



OSINT, OSANT CADA DÍA TE QUIERO MÁS: CASO REAL DE OSINT

INVESTIGACIONES OSINT EL PRESENTE Y FUTURO

JORGE CORONADO CEO DE QUANTIKA14



GDG Sevilla

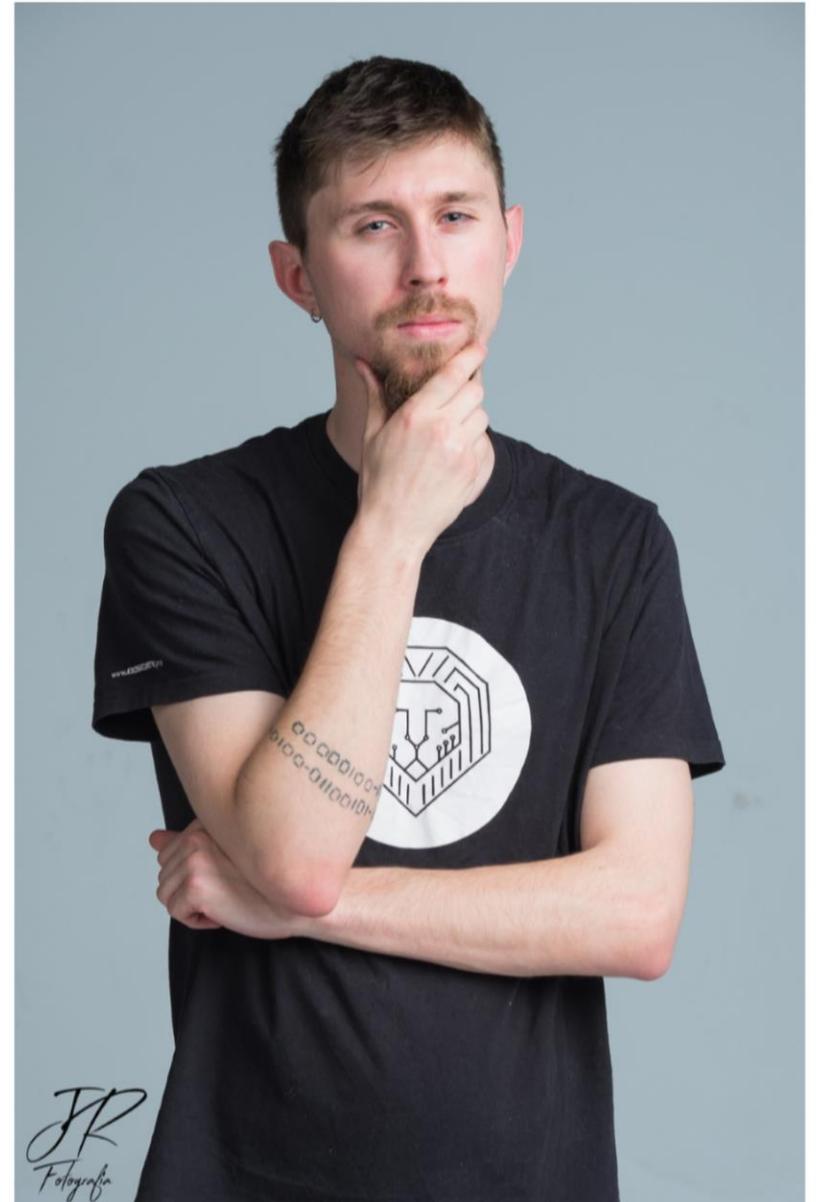
Quién es Jorge Coronado

- Fundador y CEO de **QuantiKa14**
- Colaborador de **Canal Sur Radio** desde 2015
- **Profesor** en el curso de **detectives de la Universidad Pablo Olavide** de Sevilla
- Co-autor del primer **“Protocolo institucional en España ante la violencia de género en las redes sociales”**
- **Formación a cuerpos de seguridad en investigación a través de Internet** desde la ESPA y otros cursos
- Creador del **protocolo de actuación para la búsqueda de personas desaparecidas a través de las tecnologías de la información y comunicación**
- **Vocal** de la **asociación de peritos tecnológicos de Andalucía (APTAN)**
- Dinamizador del **Hack&Beers Sevilla**
- Autor del canal de Youtube **Investiga Conmigo desde el Sü**
- **Podcast -> Embaja de Triana**
- Creador de aplicaciones (free) como: **Guasap Forensic, Shodita, EO-Ripper, Dante Gates, Killo.io, etc**
- Director y profesor del curso de verano de la Pablo de Olavide sobre ciberdelincuencia de género (2018 y 2019)



