

# OSINT With Datasploit

With all this talk of shifting security left, it's important for developers to know that different tools available for security testing. Today, we check out Datasploit.



by Russ McRee MVB · Aug. 23, 17 · Security Zone · Tutorial

Like (2) Comment (0) Save Tweet

3,931 Views



I was reading an interesting [Motherboard](#) article, [Legal Hacking Tools Can Be Useful for Journalists, Too](#), that includes a reference to one of my all-time OSINT favorites, [Maltego](#). [Joseph Cox](#)'s article also mentions Datasploit, a 2016 favorite for fellow tools aficionado, [Toolswatch.org](#), see [2016 Top Security Tools as Voted by ToolsWatch.org Readers](#). Having not yet explored Datasploit myself, this proved to be a grand case of "no time like the present."

Datasploit is "an #OSINT Framework to perform various recon techniques, aggregate all the raw data, and give data in multiple formats." More specifically, as stated on [Datasploit documentation page](#) under [Why Datasploit](#), it utilizes various Open

Datasploit documentation page under [why Datasploit](#), it utilizes various Open Source Intelligence (OSINT) tools and techniques found to be effective and brings them together to correlate the raw data captured, providing the user relevant information about domains, email address, phone numbers, personal data, etc. Datasploit is useful to collect relevant information about a target in order to expand your attack and defense surface very quickly.

The feature list includes:

- Automated OSINT on domain/email/username/phone for relevant information from different sources.
- Useful for penetration testers, cyber investigators, defensive security professionals, etc.
- Correlates and collaborate results, shows them in a consolidated manner.
- Tries to find out credentials, API keys, tokens, subdomains, domain history, legacy portals, and more as related to the target.
- Available as a single consolidating tool as well as standalone scripts.
- Performs Active Scans on collected data.
- Generates HTML and JSON reports along with text files.

## Resources

GitHub: <https://github.com/dataspl0it/dataspl0it>

Documentation: <http://dataspl0it.readthedocs.io/en/latest/>

YouTube: [Quick guide to installation and use](#)

youtube: [Quick guide to installation and use](#)

## Pointers

Second, a few pointers to keep you from losing your mind. This project is very much a work in progress, with lots of very frustrated users filing bugs and wondering where the support is. The team is doing their best, be patient with them, but read through the GitHub [issues](#) to be sure any bugs you run into haven't already been addressed.

- 1) Datasplit does *not* error gracefully, it just crashes. This can be the result of unmet dependencies or even a missing API key. Do not despair, take note, I'll talk you through it.
- 2) I suggest, for ease, and best match to documentation, run Datasplit from an Ubuntu variant. Your best bet is to grab [Kali](#), VM, or dedicated and load it up there, as I did.
- 3) My installation guidance and recommendations should hopefully get you running trouble free, follow it explicitly.
- 4) Acquire as many API keys as possible, see further detail below.

## Installation and Preparation

From Kali bash prompt, in this order:

1. `git clone https://github.com/datasplit/datasplit /etc/datasplit`
2. `apt-get install libxml2-dev libxslt-dev python-dev lib32z1-dev zlib1g-dev`

3. `cd /etc/datasploit`
4. `pip install -r requirements.txt`
5. `mv config_sample.py config.py`
6. With your preferred editor, open `config.py` and add API keys for the following, at a minimum. They are for all intents and purposes required, and detailed instructions to acquire each is [here](#):
  1. Shodan API
  2. Censys ID and Secret
  3. Clearbit API
  4. Emailhunter API
  5. Fullcontact API
  6. Google Custom Search Engine API key and CX ID
  7. Zoomeye Username and Password

If, and only if, you've done all of this correctly, you might end up with a running instance of Datasploit. Seriously, this is some of the glitchiest software I've tussled with in quite a while, but the results paid handsomely. Run `python datasploit.py domain.com`, where `domain.com` is your target. Obviously, I ran `python datasploit.py holisticinfosec.org` to acquire results pertinent to your author.

