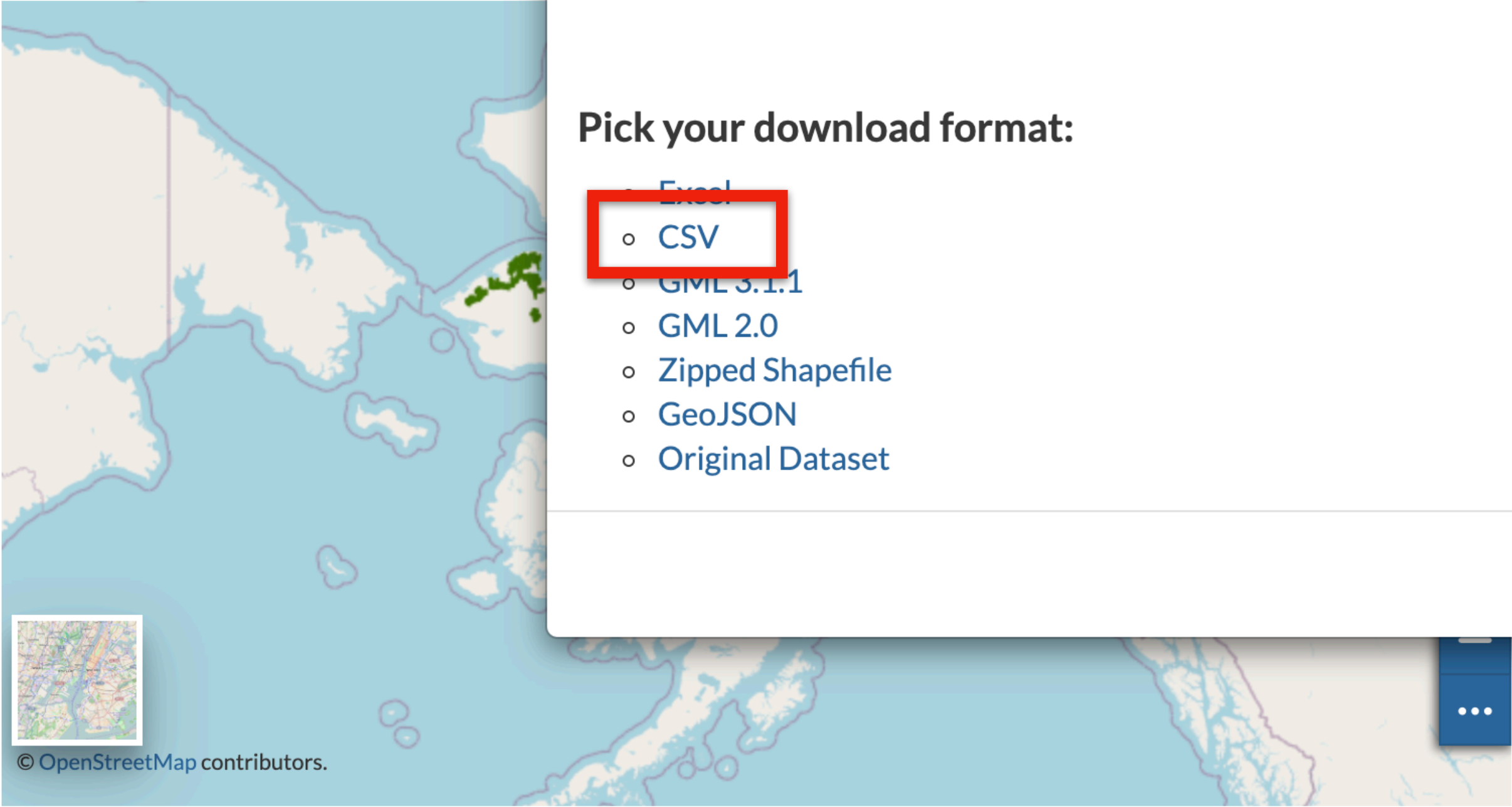
Maps

Search

Register

Sign in

swamp



© OpenStreetMap contributors.