2019

11/10/2019



DEVFEST 2019 - GDG SEVILLA

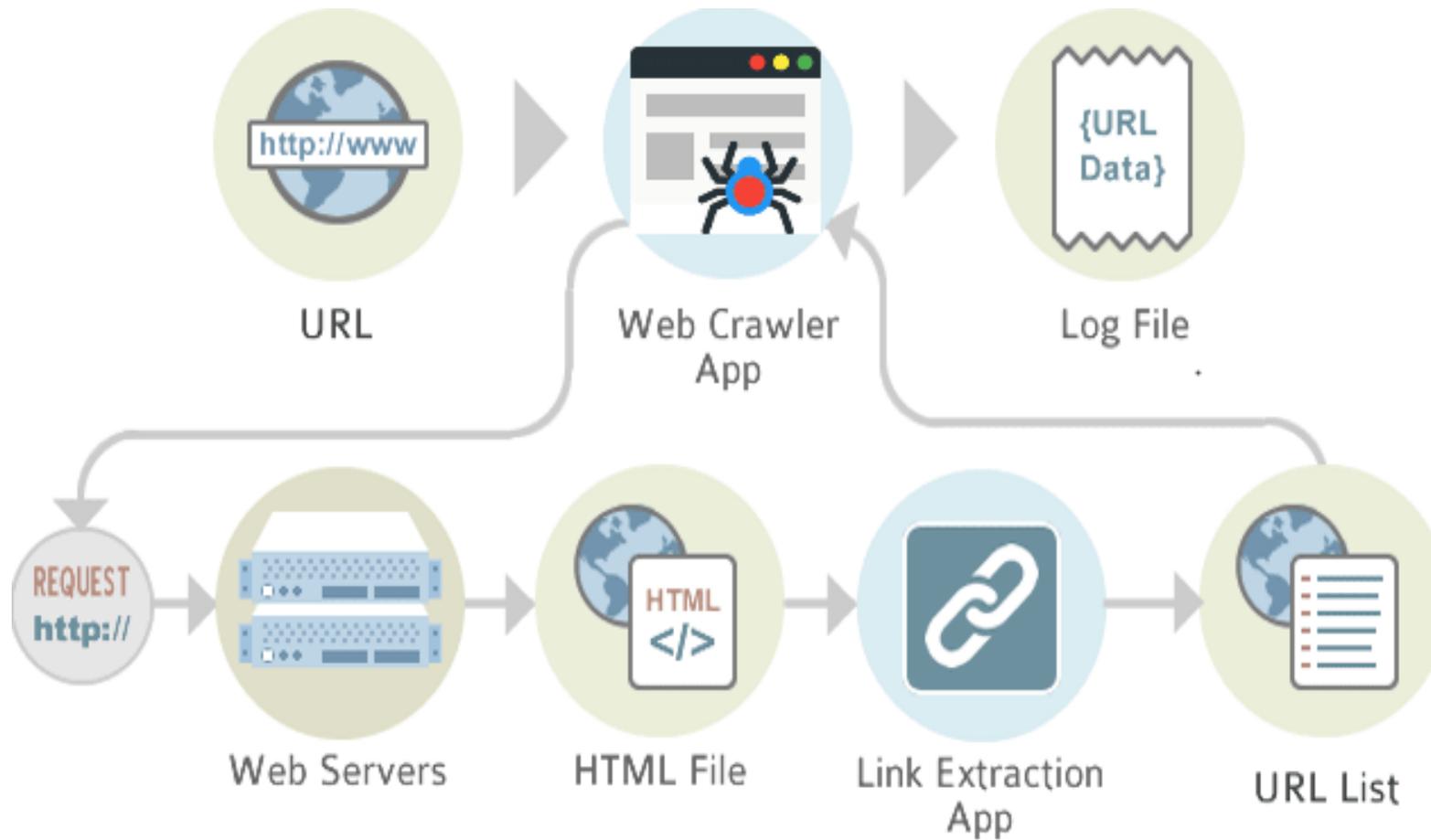
¿Qué vamos a ver?

- 1. Conceptos básicos
 - OSINT
 - SOCMINT
 - BOTS
 - CRAWLERS
- 2. Qué es Dante's Gates Minimal Version
- 3. Cómo instalarlo
- 4. EO-ripper
- 5. Demo time



1.3 ¿Qué es un bot?

1.4 ¿Qué es un Crawler?

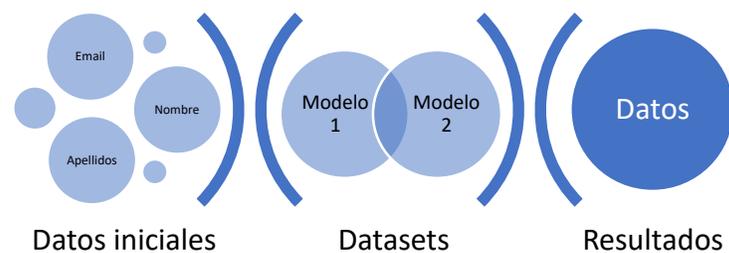




¿Qué son los datos iniciales?

Metodología

1. Identificación de datos iniciales
2. Análisis automático con buscadores (Dante's Gates Minimal Version, EO-ripper, etc)
3. Análisis manual
4. Creación de informe



