CeWL (kul)

A guide by Harry Prosser

What is CeWL?

- Custom Wordlist generator
- Output can be used with other programs
  - John the Ripper
  - Hashcat
  - Hydra
  - Burpsuite

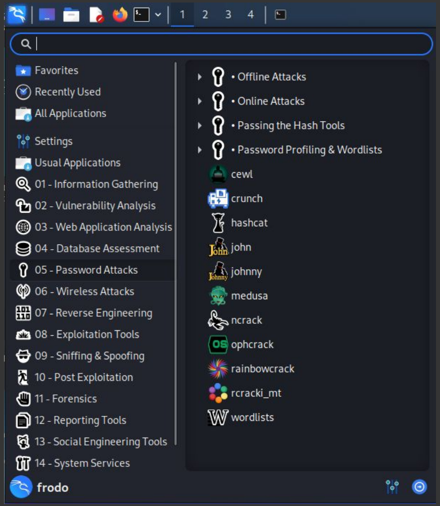
Background

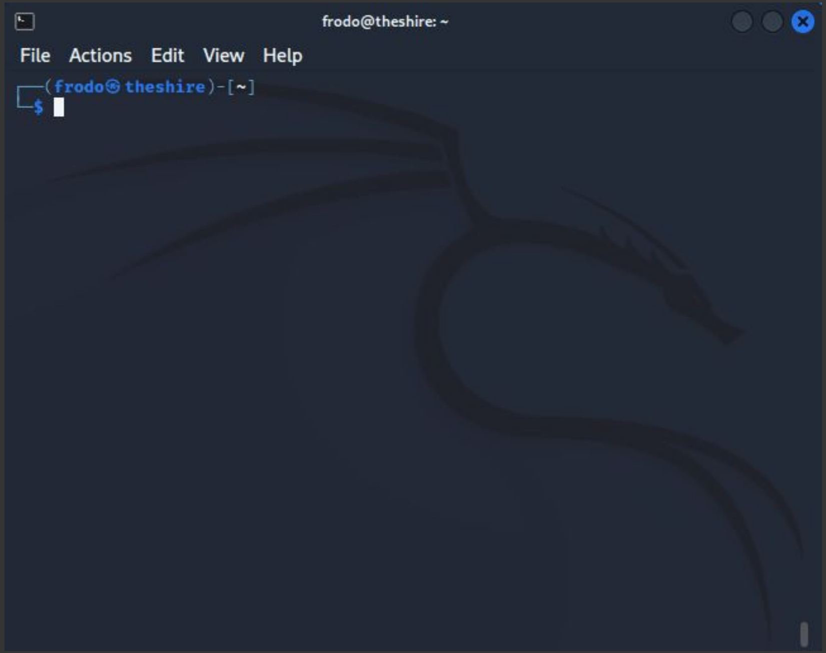
- ➔ **Ruby program**
- ➔ **Created by Robin Wood**

  Aka digininja
  https://github.com/digininja
  Founder of STEELCON
- ➔ **CeWL's function**

