

Web Scraping

Dr. Paul Escapil-Inchauspé
Universidad Adolfo Ibáñez, Chile

Proyecto 0
24 de agosto de 2022

1 Presentación del proyecto

Creación de Issue (state=open)

Error al realizar el push #34

[Edit](#)[New issue](#)

 **Open** cechiang opened this issue 8 hours ago · 1 comment



cechiang commented 8 hours ago



Hola,

Me sale este error al realizar el push.

```
$ git push --set-upstream origin users
remote: Permission to pescap/WebScrapingMDS.git denied to cechiang.
fatal: unable to access 'https://github.com/pescap/WebScrapingMDS/': The requested URL returned error: 403
```



pescap commented 7 hours ago

Owner



Relacionado con:

[#29](#)



pescap added **help wanted** **good first issue** labels 7 hours ago

Assignees

No one—assign yourself



Labels

good first issue **help wanted**



Projects

None yet



Milestone

No milestone



Development

[Create a branch](#) for this issue or link a pull request.



Notifications

[Customize](#)



Unsubscribe

Resolución de Issue (state=closed)

Regla = Closed + Assigned

Extraer data visible pero sin formato HTML #28

🔒 Closed

JoseAlexanderCL opened this issue 4 days ago · 2 comments

Edit

New issue



JoseAlexanderCL commented 4 days ago

Contributor



Alguien sabe como obtener extraer los nombres y estadísticas asociadas a las oportunidades de inversión que están aquí:

<https://cumplio.cl/oportunidades-de-inversion/>

De antemano gracias.



pescap commented 4 days ago

Owner



Hola! Debería funcionar todo perfecto con bs4. Toda la información aparece dentro del código html:

[illegible]

Assignees



JoseAlexanderCL

Labels



None yet

Projects



None yet

Mileston



No milestone

Development

[Create a branch](#)

Create a branch for this issue or link a pull request.

Notifications

Customize

 Unsubscribe

- state: {open, closed}
- author: usuario del author
- assignees: usuarios asignados

- Si todos los alumnos crean issues: 3×30 issues que monitorear
- Más trabajo todavía con los issues cerrados
- Más justo (sólo assignees por el profesor y/o ayudante)
- Más motivación

Objetivo final: Evaluación en tiempo real de la nota de los alumnos

1-Pseudo-code:

```
for cada periodo:  
  for cada issue:  
    buscar: state, author, assignees  
    agregarlo a una fila de DataFrame
```

2-Realizar manualmente la operación para algunos casos

3-Programar la operación “buscar”

4-Escalar a todos los issues

5-Escalar a todos los periodos

6-Desplegar y monitorear

- No hay API
- Código HTML clásico
- No hay acciones o código dinámico
- No parece haber otras restricciones de conexión
- No se necesita login

Error al realizar el push #34

[Edit](#)[New issue](#)

 Open cechiang opened this issue 10 hours ago · 1 comment

author

state



cechiang commented 10 hours ago



Hola,

Me sale este error al realizar el push.

```
$ git push --set-upstream origin users
remote: Permission to pescap/WebScrapingMDS.git denied to ceciang.
fatal: unable to access 'https://github.com/pescap/WebScrapingMDS/': The requested URL returned error: 403
```



pescap commented 9 hours ago

Owner



Relacionado con:

[#29](#)

Assignees

No one—assign yourself



Labels

[good first issue](#)

[help wanted](#)



Projects

None yet



Milestone


No milestone

















Development

Create a branch for this issue or link a pull request.





 **pescap / WebScrapingMDS** Public



 Pin  Unwatch **4**  Fork **29**  Starred **6**

 Code  **Issues 9**  Pull requests  Discussions  Actions  Projects  Wiki  Security  Insights  Settings

Extraer data visible pero sin formato HTML #28



 Closed JoseAlexanderCL opened this issue 5 days ago · 2 comments

 state

**JoseAlexanderCL** commented 5 days ago Contributor 

Alguien sabe como obtener extraer los nombres y estadísticas asociadas a las oportunidades de inversión que están aquí:
<https://cumplio.cl/oportunidades-de-inversion/>


De antemano gracias.

**pescap** commented 4 days ago Owner 

Hola! Debería funcionar todo perfecto con bs4. Toda la información aparece dentro del código html:

```
print('century background: rgb(255, 255, 255);background-color: #f0f0f0;...'
```

Assignees

 **JoseAlexanderCL**

Labels

assignees

None yet

Projects

None yet

Milestone

No milestone

Producto esperado - I

```
In [106]: df
Out[106]:
```

	issue	author	state	assignees
0	1	pescap	open	No one assigned
1	2	pescap	closed	pescap
2	2	pescap	closed	caspoerer
3	4	pescap	closed	No one assigned
4	5	GeorgeAdrock	open	GeorgeAdrock
5	7	pescap	closed	No one assigned
6	15	DenisOsses	closed	No one assigned
7	17	CarlosZuritaEstolaza	open	CarlosZuritaEstolaza
8	20	CarlosZuritaEstolaza	open	No one assigned
9	21	gavalenz	open	No one assigned
10	24	shorokein	open	No one assigned
11	25	josemoragonzalez	closed	pescap
12	25	josemoragonzalez	closed	josemoragonzalez
13	27	pescap	open	No one assigned
14	28	JoseAlexanderCL	closed	JoseAlexanderCL
15	29	iggypoprl	open	No one assigned
16	34	cechiang	open	No one assigned

Número de columnas: 4.

Estimación número de filas: $38 \times 6 = 228$.

Orden de magnitud esperado 1000 filas aproximadamente.

Producto esperado - II