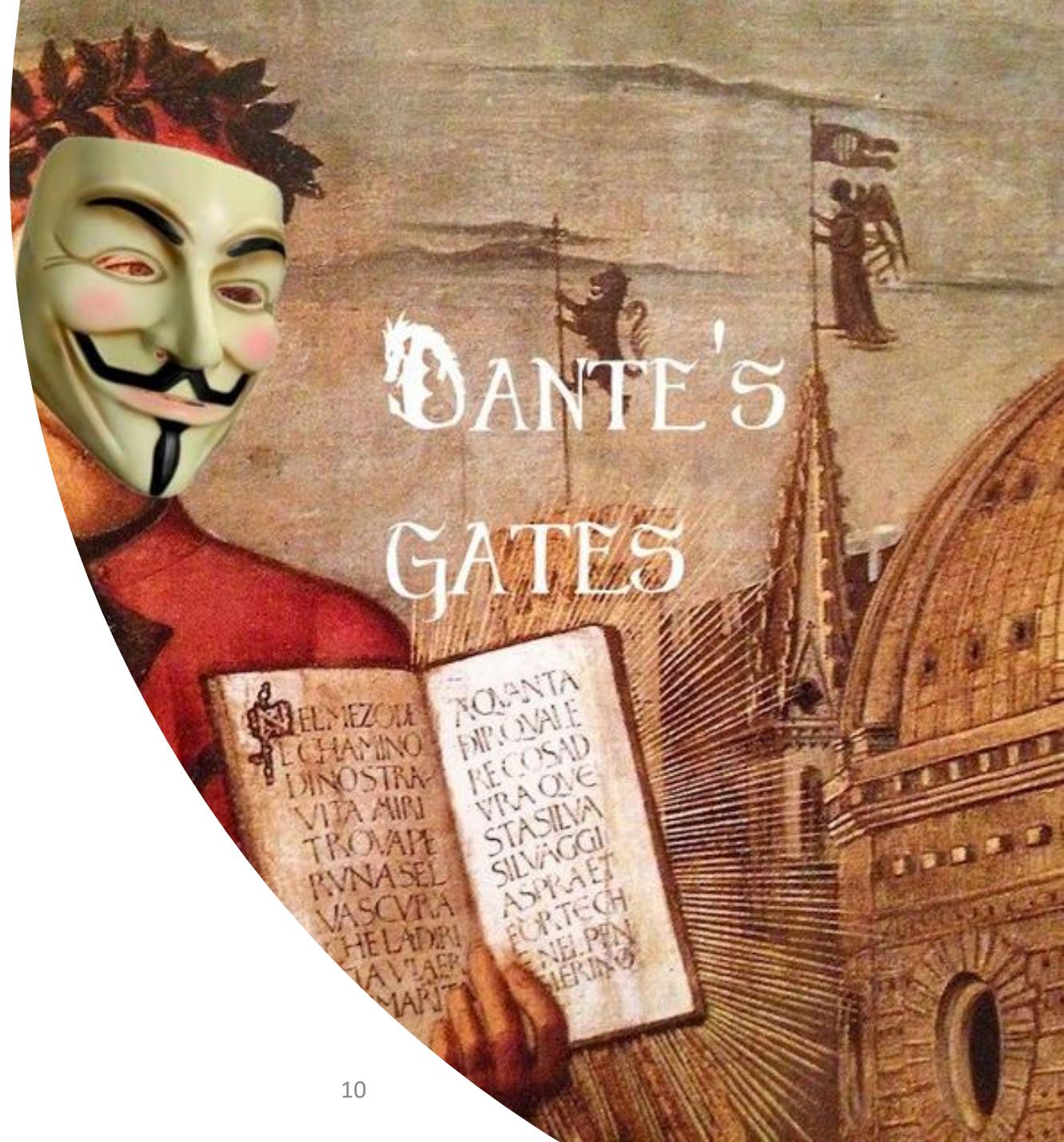
Con los **datos iniciales** creamos un dataset de información que tendrá que ser verificado.

2. ¿Qué es Dante's Gates Minimal Version?

- Código abierto
- Python 2.7
- Suit de herramientas para hacer OSINT en España
- [CÓMO DESCARGAR:](https://github.com/Quantika14/osint-suite-tools)
<https://github.com/Quantika14/osint-suite-tools>

11/10/2019

DEVFEST 2019 - GDG SEVILLA

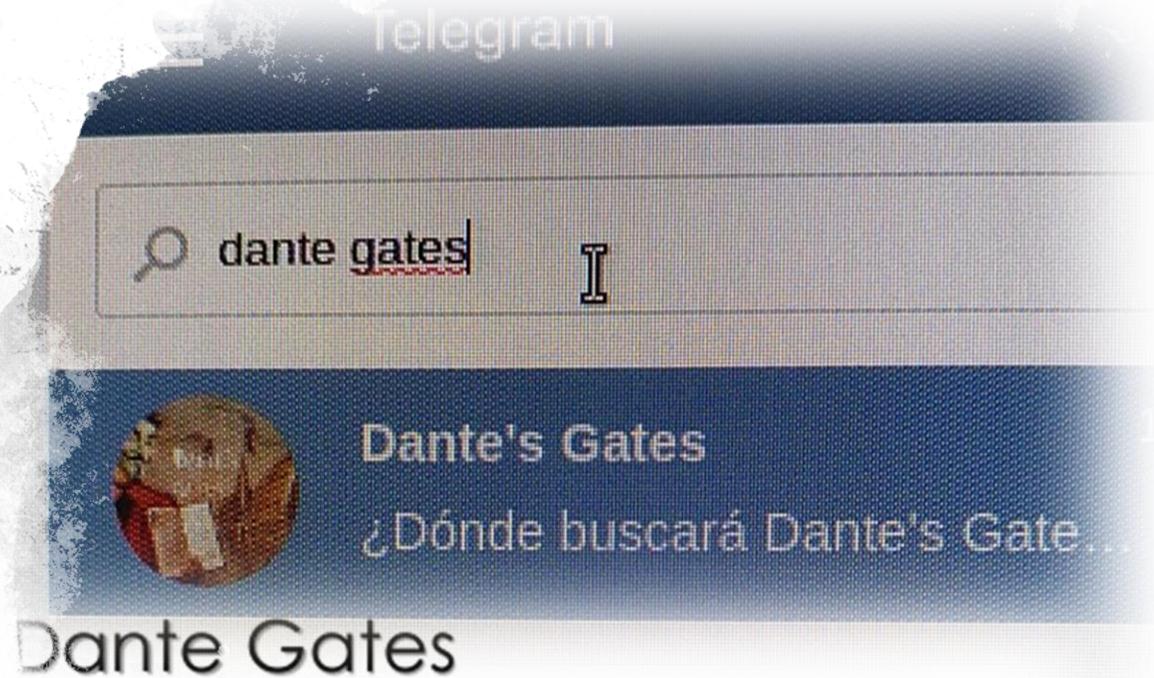


¿De dónde viene DG MV?

Un Proyecto de QuantiKa14.