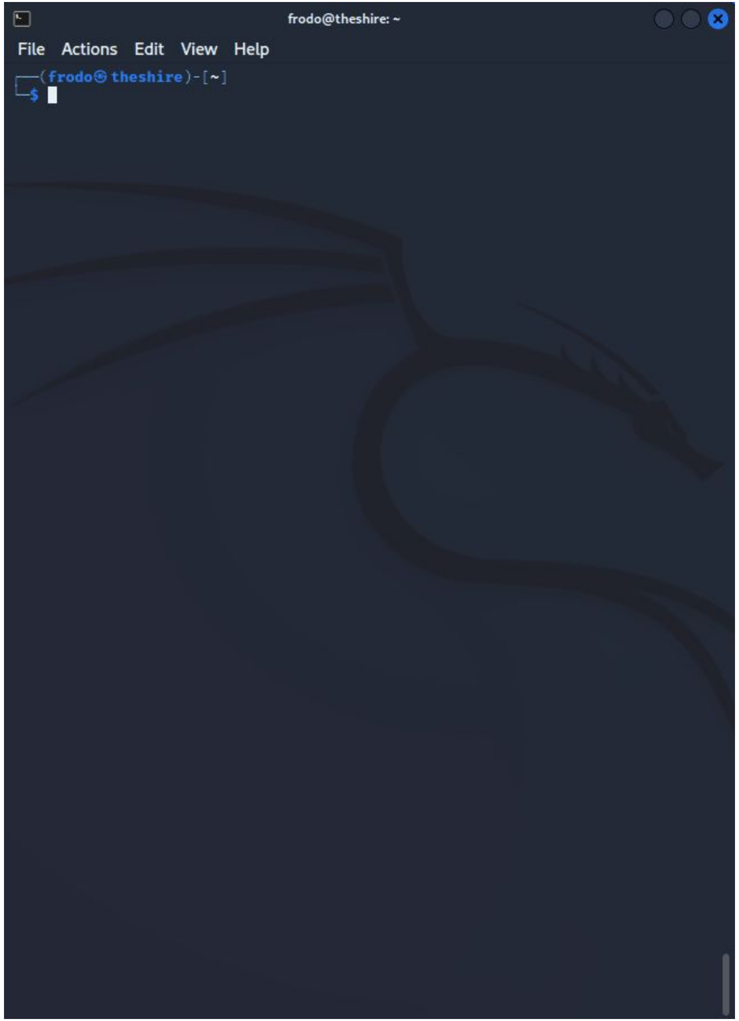
  Cewl **actively** spiders a given URL to a specified depth, optionally following external links, and returns a list of words which can then be used

How to CeWL





```
┌──(frodo㉿theshire)-[~]
└─$ 
```



```
┌──(frodo㉿theshire)-[~]
└─$ 
```

CeWL Common Options

-h | shows help file

-d | spider depth

-m | minimum word length

-o | lets spider follow external links

-w | write out to file (>or>>)

--with-numbers |include integers

-a | include meta data

-e | include email addresses

-c | count for each found word

-v | verbose

Usage

- **Is it legal?**

  Crawls public websites

  Follows links

- **Be Careful**

  Creating the wordlist = **OKAY**

  Using the wordlist = **MAYBE NOT**

- **Ruby file can be found at:**

  <user>/usr/bin/cewl

  use: dpkg-query -L cewl

Disclaimer!

- Do you have written consent?
- Is the target URL protected?
- What are you using the output for?
- Can using CeWL get me in trouble?