Info

Attributes

Share

Ratings

Comments

Title swamp

Download Layer

Images

Data

Do you want to filter it?

Pick your download format:

Excel

CSV

GML 3.1.1

GML 2.0

Zipped Shapefile

GeoJSON

Original Dataset

Close

Download Layer

Metadata Detail

View Layer

Download Metadata

Maps using this layer

List of maps using this layer:

Water-Alaska

name: GeoNode Layer CSV scraper

on:

push:

workflow_dispatch:

schedule:

– cron: '10,30,50 * * * *'

jobs:

scheduled:

runs-on: ubuntu-latest

steps:

– name: Check out this repo

uses: actions/checkout@v2

– name: Fetch latest data

run: |-

curl "https://master.demo.geonode.org/geoserver/ows?

service=WFS&version=1.0.0&request=GetFeature&typename=geonode%3Aswamp&outputFormat=csv&srs=EPSG%3A2964" >

swamp.csv

– name: Commit and push if it changed

run: |-

git config user.name "Automated"

git config user.email „actions@users.noreply.github.com“

git pull

git add -A

timestamp=\$(date -u)

git commit -m "Latest data: \${timestamp}" || exit 0

git push

 Lock conversation

Now comes the fun part.
We're harvesting geospatial data
and visualize it with GitHub.

As to make the magic happen we need to
take care of some steps first.