- Página web -> acceso limitado
- Bot de Telegram -> DEAD

Proyecto open source: Dante's Gates
Minimal version



2.1 ¿A quién le puede interesar DG.MV?

- Detectives
- Periodistas
- Peritos informáticos
- Cuerpos de seguridad
- Recursos humanos

2.2 ¿En qué puede ayudar?

- **Investigaciones privadas usando OSINT** (casos de corrupción, desfalco, terrorismo, etc)
- **Auditorías de seguridad informática** (proceso de footprinting y figerprinting)
- **Marketing** (estudios de mercados, competencia, etc)
- **Prueba digital** para un peritaje informático

Transparencia ciudadana

DISCLAIMER:

El autor de esta aplicación no se ha responsable de su uso. Su intención es formativa y capacitar a los ciudadanos a ejercer su derecho a la transparencia en España con el uso de herramientas OSINT.

2.3 ¿Cómo se estructura Dante's Gates Minimal Version?

- Buscadores (scripts):
 - Buscador de personas
 - Buscador de niks
 - Buscador de emails
 - Buscador de Ips (-)
 - Buscador de TLFN(-)
 - Buscador de empresas(-)
- Bots y crawlers:
 - LinkedIn
 - Web
 - Pdfs
 - BOE y BORME
 - TESTRA
 - Adjudicaciones (-)
 - Imágenes
- Archivos: JSON, CSV, PDF e imágenes (por major y guardar en DB)
- Base de datos en local con MongoDB (crear una API)



2.5 APIS y librerías

- Wikipedia:
<https://pypi.org/project/wikipedia/>
- Libreforme API:
<https://libreforme.readthedocs.io/es/latest/api/>
- Dryscape: para scraping en webs con JS
- BeautifulSoup: parsing
- PDFGREP: para extraer texto PDF

```
>>> import wikipedia
>>> print wikipedia.summary("Wikipedia")
# Wikipedia (/ˌwɪkɪˈpiːdiə/ or /ˌwɪkiˈpiːdiə/ WIK-i-PEI...

>>> wikipedia.search("Barack")
# [u'Barak (given name)', u'Barack Obama', u'Barack (B...
```

```
$ curl -s "https://libreforme.net/borme/api/v1/empresa/search/?q=Gowex+Malaga&page=1" | python -m
{
  "objects": [
    {
      "name": "GOWEX MALAGA",
      "resource_uri": "/borme/api/v1/empresa/gowex-malaga/",
      "slug": "gowex-malaga"
    }
  ]
}
```

2.6 Buscador de Nicks

- 90 webs
- Verificación por URL
- Mejoras:
 - Base de datos
 - Usar buscadores (DDG, Google, Bing, etc)

```
os +hy-
`hh .ys` .hs /y/ +d-
-yd- +hs. ./ `+dd- - `odo` yh:
`hy `ohh/./-syyddss+-o ./yds. -d+
:d+ `oydhs/-shhddhy+-:ohdhs. `ys
+h:- `ohdmmmmmmmmddmmhs. `+y.
/hdhsyhdddmmNNmmmmmmhyyyshhy.
-/oyhydmmmmmdmmdyyhyyyo/.
`yhddmmdymdhoyddhys`
`.:+sydyhdddhdddhhhdysyo+/.
`.:+o+.:. :dmmmmmmmmddmddd: -+o+-
-+oo:.` +dmmmmmmmmssdmhy/ -/+::`
.+::.` `shmmNdo/ymmmmmy/+shsyo .::/`
:+` -sydmmhs- /nmhmy+/+sdhso` -+`
o /syhmmdy:ommys//+ymhoo. -/
/- oyyhmmh+omddo++oymyss- o`
++ `ohhdmmmy+-hddy+//+shhss: :/
+::.` :shhyddNmyo:dddh+//+shyyhy+ `o:
`y.:hyydNmdsymmmd++++sdsyo-/y` `::`
`y/ /yyhmNmhmhmmmmho+odyss. y+
os /yydmmmmmmmyohdyo. -h-
:h- -+sdmmmmddmmyho:` +s.
:y+ ./shhhyyhhho-` `ss`
o-` ...` .o`
.s` :/
.o` ::
s` o`
.s` :/
:s-` .o+
..`
```

OSINT
PARA
TODOS
E
INVESTIGA
CONMIGO

ANTE'S GATES MINIMAL v 1.0 | <<TIP-1337>> | Buscador De Nicks | QUANTIKA14 | @JORGEWEBSEC
VERSION: 1.0 | 19/02/2019 | INVESTIGA CONMIGO DESDE EL SU | WWW.QUANTIKA14.COM

El buscador de nicks no es perfecto. Necesita la colaboración de todos para mejorar.
Si quieres ayudarnos con Dante's Gates Minimal Version solo tienes que compartirnos tu idea
Si hay un fallo o mejoras puedes subirlo en issues aquí:
<https://github.com/Quantika14/osint-suite-tools/issues>

2.7 Tipos de bots/crawlers

- **Estáticos:** es utilizado solo una vez
- **Dinámicos:**
 - Aplicaciones que se ejecutan cada X tiempo
 - Aplicaciones que están en ejecución todo el tiempo

```
sudo crontab -e
```

Ejecutar un script de lunes a viernes a las 2:30 horas:

```
30 2 * * 1-5 /bin/ejecutar/script.sh
```

Ejecutar un script de lunes a viernes cada 10 minutos desde las 2:00 horas durante una hora:

```
0,10,20,30,40,50 2 * * 1-5 /bin/ejecutar/script.sh
```