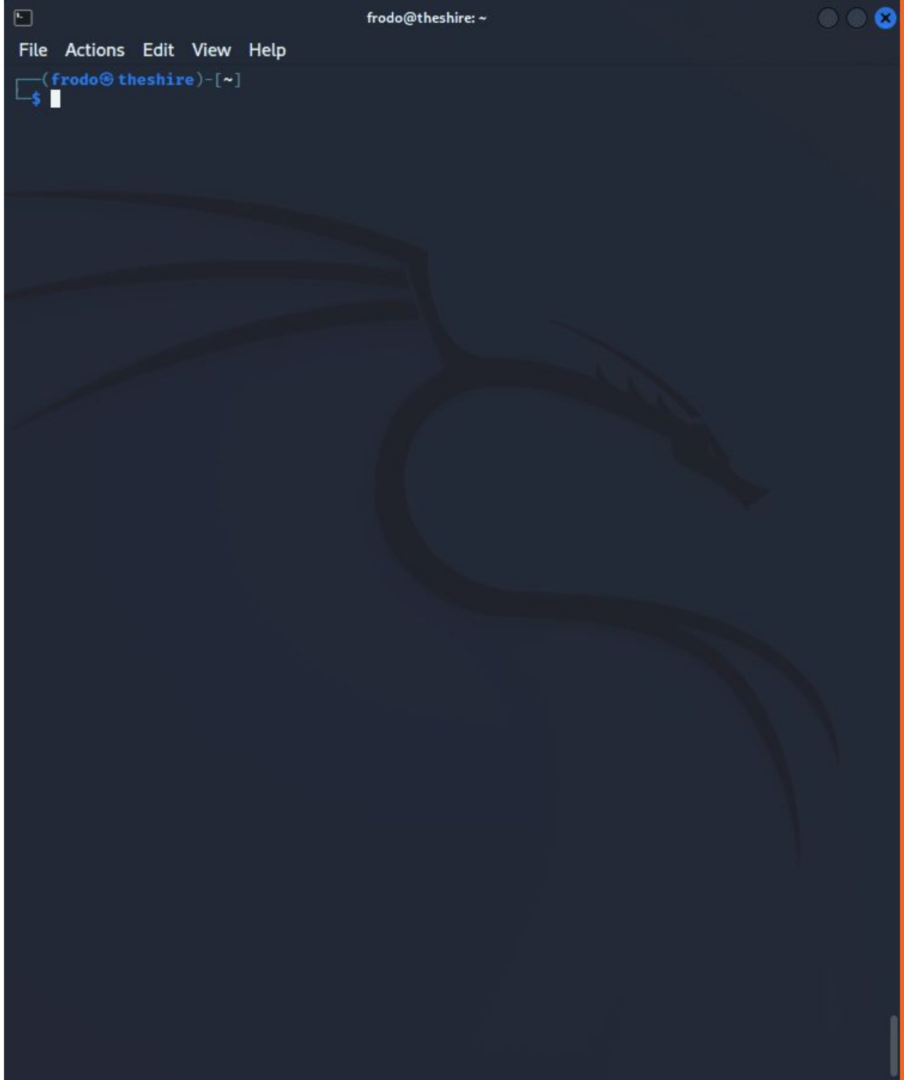
#W hat W ould J amie D o

CeWL

Megacorpone

the
MegaCorp
One
for
row
and
container
that
technology
megacorpone
com
Our
this
Email
has
nanotechnology
Bootstrap
core
SUPPORT
CAREERS
LOG
More
Info
are
Contact
any
company
About
our
CONTACT
edge
with
been
Joe
Sheer
This
Just
ABOUT
experience
have
Systems
Based
you
Security
Old
Mill
Rachel
United
States
end
col
CSS
Custom
styles
template
debugging
purposes
Don
actually
copy
line
Fixed
navbar
Toggle
navigation
HOME
nav
collapse
new
some
being
What
behind
can
FOOTER
All
rights
reserved
fictitious
brought
Offensive
Social
Media
Location
footerwrap
JavaScript
Placed
document
pages
load
faster
future
years
bleeding
technologies
available
working
into
smart
weapons
products
CEO
Not
Found
not
WRAP
Name
must
view
wrap
Twitter
Nanotechnology
Future
opportunities
offer
research
computer
now
way
service
Why
Work
With
advanced
deem
impossible
engineered
provide
cutting
decades
ahead
competition
used
nanomedicines
cell
regeneration
military
applications
twrap
MIDDLE
CONTENT
Services
Cell
Regeneration
Immune
Supplements
Micromachine
Cyberisation
Repair
Nanomite
Weaponry
Nanoprobe
Entity
Assimilation
Perlin
VanHook
Chemical
Dispersal
FAQ
ethics

