

Identificando informações de um Website

Identificando o proprietário de um Website utilizando a biblioteca
python-whois

Identificando o proprietário de um Website

Introdução

Para alguns sites, pode ser importante para nós descobrirmos quem é o proprietário. Por exemplo, se uma empresa for conhecida por bloquear *web crawlers*, seria interessante sermos mais conservadores na taxa de download.

Para descobrir quem é o proprietário de um site, podemos usar o protocolo *WHOIS*. Existe um módulo do *Python* que utiliza este protocolo, o *python-whois*.

Identificando o proprietário de um Website

Instalando a biblioteca python-whois

Podemos instalar utilizando o comando:

\$ pip install python-whois

```
C:\Users\evaldo
λ pip install python-whois
Collecting python-whois
  Downloading https://files.pythonhosted.org/packages/63/8a/8ed58b8b28b6200ce1cdfe4e4f3bbc
8b85a79eef2aa615ec2fef511b3d68/python-whois-0.7.0.tar.gz (82kB)
    100% |████████████████████| 92kB 169kB/s
Collecting future (from python-whois)
  Downloading https://files.pythonhosted.org/packages/00/2b/8d082ddfede935f3608cc61140df6dc
bf0edealbc3ab52fb6c29ae3e81e85/future-0.16.0.tar.gz (824kB)
    100% |████████████████████| 829kB 147kB/s
mysql-connector-python 8.0.11 requires protobuf>=3.0.0, which is not installed.
Installing collected packages: future, python-whois
  Running setup.py install for future ... done
  Running setup.py install for python-whois ... done
Successfully installed future-0.16.0 python-whois-0.7.0
```

Identificando o proprietário de um Website

Executando o python-whois

Para executar, devemos importar o módulo *whois*:

```
>>> import whois
```

E executar o método *whois* passando o endereço do site, como no exemplo abaixo:

```
>>> whois.whois('globo.com')
```

Identificando o proprietário de um Website

Executando o python-whois

Veja o resultado de uma execução:

```
Cmder
>>> print(whois.whois("globo.com"))
{
  "domain_name": "GLOBO.COM",
  "registrar": "1API GmbH",
  "whois_server": "whois.1api.net",
  "referral_url": null,
  "updated_date": [
    "2018-05-23 16:27:42",
    "2018-05-23 16:27:41"
  ],
  "creation_date": "1998-12-21 05:00:00",
  "expiration_date": "2020-12-21 05:00:00",
  "name_servers": [
    "NS01.GLOBO.COM",
    "NS02.GLOBO.COM",
    "NS03.GLOBO.COM",
    "NS04.GLOBO.COM",
    "ns01.globo.com 131.0.24.26 2804:0294:0100:0803:0131:0000:0024:0026",
    "ns02.globo.com 186.192.89.18",
    "ns03.globo.com 186.192.89.5 2804:0294:4000:8001:0000:0000:0000:0005",
    "ns04.globo.com 177.53.95.213 2804:0294:8000:0200:0177:0053:0095:0213"
  ],
  "status": [
```

```
Cmder
],
  "status": [
    "clientDeleteProhibited https://icann.org/epp#clientDeleteProhibited",
    "clientTransferProhibited https://icann.org/epp#clientTransferProhibited",
    "clientDeleteProhibited - http://www.icann.org/epp#clientDeleteProhibited",
    "clientTransferProhibited - http://www.icann.org/epp#clientTransferProhibited"
  ],
  "emails": [
    "abuse@1api.net",
    "dns-admin@corp.globo.com",
    "dns-tech@corp.globo.com"
  ],
  "dnssec": "unsigned",
  "name": "Globo Comunicacao e Participacoes S/A",
  "org": "27.865.757/0001-02",
  "address": "Rua Lopes Quintas",
  "city": "Rio de Janeiro",
  "state": "RJ",
  "zipcode": "22460-010",
  "country": "BR"
}
>>> |
```

FIM