Esto quizá puede ser largo. La sintaxis de crontab permite lo siguiente. Imaginemos que queremos ejecutarlo cada 5 minutos:

```
*/5 2 * * 1-5 /bin/ejecutar/script.sh
```

```
vm@makina:~/Escritorio/osint-suite-tools-master/bots/PDF_crawler_webs$ python pdfget.py generalisimofranco http
start crawling:
pdfget.py:37: UserWarning: No parser was explicitly specified, so I'm using the best available HTML parser for
m, or in a different virtual environment, it may use a different parser and behave differently.
```

```
The code that caused this warning is on line 37 of the file pdfget.py. To get rid of this warning, pass the add
```

1
9

2.8 Ejemplo bot estático:

```
soup = BeautifulSoup(context)
HTTP Error 404: Not Found
going to fetch: http://www.generalisimofranco.com/GC/asesinados01/PFD_024.pdf
going to fetch: http://www.generalisimofranco.com/GC/asesinados01/PFD_024.pdf
going to fetch: http://www.generalisimofranco.com/GC/asesinados01/PFD_024.pdf
going to fetch: http://www.generalisimofranco.com/GC/asesinados01/PFD_024.pdf
going to fetch: http://www.generalisimofranco.com/GC/BCL/PDF/023.pdf
going to fetch: http://www.generalisimofranco.com/GC/BCL/PDF/023.pdf
going to fetch: http://www.generalisimofranco.com/GC/BCL/PDF/023.pdf
going to fetch: http://www.generalisimofranco.com/GC/BCL/PDF/023.pdf
going to fetch: http://www.generalisimofranco.com/GC/URSS/PDF011.pdf
going to fetch: http://www.generalisimofranco.com/GC/URSS/PDF011.pdf
going to fetch: http://www.generalisimofranco.com/GC/BCL/PDF/022.pdf
going to fetch: http://www.generalisimofranco.com/GC/BCL/PDF/022.pdf
going to fetch: http://www.generalisimofranco.com/GC/BCL/PDF/021.pdf
going to fetch: http://www.generalisimofranco.com/GC/BCL/PDF/021.pdf
going to fetch: http://www.generalisimofranco.com/VIDAS/jose_diaz_ramos/PDF02.pdf
going to fetch: http://www.generalisimofranco.com/VIDAS/jose_diaz_ramos/PDF02.pdf
going to fetch: http://www.generalisimofranco.com/GC/BCL/PDF/019.pdf
going to fetch: http://www.generalisimofranco.com/GC/BCL/PDF/019.pdf
going to fetch: http://www.generalisimofranco.com/GC/URSS/PDF010.pdf
going to fetch: http://www.generalisimofranco.com/GC/URSS/PDF010.pdf
```

- PDF_CRAWLERS WEBS:
<https://github.com/Zealcui/pdfcrawler>
- \$python pdfcrawler.py %dir %URL

```
HTTP Error 404: Not Found
HTTP Error 403: Forbidden
```

2.9 Ejemplo de bot dinámico y continuo

- Adjudicaciones
- BOE y BORME
- WEB crawler
- Redes Sociales
- Leak Offshore
(<https://offshoreleaks.icij.org/>)

s_Acoruna.py
s_Andalucia.py
s_AyuntamientoAlbacete.py
s_AyuntamientoAlicante.py
s_AyuntamientoAlmeria.py
s_AyuntamientoAvila.py
s_AyuntamientoBarcelona.py
s_AyuntamientoBilbao.py
s_AyuntamientoCadiz.py
s_AyuntamientoCeuta.py
s_AyuntamientoCordoba.py
s_AyuntamientoGirona.py
s_AyuntamientoGranCanaria.py
s_AyuntamientoGuadalajara.py
s_AyuntamientoHuelva.py
s_AyuntamientoLleida.py
s_AyuntamientoMadrid.py
s_AyuntamientoMurcia.py
s_AyuntamientoSevilla.py
s_AyuntamientoTarragona.py
s_AyuntamientoValladolid.py
s_AyuntamientoZamora.py
s_AyuntamientoZaragoza.py
s_Canarias.py
s_CastillayLaMancha.py
s_Cataluña.py
s_ComunidadMadrid.py
s_DiputacionCadiz.py
s_DiputacionGranada.py
s_DiputacionHuelva.py
s_Larioja.py
s_Murcia.py
s_Paisvasco.py

2.10 Ejemplo de bot dinámico

- **TESTRA:** ejecutar la aplicación todos los días a las 10:30 de la mañana

```
vm@makina: ~/Escritorio/osint-suite-tools-master/bots/PDF_crawler_webs
Archivo Editar Ver Buscar Terminal Ayuda
GNU nano 2.9.3 /tmp/crontab.zKiFwW/crontab Modifi

# daemon's notion of time and timezones.
#
# Output of the crontab jobs (including errors) is sent through
# email to the user the crontab file belongs to (unless redirected).
#
# For example, you can run a backup of all your user accounts
# at 5 a.m every week with:
# 0 5 * * 1 tar -zcf /var/backups/home.tgz /home/
#
# For more information see the manual pages of crontab(5) and cron(8)
#
# m h dom mon dow   command
30 10 * * 0-6 /opt/DGMV/TESTRA/testra.py
```

2.11 Tor Spider:

- Enlace: https://github.com/absingh31/Tor_Spider
- En el archive onions copiamos las rutas que queremos descargar
- Copiamos los archivos generados dentro de la carpeta data