etc.



```
┌──(frodo㉿theshire)-[~]
└─$ 
```

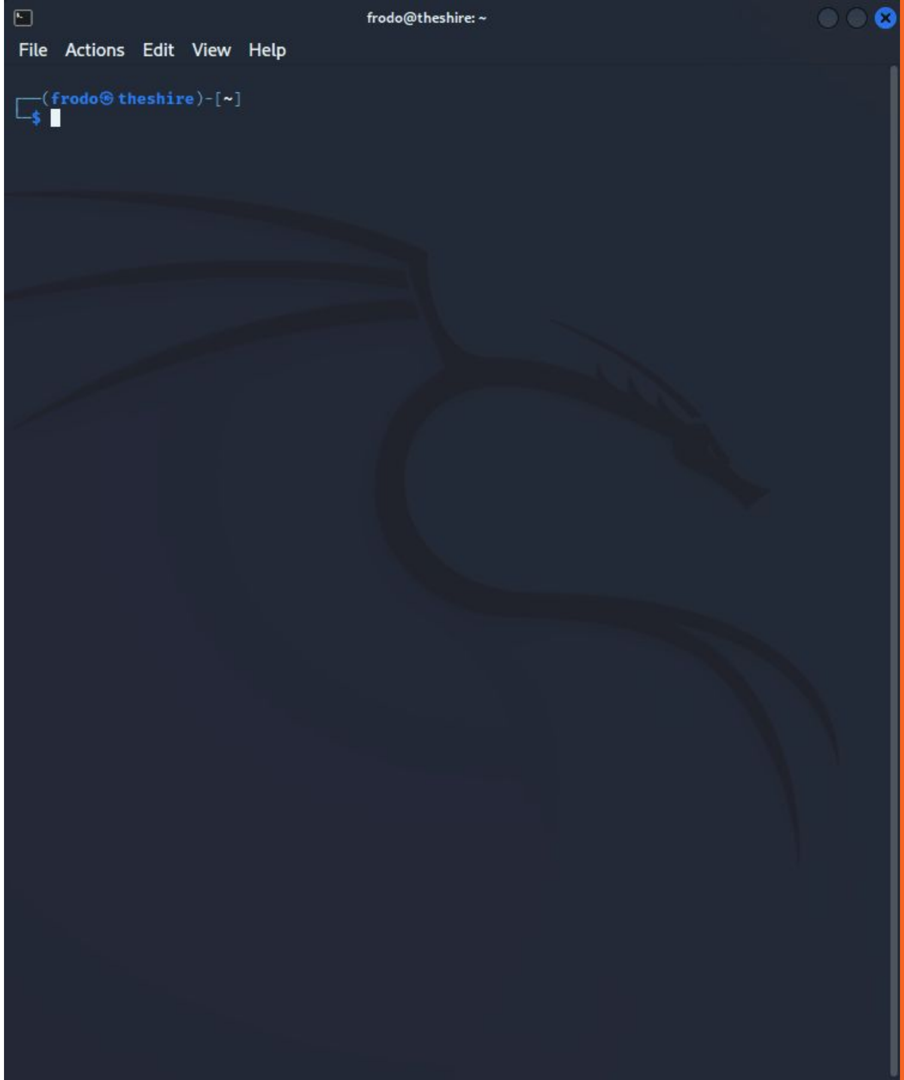
CeWL

Megacorpone

COUNT

the	the, 55
MegaCorp	MegaCorp, 48
One	One, 48
for	for, 28
row	row, 20
and	and, 20
container	container, 20
that	that, 18
technology	technology, 15
megacorpone	megacorpone, 14
com	com, 14
Our	Our, 13
this	this, 12
Email	Email, 12
has	has, 11
nanotechnology	nanotechnology, 11
Bootstrap	Bootstrap, 10
core	core, 10
SUPPORT	SUPPORT, 10
CAREERS	CAREERS, 10
LOG	LOG, 10
More	More, 10
Info	Info, 10
are	are, 10
Contact	Contact, 10
any	any, 9
company	company, 9
About	About, 8
our	our, 8
CONTACT	CONTACT, 7
edge	edge, 7
with	with, 7
been	been, 7
Joe	Joe, 7
Sheer	Sheer, 7
This	This, 7
Just	Just, 6
ABOUT	ABOUT, 6
experience	experience, 6
have	have, 6

etc.