swamp — master.demonode.org

master.demonode.org/layers/geonode_master_data:geonode:swamp

GeoNode

Data

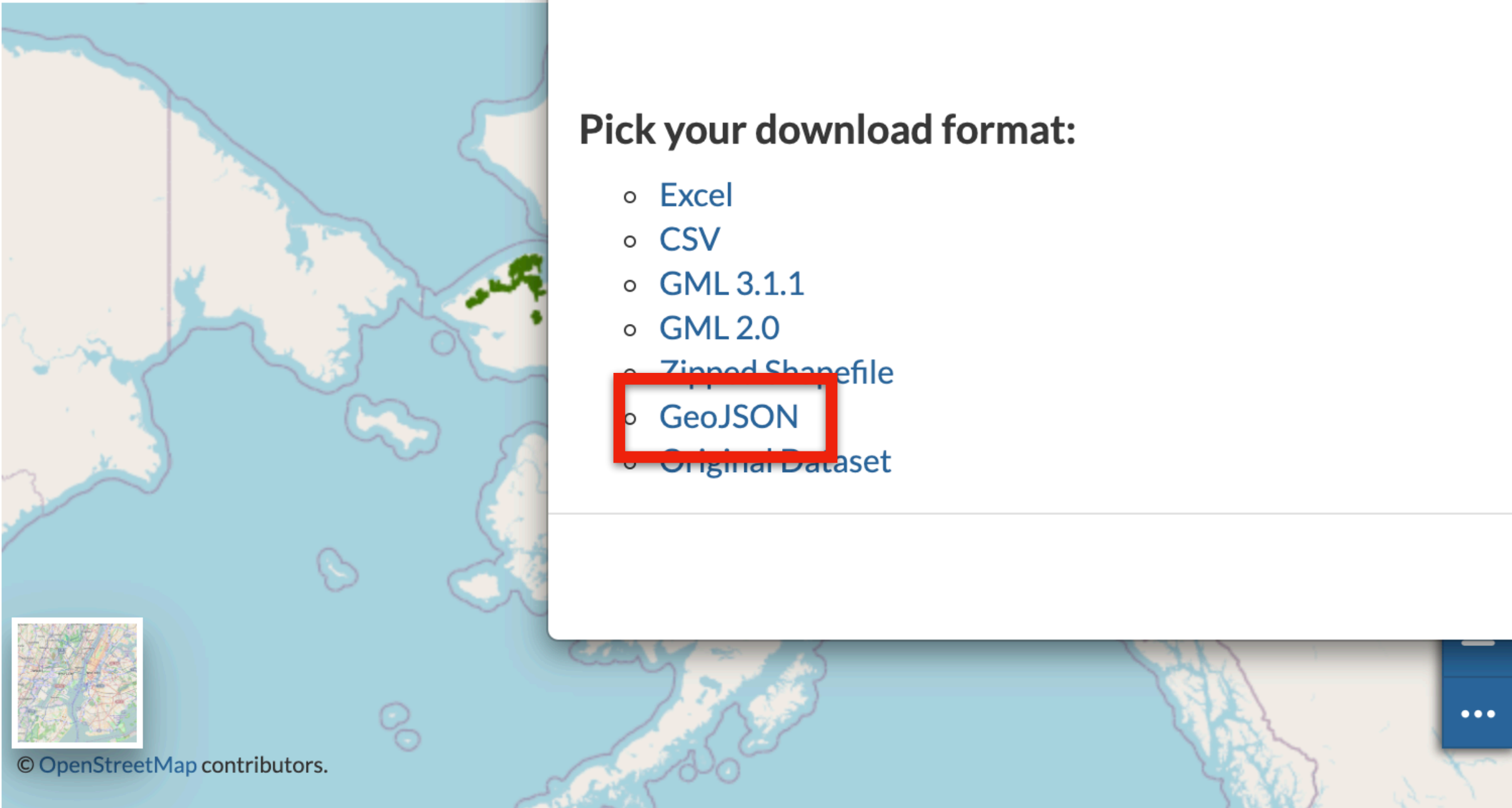
Maps

Search

Register

Sign in

swamp



© OpenStreetMap contributors.

Info

Attributes

Share

Ratings

Comments

Title

swamp

Download Layer

Images

Data

Do you want to filter it?

Pick your download format:

Excel

CSV

GML 3.1.1

GML 2.0

Zipped Shapefile

GeoJSON

Original Dataset

Close

Download Layer

Metadata Detail

View Layer

Download Metadata

Maps using this layer

List of maps using this layer:

Water-Alaska

name: GeoNode Layer scraper

on:

workflow_dispatch:

schedule:

- cron: '8,28,48 * * * *'

jobs:

scheduled:

runs-on: ubuntu-latest

steps:

- uses: actions/checkout@v2

- uses: actions/setup-node@v1

with:

node-version: '12.x'

- run: npm install -g geojson-rewind

- run: |-

curl "https://master.demo.geonode.org/geoserver/ows?

service=WFS&version=2.0.0&request=GetFeature&typename=geonode%3Aswamp&outputFormat=json&srs=EPSG%3A4326&srsName=EPSG%3A4326" | jq 'del(.requested_time,.crs,.timeStamp)' > swamp_tmp.geojson

- run: geojson-rewind swamp_tmp.geojson > swamp.geojson

- run: rm swamp_tmp.geojson

- run: |-

git config user.name "Automated"

git config user.email "actions@users.noreply.github.com"

git pull

git add -A

timestamp=\$(date -u)

git commit -m "Latest data: \${timestamp}" || exit 0

git push

name: GeoNode Layer scraper

on:

workflow_dispatch:

schedule:

- cron: '8,28,48 * * * *'

jobs:

scheduled:

runs-on: ubuntu-latest

steps:

- uses: actions/checkout@v2

- uses: actions/setup-node@v1

with:

node-version: '12.x'

- run: npm install -g geojson-rewind

- run: |-

curl "https://master.demo.geonode.org/geoserver/ows?

service=WFS&version=2.0.0&request=GetFeature&typename=geonode%3Aswamp&outputFormat=json&srs=EPSG%3A4326&srsName=EPSG%3A4326" | jq 'del(.requested_time,.crs,.timeStamp)' > swamp_tmp.geojson