The screenshot shows the GitHub interface for a repository named 'der'. At the top right, there are buttons for 'Watch' (3), 'Star' (16), and 'Fork'. Below these are navigation links for 'Pull requests 0', 'Projects 0', 'Wiki', and 'Insights'. The main content area features a text snippet: 'nd scrap the lesser known deep web or one can say dark web. Just provide the onion link and get started.' Below this are several topic tags: 'socks', 'stem', 'python3', 'tor-config', 'tor-spider', 'ioc', 'file-manager', and 'scraping'. Further down, it indicates '1 branch', '0 releases', and '1 contributor'. A navigation bar includes a 'request' button and options to 'Create new file', 'Upload files', 'Find File', and 'Clone or download'. The commit history table below shows the following entries:

| Commit Message | Time |
|---|---------|
| remove | a year |
| onion link example added | 2 years |
| update | a year |
| update | a year |
| made corresponding change to crawl_bot.py | 2 years |
| some functions to manage the file | 2 years |
| extract the domain names | 2 years |
| onion link example added | 2 years |
| starting to get the crawler run | 2 years |
| update | a year |
| update | a year |

2.12 Problemas actuales

- Python 2.7
- Buscadores como Google, DuckDuckGo, Bing, Shodan, etc.
- Open data de la administración
- Lento



3. ¿Cómo instalar Dante's Gates Minimal Version?

1. En Linux descargamos el repositorio:
<https://github.com/Quantika14/osint-suite-tools>
2. Lo descomprimos en /opt/
3. Instalamos PDFgrep: “sudo apt install pdfgrep”
4. Instalamos las librerías: “sudo pip install -r requirements.txt”
- 5. Ejecutamos nuestros bots para obtener información
- 6. Ejecutamos el buscador que queremos:
 - BuscadorPersonas.py
 - BuscadorNiks.py



2.13 Mejoras

- Python 3
- Configurar conexiones y búsquedas por base de datos
- Disminuir archivos en data, insertarlos y crear un índice
- Multihilos
- Expresiones regulares
- Sistema de informes
- Interfaz web
- Paraísos fiscales
- Registros mercantiles de otros países

CÓMO CREAR UN MOTOR DE BÚSQUEDA DE LA



4. ¿Cómo crear un buscador de la Deepweb?

- Darksearch api
- Intelx.io
- Crawlers buscando .onion
- Crawler recursive buscando .onion en la deepweb

<https://blog.quantika14.com/blog/2019/08/13/como-crear-tu-propio-motor-de-busqueda-de-la-deepweb-python-y-api-de-intelx-io/>

<https://blog.quantika14.com/blog/2019/08/05/como-crear-tu-propio-buscador-de-la-deepweb-python-y-api-de-darksearch/>

<https://blog.quantika14.com/blog/2019/07/31/como-crear-tu-propio-buscador-de-la-deepweb/>

4.1 Configurar API DarkSearch

```
session = requests.session()
session.proxies = {}

session.proxies['http'] = 'socks5h://localhost:9050'
session.proxies['https'] = 'socks5h://localhost:9050'
```

```
19 JSON = json.loads(data_JSON)
20
21 pages = JSON["last_page"]
22
23 for p in range(1,pages):
24     for i in range(0,20):
25         try:
26             URL = "https://darksearch.io/api/search?query="+ query + "&page=" + str(p)
27             r = session.get(URL)
28             data_JSON = r.text
29             JSON = json.loads(data_JSON)
30             print(JSON["data"][i]["title"])
31         except:
32             print("[INFO][>] DarkSearch has broken the connection...")
33
```

Params :

```
{
  "query": string,
  "page": int
}
```

Response :

```
{
  "total": int,
  "per_page": int,
  "current_page": int,
  "last_page": int,
  "from": int,
  "to": int,
  "data": [
    {
      "title": string,
      "link": string,
      "description": string
    }
  ]
}
```

Captura de pantalla del funcionamiento de la API y la estructura JSON.

<https://blog.quantika14.com/blog/2019/08/05/como-crear-tu-propio-buscador-de-la-deepweb-python-y-api-de-darksearch/>

Code | Issues 0 | Pull requests 0 | Security | Insights

Dismiss

Join GitHub today

GitHub is home to over 40 million developers working together to host and review code, manage projects, and build software together.

4.2 Configurar API IntelX es aún más fácil

Branch: master | SDK / Python / ix_search.py | Find file | Copy path

Kleissner Minor fixes. Updating the Python code. ad5c58d on 27 Jul

https://github.com/IntelligenceX/SDK/blob/master/Python/ix_search.py

107 lines (87 sloc) | 3.34 KB | Raw | Blame | History

```

1  #!/usr/bin/env python
2  import os
3  import re
4  import sys
5  import time
6  import json
7  import urllib
8  import requests