Datasploit rapidly pulled results as follows:

211 domain references from Github:

```

----> Searching Github for domain results
[+] Found 211 results on github.
Top 30 results shown below
01. File: https://github.com/calh/Maltego-pyCSV/blob/064acc3851b44cc543392e1a4491ea464239de86/README.md
   Owner: calh/Maltego-pyCSV
   Repository: https://github.com/calh/Maltego-pyCSV
02. File: https://github.com/holisticinfosec/toolsmith_R/blob/42ad034a1916c866d1719ae7b5e9ce7ed3bb7b0e/README.md
   Owner: holisticinfosec/toolsmith_R
   Repository: https://github.com/holisticinfosec/toolsmith_R
03. File: https://github.com/CERTCC-Vulnerability-Analysis/Vulnerability-Data-Archive/blob/917c6ad3c663997a4776754b182e87d
   /vuln/166730/vendor-mmap-7axxed.json
   Owner: CERTCC-Vulnerability-Analysis/Vulnerability-Data-Archive
   Repository: https://github.com/CERTCC-Vulnerability-Analysis/Vulnerability-Data-Archive
04. File: https://github.com/aboutsecurity/Bro-samples/blob/88f66921f938237b07c8676d5876c2e99cc17378/README.md
   Owner: aboutsecurity/Bro-samples
   Repository: https://github.com/aboutsecurity/Bro-samples
05. File: https://github.com/recuweb-data/en_categories/blob/c7d2cde92c3f2ee48ba1b569519b5a218a26f6bf/_hun/hunting/latest_
   Owner: recuweb-data/en_categories
   Repository: https://github.com/recuweb-data/en_categories
06. File: https://github.com/Resistor52/processpcaps/blob/098d7da2faa8e63e3ea5629916e63b0f3279a5a/downloadSamplePCAPS.sh
   Owner: Resistor52/processpcaps
   Repository: https://github.com/Resistor52/processpcaps
07. File: https://github.com/SAFETAG/SAFETAG/blob/73b80e7a43ca1353872073becf30ab46a9192fcc/en/references/recon-ng.md
   Owner: SAFETAG/SAFETAG
   Repository: https://github.com/SAFETAG/SAFETAG

```

### GitHub results

Luckily, no results from Shodan.

Four results from Paste(s):

```

----> Finding Paste(s)..
https://www.googleapis.com/customsearch/v1?key=A1zaSyC24lr0sIAAc9ERj8lF3aUyFZiW3npHdn04cx=013584890790507044394
sec.org?&start=1
[+] 4 results found

Title: SPA: Single Packet Authorization - The Ghost in the Machine [PDF] ...
URL: https://www.reddit.com/r/netsec/comments/am72w/spa_single_packet_authorization_the_ghost_in_the/
Snippet: Jan 6, 2010 ... SPA: Single Packet Authorization - The Ghost in the Machine [PDF] (
holisticinfosec.org). submitted 7 years ago by sanitybit · 1 comment; share.

Title: El Jefe: The Boss Will See You Now [PDF] : netsec
URL: https://www.reddit.com/r/netsec/comments/fdxqa/el_jefe_the_boss_will_see_you_now_pdf/
Snippet: Feb 2, 2011 ... El Jefe: The Boss Will See You Now [PDF] (holisticinfosec.org). submitted 6 years
ago by dguide · comment; share; report. no comments (yet).

Title: message to the world I AM PH0K3 the egypt gov is spying on us i ...
URL: https://pastebin.com/906K7WAJ
Snippet: Oct 13, 2015 ... russ@holisticinfosec.org. linux@linux.cz. sales@woodlandmanufacturing.com.
doaa_nader@hotmail.com. ashraf@ntra.gov.eg.

Title: A drug is introduced which permanently makes you 50% happier but ...
URL: https://pastebin.com/2346wvRY
Snippet: Sep 30, 2012 ... scarybeasts/status/32941540531765248. permalinkcontextfull comments. 7. El
Jefe: The Boss Will See You Now [PDF] (holisticinfosec.org).

```

### Pastebin and Pastie results

Datasploit pulled russ at holisticinfosec dot org as expected, per email harvesting.