- run: geojson-rewind swamp_tmp.geojson > swamp.geojson

- run: rm swamp_tmp.geojson

- run: |-

git config user.name "Automated"

git config user.email "actions@users.noreply.github.com"

git pull

git add -A

timestamp=\$(date -u)

git commit -m "Latest data: \${timestamp}" || exit 0

git push



name: GeoNode Layer scraper

on:

workflow_dispatch:

schedule:

- cron: '8,28,48 * * * *'

jobs:

scheduled:

runs-on: ubuntu-latest

steps:

- uses: actions/checkout@v2

- uses: actions/setup-node@v1

with:

node-version: '12.x'

- run: npm install -g geojson-rewind

- run: |-

curl "https://master.demo.geonode.org/geoserver/ows?

service=WFS&version=2.0.0&request=GetFeature&typename=geonode%3Aswamp&outputFormat=json&srs=EPSG%3A4326&srsName=EPSG%3A4326" | jq 'del(.requested_time,.crs,.timeStamp)' > swamp_tmp.geojson

- run: geojson-rewind swamp_tmp.geojson > swamp.geojson

- run: rm swamp_tmp.geojson

- run: |-

git config user.name "Automated"

git config user.email "actions@users.noreply.github.com"

git pull

git add -A

timestamp=\$(date -u)

git commit -m "Latest data: \${timestamp}" || exit 0

git push

name: GeoNode Layer scraper

on:

workflow_dispatch:

schedule:

- cron: '8,28,48 * * * *'

jobs:

scheduled:

runs-on: ubuntu-latest

steps:

- uses: actions/checkout@v2

- uses: actions/setup-node@v1

with:

node-version: '12.x'

- run: npm install -g geojson-rewind

- run: |-

curl "https://master.demo.geonode.org/geoserver/ows?

service=WFS&version=2.0.0&request=GetFeature&typename=geonode%3Aswamp&outputFormat=json&srs=EPSG%3A4326&srsName=EPSG%3A4326" | jq 'del(.requested_time,.crs,.timeStamp)' > swamp_tmp.geojson

- run: geojson-rewind swamp_tmp.geojson > swamp.geojson

- run: rm swamp_tmp.geojson

- run: |-

git config user.name "Automated"

git config user.email "actions@users.noreply.github.com"

git pull

git add -A

timestamp=\$(date -u)