```

```
Editor VIM Buscar Terminal Ayuda
root@kali:~/Workspaces/email-osint-ripper-master# sudo python eo-ripper.py

O-RIPPER.P

Author: Jorge Hebeac | Twitter: @JorgeHebeac | jorge.coronado@quantika.es
Can I know with your email?
Only 1 email or emails list
Verify emails
Verify LinkedIn, WordPress, Amazon[ES], Tumblr, Netflix and DOC Hacking
Pastebin

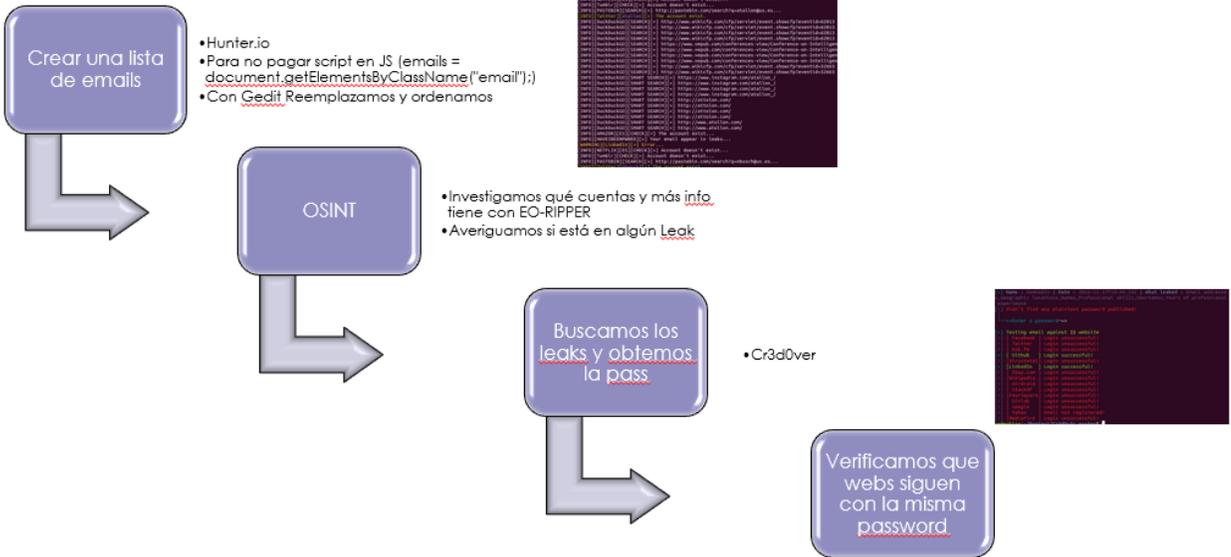
Usage: 09/01/2017 | Version: 1.0
Test version: 11/01/2017 | Version: 1.0.1
Test version: 17/07/2018 | Version: 1.0.8
Test version: 11/07/2018 | Version: 1.2.1

Emails list (default: emails.txt)
Only one target
Email spoofing generate

1/3/19
```

5. Buscador de Email: EO-Ripper

<https://github.com/Quantika14/email-osint-ripper>



A magnifying glass is positioned over a white computer keyboard. The text '¿Investigas con migo?' and 'DEMO TIME' is overlaid in white. The background is a dark blue gradient.

¿Investigas con migo? DEMO TIME

Archivo Editar Ver Historial Marcadores Herramientas Ayuda

Investiga-conmigo/sevilla.json x Dante Gates 3 1 x Teseo

dantegates.pro/resolution

Comprar Creditos 74 RTL

Investiga conmigo desde el Sü: capítulo I

TIPOS DE BÚSQUEDA

- Personas
- Medios
- Twitter

El uso de la obtención de información de fuentes abiertas en Internet de forma inteligente. DantesGates tiene el objetivo de ser un <all in one OSINT> para investigadores. Encuentra más información en el blog de Quantika14. Gracias.

ADJUDICACIONES

Adjudicatario: CARTUJA INMOBILIARIA, S.A.U

NIF: A78941960

Fecha de la adjudicación: 14/11/2008

Caracter definitivo: No

Importe: 3.851.337,95€

Adjudicatario: CARTUJA INMOBILIARIA, S.A.