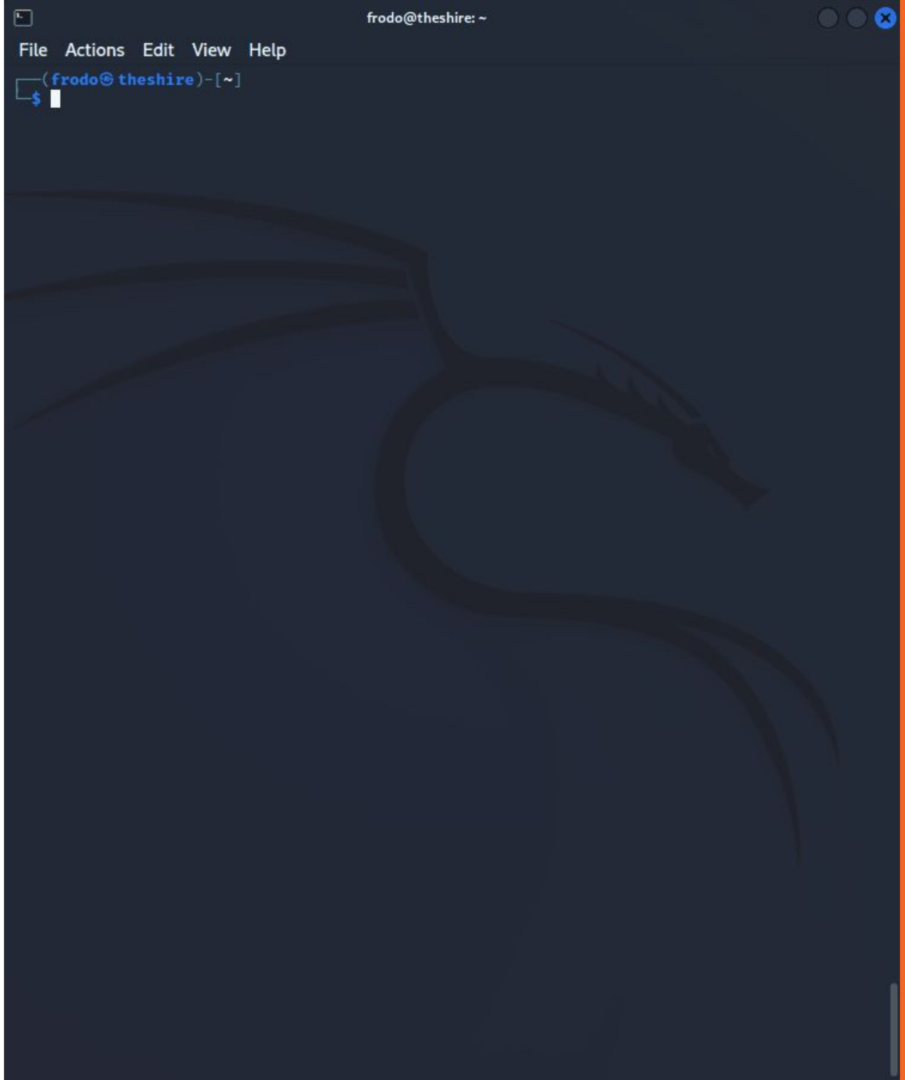


```
┌──(frodo㉿theshire)-[~]
└─$ 
```

Let's use some OPTIONS

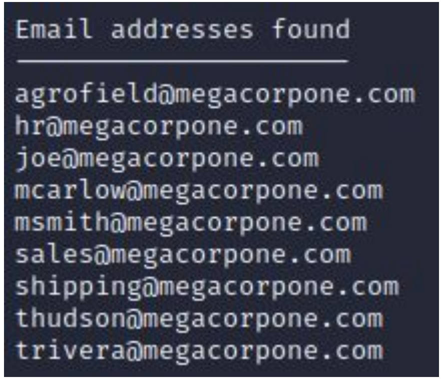
-d 3	crawls 3 pages deep
-m 8	minimum length of 8 characters
-e	prints email addresses
-v	verbose, tells you source info

