

intitle:"Google Hacking"

Presented by Robert Vinson
security@uiowa.edu

Because having direction is good...

- What - is Google (hacking)?
- Why/When - do we care?
- How - can we find “stuff”?
- Where - do we come in?

Google

- Google
 - It was created by two guys.
 - They have lots of money now.
 - Motto: Do no evil.
 - Goal: “Organize the world’s information and make it universally accessible and useful”

Hacking?

- Google hacking **is not** hacking Google.
- Google hacking **is** using Google in creative ways to find nifty tidbits.

We care when:

- Google can be used to compromise the security of:
 - An establishment (i.e. our university)
 - An individual

Google operators

- Used to make searches less ambiguous
- Some of the more useful operators:
 - site (e.g. site:uiowa.edu)
 - intitle/allintitle
 - inurl
 - filetype

Searching strategy

- Search for phrases where possible.
- Use advanced operators to your advantage.
- Make searches as specific as possible to narrow results.
 - If the search is too specific. Try using a more generic search, and then refine it.

Be good!

- The information in the following searches, and from Google hacking in general, has the possibility of being used for malicious purposes. This demonstration is delivered for illustrative purposes, not as a way of enabling illegal and/or harmful actions. However, it is our hope that this demonstration enables administrators to locate and resolve insecurities in their environments.

Threats to individuals - examples

- [resume OR vitae filetype:doc "social security number" 000000000..999999999](#)
- [inurl:customers.xls](#)

Threats to establishments -examples

- [intext:"Tobias Oetiker" "traffic analysis" site:edu](#)
- [filetype:log site:edu "set password for"](#)
- [filetype:config OR filetype:conf site:edu - Google Search](#)

And, because Jason Alexander
went to Iowa State...

- [site:iastate.edu intitle:"index of" modified](https://www.iastate.edu/index.html)

Creepy Crawlers: Worms and Spiders

- There has already been a worm that harvested email address from google searches in order to spread.
- A program could query for server specific messages to search for vulnerable servers.

Creepy Crawlers (cont.)

- A program could search for user information, and save results that seem relevant for later review by an identity thief.
- One could even enumerate servers in a domain by doing a `site:domain.com` search and parsing URLs for server names.

Protecting Ourselves

- Do not enter personal information in public areas.
- Turn off directory listings!!!
- Change default server error strings/replies and program names.
- Use a robots.txt file.

Prevent Directory Browsing - IIS

- Include “index.html” in the directory.
- IIS – Turn off/manage Directory Browsing
 - <http://support.microsoft.com/kb/313075/EN-US/>

Prevent Directory Listings - Apache

- Apache

- <Directory /somedir>

- Options -Indexes

- </Directory>

- Use .htaccess for individual directories.

- <http://httpd.apache.org/>

robots.txt file

- A robots.txt file is a way to keep search engines' spiders from indexing specified parts of a site.

User-agent: *

Disallow: /directory/

– <http://www.robotstxt.org>

Changing defaults

- Change the default filename of applications if plausible.
- Consider using `mod_headers` with Apache or `IISLockDown` with IIS to change default banners.
- Consider changing default error pages.

Future threats?

- More intelligent/devious programs designed to harvest information?
- Combining the power of facial recognition software and Google's image search?
- Using maps.google.com to get a visual of a person's home.

Resources

- “Dangerous Google – Searching for Secrets” – www.hakin9.org/en
- Google Hacking for Penetration Testers – books 24x7
www.lib.uiowa.edu
- <http://johnny.ihackstuff.com/>
- www.google.com

Questions?

security@uiowa.edu