NIF: A78941960

Fecha de la adjudicación: 12/12/2008

Localización: Andalucía

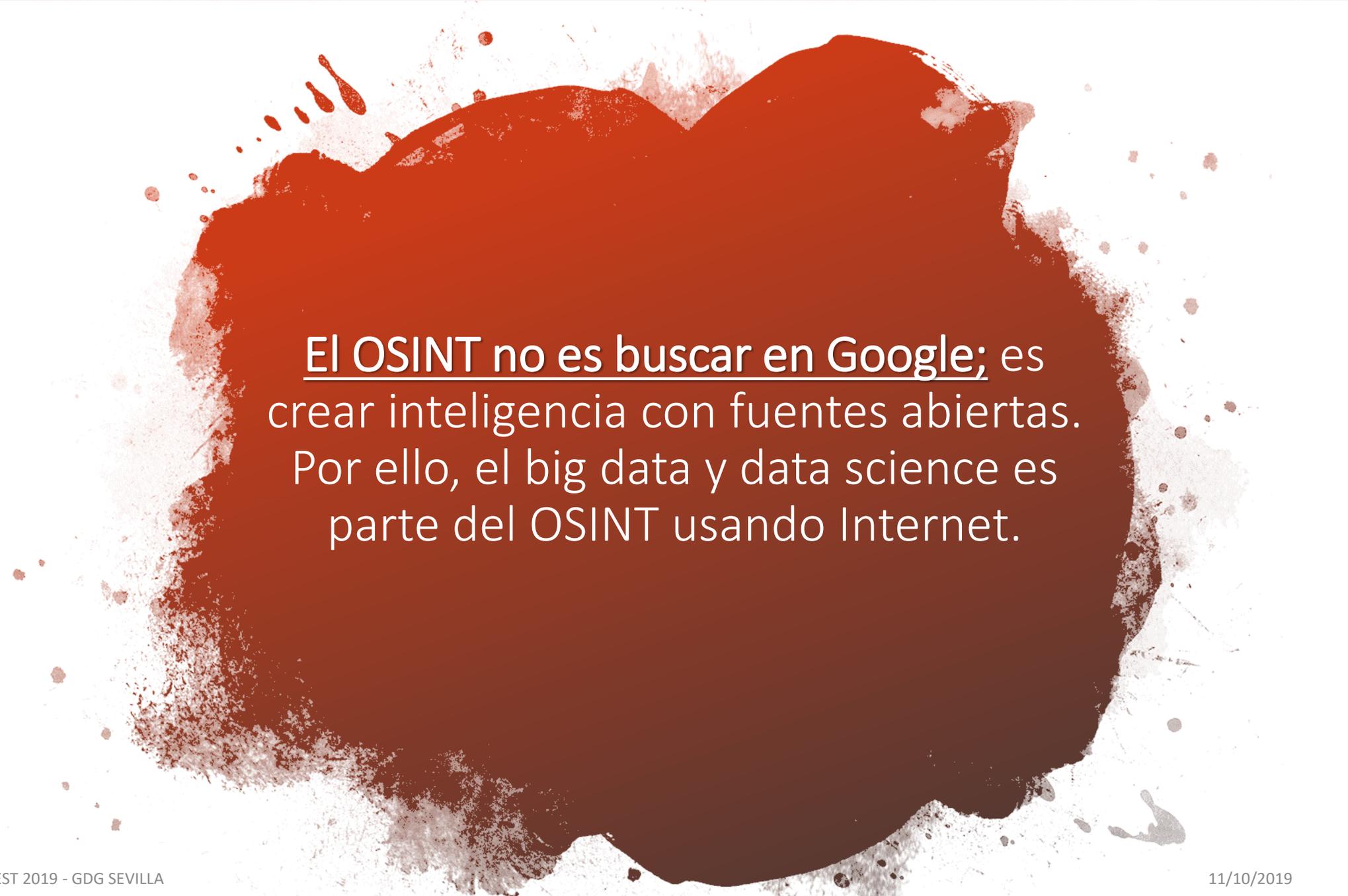
DEVFEST 2019 - GDG SEVILLA

Caracter definitivo: No

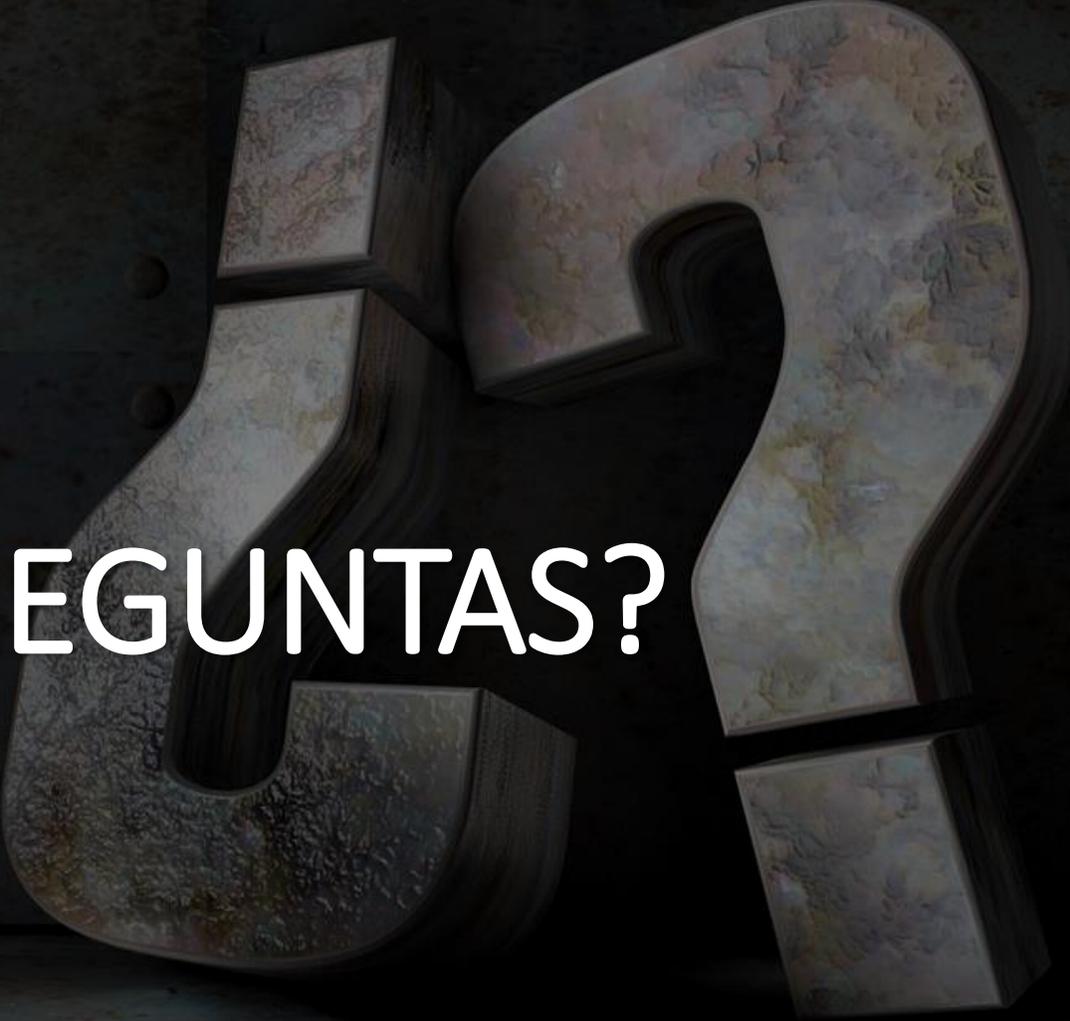
11/10/2019

Importe: 1.334.000,00€

- <https://www.youtube.com/watch?v=zI11NiQCX-I>



El OSINT no es buscar en Google; es crear inteligencia con fuentes abiertas. Por ello, el big data y data science es parte del OSINT usando Internet.



¿PREGUNTAS?



¡Muchas
gracias por su
atención!

www.quantika14.com

PRODUCIDO POR QUANTIKAI14
Investiga conmigo
desde el sü

¡NO TE OLVIDES DE VER!

INVESTIGA CONMIGO DESDE EL SÜ

https://www.youtube.com/channel/UCotPHyHsSSyhly_RN02jvSg