git commit -m "Latest data: \${timestamp}" || exit 0

git push

4 GeoNode Api endpoints

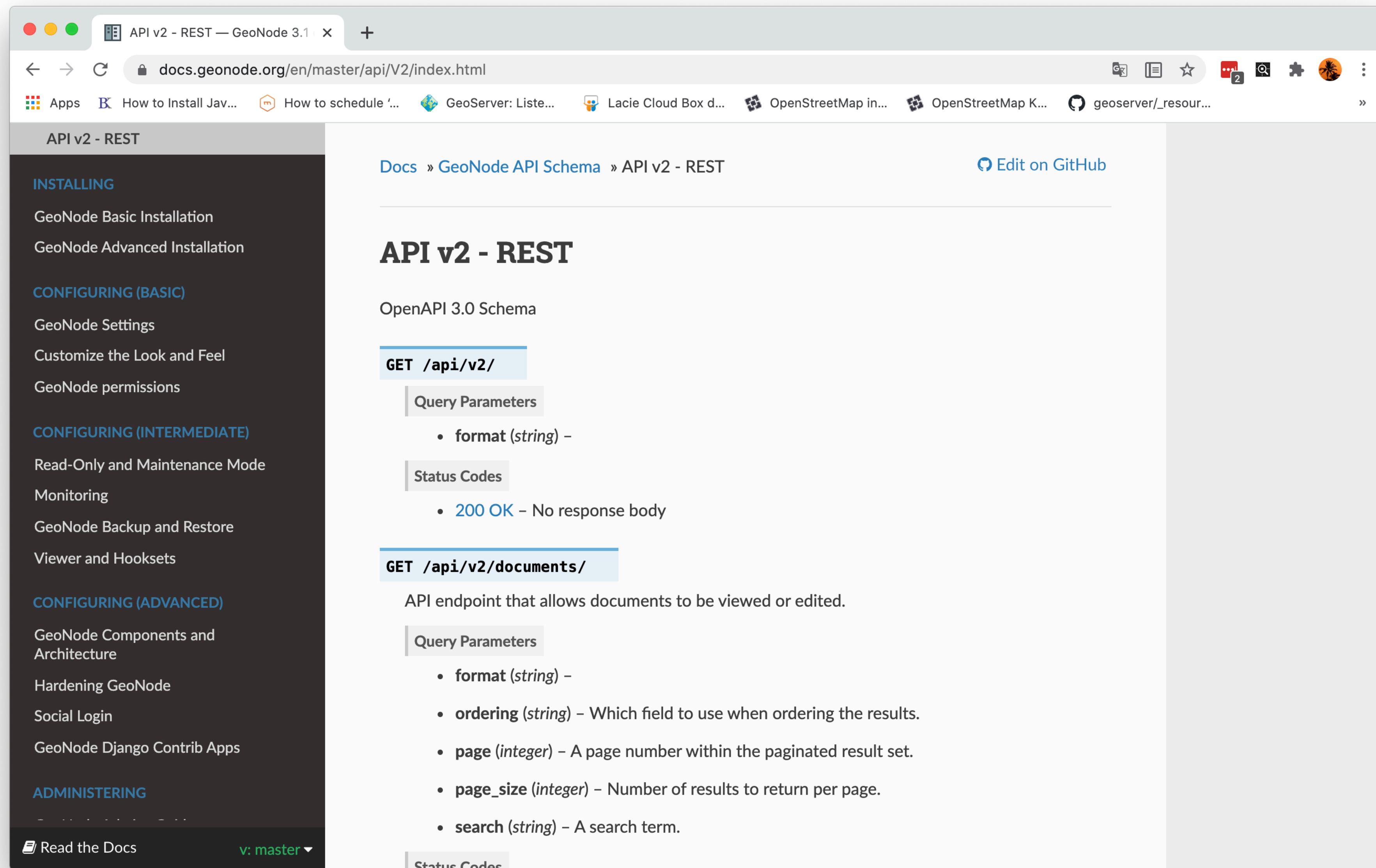
Visit the GeoNode API

<https://master.demo.geonode.org/api>

4 GeoNode Api endpoints

```
// 20201130213751
// https://master.demo.geonode.org/api/

{
  "base": {
    "list_endpoint": "/api/base/",
    "schema": "/api/base/schema/"
  },
  "categories": {
    "list_endpoint": "/api/categories/",
    "schema": "/api/categories/schema/"
  },
  "documents": {
    "list_endpoint": "/api/documents/",
    "schema": "/api/documents/schema/"
  },
  "featured": {
    "list_endpoint": "/api/featured/",
    "schema": "/api/featured/schema/"
  },
  "group_profile": {
    "list_endpoint": "/api/group_profile/",
    "schema": "/api/group_profile/schema/"
  },
  "groupcategory": {
    "list_endpoint": "/api/groupcategory/",
    "schema": "/api/groupcategory/schema/"
  },
  "groups": {
    "list_endpoint": "/api/groups/",
    "schema": "/api/groups/schema/"
  },
  "keywords": {
    "list_endpoint": "/api/keywords/",
    "schema": "/api/keywords/schema/"
  },
  "layers": {
```



<https://docs.geonode.org/en/master/api/V2/index.html>

Keep in mind.

Every action which requests content, puts load on your server.

Security: You can set your repository to private access. **GithubSecrets** and **basic auth headers** can be used for requesting private resources.

Keep in mind.

I would not suggest to use this technic with heavy layers and of course not with binary data.

Find my examples here

<https://github.com/t-book/track-geonode>

Thanks

Danke

Tak

Toda

Grazie

děkuji

gracias

dank je wel

blagodaram

شكرا

merci