cewl -d 3 -m 8 -e -v http://www.megacorpone.com

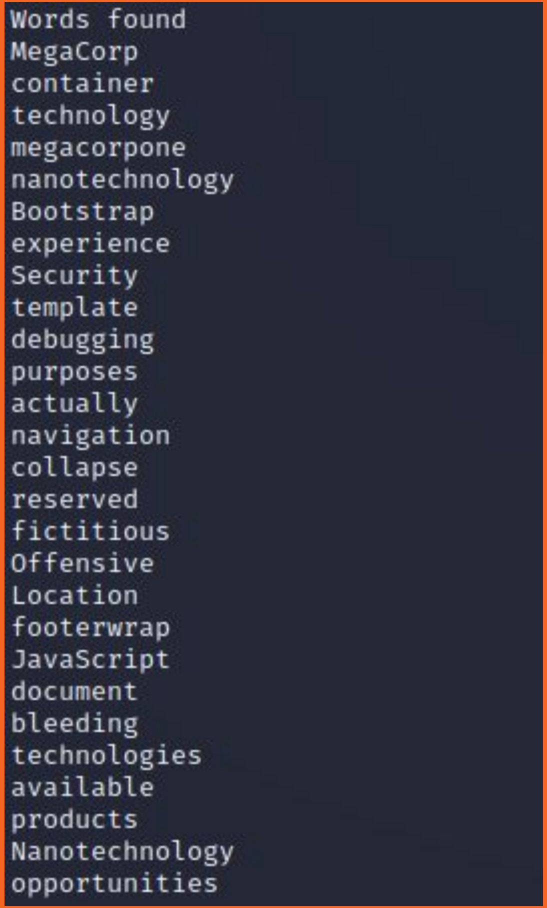


```
┌──(frodo㉿theshire)-[~]
└─$ 
```

What did we FIND

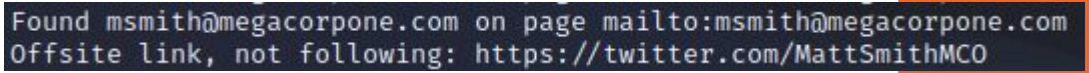


```
Email addresses found
---------------------
agrofield@megacorpone.com
hr@megacorpone.com
joe@megacorpone.com
mcarlow@megacorpone.com
msmith@megacorpone.com
sales@megacorpone.com
shipping@megacorpone.com
thudson@megacorpone.com
trivera@megacorpone.com
```

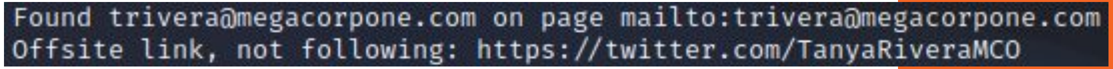


```
Words found
MegaCorp
container
technology
megacorpone
nanotechnology
Bootstrap
experience
Security
template
debugging
purposes
actually
navigation
collapse
reserved
fictitious
Offensive
Location
footerwrap
JavaScript
document
bleeding
technologies
available
products
Nanotechnology
opportunities
```

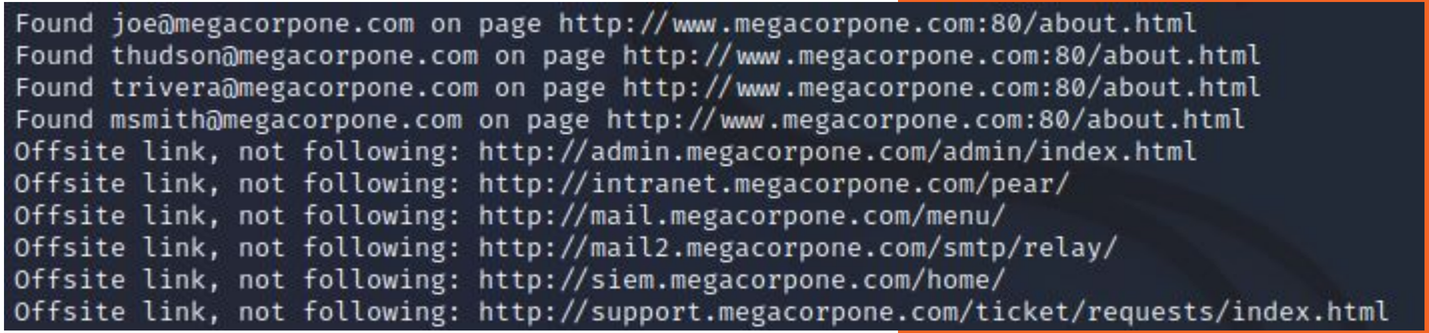
-v | verbose output



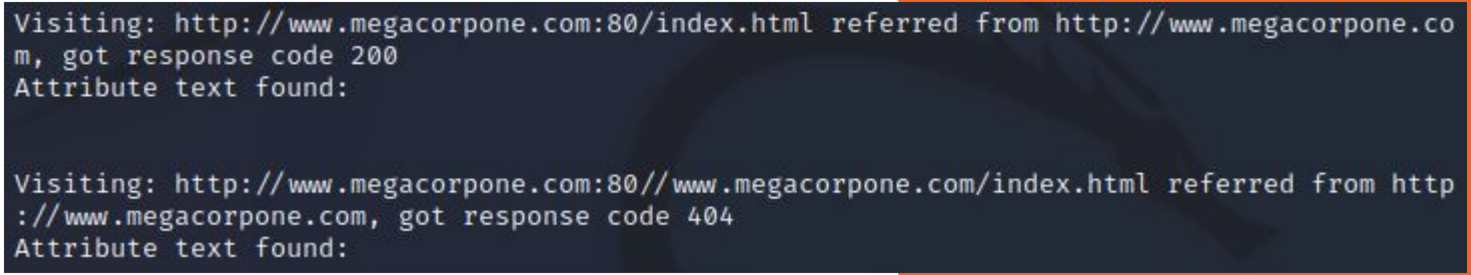
```
Found msmith@megacorpone.com on page mailto:msmith@megacorpone.com
Offsite link, not following: https://twitter.com/MattSmithMCO
```



```
Found trivera@megacorpone.com on page mailto:trivera@megacorpone.com
Offsite link, not following: https://twitter.com/TanyaRiveraMCO
```