## Accurate HolisticInfoSec host location data from Zoomeye:

```

-->> Finding hosts from ZoomEye
2. https://www.holisticinfocsec.org/
IP: [u'70.40.190.70']
Site: www.holisticinfocsec.org
Title: 302 found
Headers: HTTP/1.1 302 found
Server: nginx/1.13.8
Date: Sun, 26 Jul 2017 17:19:51 GMT
Content-Type: text/html; charset=iso-8859-1
Content-Length: 209
Connection: keep-alive
Location: https://www.holisticinfocsec.org/

Location: {u'city': {u'geoname_id': 5780026, u'names': {u'zh-CN': u'乌兹别克斯坦首都', u'en': u'Provo'}}, u'country': {u'geoname_id': 6252001, u'code': u'US', u'names': {u'zh-CN': u'美国', u'en': u'United States'}}, u'isp': u'Unified Layer', u'asn': 48800, u'subdivisi
on': {u'geoname_id': 5549030, u'code': u'UT', u'names': {u'zh-CN': u'犹他州', u'en': u'Utah'}}, u'location': {u'lat': 40.2188,
u'lon': -111.6138}, u'organization': u'Unified Layer', u'asn': u'Unified Layer', u'continent': {u'geoname_id': 6255049, u'code': u'NA', u'
names': {u'zh-CN': u'北美洲', u'en': u'North America'}}}

```

Details regarding HolisticInfoSec sub-domains and page links:

```

----> Finding subdomains, will be back soon with list.
sub access token= 40c630c3c002c49741c3611913f0d3092fdee9"
List of subdomains found: 4e0c-b810-0e5501ds1806"
sysio id= 00000000-d800-4000-910f-a65a315dca61"
holisticinfosec.org 0157d0f5k1fg7wFJK03xwFclp9"
www.holisticinfosec.org
e-cm-0000-0000-001A4c0ERJ81F3edYfZiN3ngMmQ"
----> Finding Pagelinks: 70507044396-zwsixu3qjw"
[OK] done

in Visible links
https://holisticinfosec.org/
https://holisticinfosec.org/df/news
https://holisticinfosec.org/df/simplicity
https://holisticinfosec.org/df/toolsmith
https://holisticinfosec.org/
https://holisticinfosec.org/images/logo.png
http://holisticinfosec.blogspot.com/search?q=toolsmith&max-results=20&by-date=true
http://holisticinfosec.blogspot.com/search?q=toolsmith&max-results=20&by-date=true
https://holisticinfosec.org/df/toolsmith#00
http://holisticinfosec.blogspot.com/search?q=toolsmith&max-results=20&by-date=true
http://pico.dev7studios.com/ 0157d0f5k1fg7wFJK03xwFclp9"
http://github.com/zohma/plumbing/ 0157d0f5k1fg7wFJK03xwFclp9"
https://twitter.com/holisticinfosec"
https://github.com/0157d0f5k1fg7wFJK03xwFclp9"

in Hidden links: 4e0c-b810-0e5501ds1806"
https://holisticinfosec.org/advisories/hio-2009-0305-e107-multiple-e107-admin-csrf-a-xss-vulnerabilities
https://holisticinfosec.org/advisories/hio-2010-0223-web-wiz-forens-csrf-vulnerabilities
https://holisticinfosec.org/bestpractices
https://holisticinfosec.org/events
https://holisticinfosec.org/iathenevs

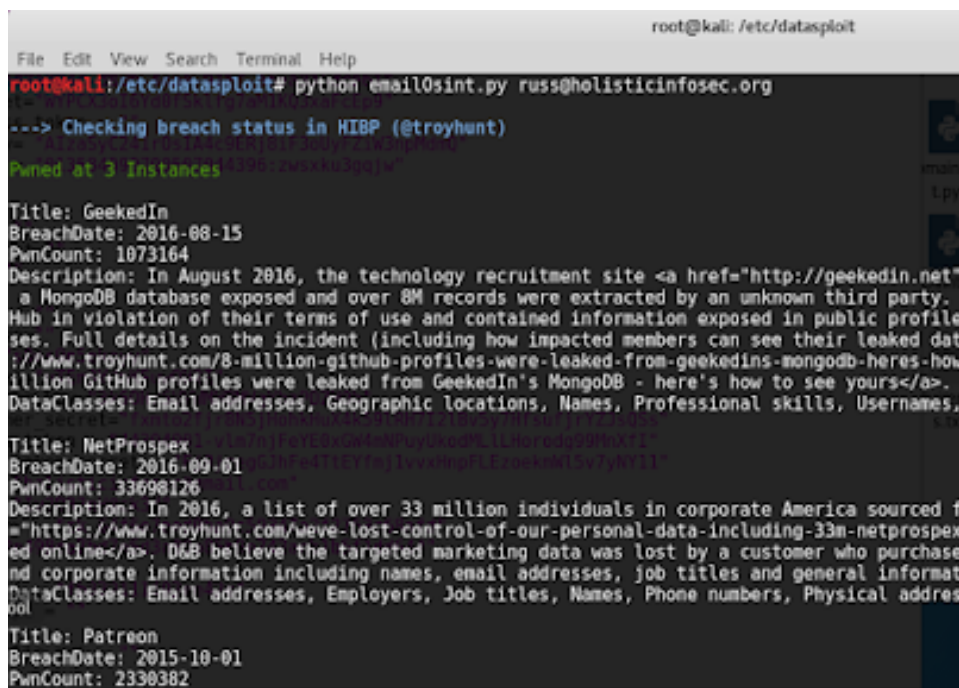
```

```
https://holisticinfosec.org/presentations
https://holisticinfosec.org/publications
https://holisticinfosec.org/templates
https://holisticinfosec.org/vulns
```

### *Sub-domains and page links*

Finally, a good return on DNS records for [holisticinfosec.org](https://holisticinfosec.org) and, thankfully, no vulnerabilities found via [PunkSpider](#).

Datasploit can also be integrated into other code and called as individual scripts for unique functions. I did a quick run with `python emailOsint.py` `russ@holisticinfosec.org` and the results were impressive:



```
root@kali: /etc/datasploit
File Edit View Search Terminal Help
root@kali:/etc/datasploit# python emailOsint.py russ@holisticinfosec.org
---> Checking breach status in HIBP (@troyhunt)
Pwned at 3 Instances 1398:2wxxk03gqjw
Title: GeekedIn
BreachDate: 2016-08-15
PwnCount: 1073164
Description: In August 2016, the technology recruitment site <a href="http://geekedin.net">a MongoDB database exposed and over 8M records were extracted by an unknown third party. Hub in violation of their terms of use and contained information exposed in public profiles. Full details on the incident (including how impacted members can see their leaked data) is available at: <a href="http://www.troyhunt.com/8-million-github-profiles-were-leaked-from-geekedins-mongodb-heres-how">http://www.troyhunt.com/8-million-github-profiles-were-leaked-from-geekedins-mongodb-heres-how</a>. 8 million GitHub profiles were leaked from GeekedIn's MongoDB - here's how to see yours</a>.
DataClasses: Email addresses, Geographic locations, Names, Professional skills, Usernames,
Title: NetProspex
BreachDate: 2016-09-01
PwnCount: 33698126
Description: In 2016, a list of over 33 million individuals in corporate America sourced from a database was exposed online. D&B believe the targeted marketing data was lost by a customer who purchased corporate information including names, email addresses, job titles and general information.
DataClasses: Email addresses, Employers, Job titles, Names, Phone numbers, Physical addresses
Title: Patreon
BreachDate: 2015-10-01
PwnCount: 2330382
```

### *Email OSINT*

I love that the first query is of [Troy Hunt's Have I Been Pwned](#). Not sure if you have

been? Better check it out. A reminder here, you'll really want to be sure to have as many API keys as possible or you may find these buggy scripts crashing. You'll definitely find yourself compromising between frustration and the rapid, detailed results. I put this offering squarely in the "shows much promise category" if the devs keep focused on it, assesses for quality, and handles errors better.

Give Datasploit a try for sure.

Cheers, until next time.

---

## Like This Article? Read More From DZone



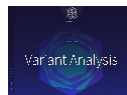
**The Difference Between  
Vulnerability Assessment and  
Penetration Testing**



**Dependencies: It's Not Just Your  
Code You Need to Secure**



**Introduction to Pentesting**



**Free DZone Refcard  
Variant Analysis**

Topics: [PENETRATION TESTING](#) , [SECURITY](#) , [VULNERABILITIES](#)

 **Like (2)**    **Comment (0)**    **Save**    **Tweet**

 **3,931 Views**

Published at DZone with permission of Russ McRee , DZone MVB. [See the original article here.](#)



Opinions expressed by DZone contributors are their own.



# Security Partner Resources

IN PROGRESS