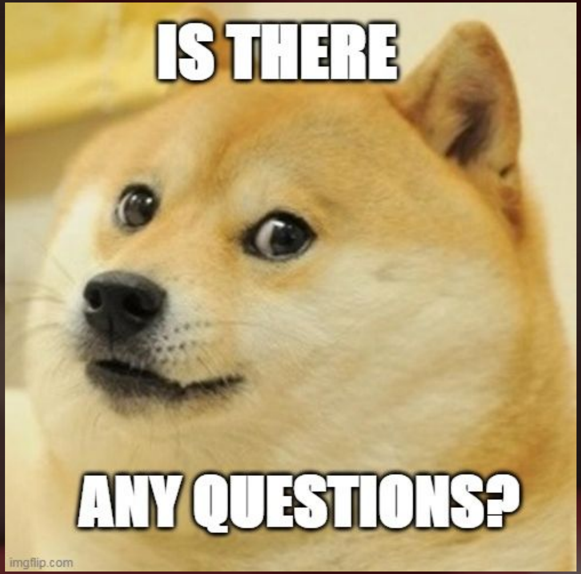
```
Found joe@megacorpone.com on page http://www.megacorpone.com:80/about.html
Found thudson@megacorpone.com on page http://www.megacorpone.com:80/about.html
Found trivera@megacorpone.com on page http://www.megacorpone.com:80/about.html
Found msmith@megacorpone.com on page http://www.megacorpone.com:80/about.html
Offsite link, not following: http://admin.megacorpone.com/admin/index.html
Offsite link, not following: http://intranet.megacorpone.com/pear/
Offsite link, not following: http://mail.megacorpone.com/menu/
Offsite link, not following: http://mail2.megacorpone.com/smtp/relay/
Offsite link, not following: http://siem.megacorpone.com/home/
Offsite link, not following: http://support.megacorpone.com/ticket/requests/index.html
```



```
Visiting: http://www.megacorpone.com:80/index.html referred from http://www.megacorpone.co
m, got response code 200
Attribute text found:


Visiting: http://www.megacorpone.com:80//www.megacorpone.com/index.html referred from http
://www.megacorpone.com, got response code 404
Attribute text found:
```

Question Time



Other great Custom Wordlist Generators

User Name Generator | https://github.com/therodri2/username_generator.git

Creates first and last name combinations

john smith > jsmith / johnsmith / smithjohn / sjohn / j.smith / j-smith

CRUNCH | also in KALI 05 - Password Attacks Menu or use apt database

Creates combinations of given character set

123 > 111 / 112 / 113 / 121 / 122 / 123 / 131 / 132 / 133 / 211 / 212 / 213 etc.

CUPP Common User Password Profiler | https://github.com/Mebus/cupp.git

Uses details about a specific target, such as their birthdate, pet name, company name, etc. Has an interactive command prompt

YO DAWG I HEAR YOU LIKE LISTS
SO WE PUT YOUR LISTS IN A LIST SO
YOU CAN LIST WHILE YOU LIST

What you have learned

What a Custom Wordlist is.

A wordlist is a file (a text file in most cases but not limited to it) that contains a set of values that the attacker requires to provide to test a mechanism. A custom one is one that is built to be specific to a target.

Basic CeWL commands

```
cewl -h
```

New Resources

CeWL

Username Generator

Crunch

CUPP

Contact Details

https://www.linkedin.com/in/harryprosser/