

*A hacker's view on the
Internet, privacy, etc.*

Virgil Griffith

Disruptive Technologist


virgil@caltech.edu

<http://virgil.gr>





Talk Outline

- What I do
 - All about WikiScanner
 - A Hacker's view of...
 - Wikipedia
 - Privacy
 - Social Networks
 - Nifty tricks to compromise privacy
- 

What I do

- Fun and games with data that have socially interesting ramifications.
- Examples
 - ◎ [Derive Mother's Maiden Names from public records](#)
 - ◎ [WikiScanner](#)
 - ◎ [Booksthatmakeyoudumb](#)
 - ◎ [FreefoodatCaltech](#) / [MIT](#)
 - ◎ [Planet Sony](#) (Dan Kaminsky)

WikiScanner - How it works

- 1) Download **ALL** of Wikipedia, getting all of the anonymous edits (Free from Wikimedia.)
 - ⊙ Found 34.5M anonymous edits, ~21% of Wikipedia

- 2) Tracing is hard, so buy a database of what organizations own which IP addresses (available from private corporations. \$1,000)
 - ⊙ 2,668,095 different orgs in database.

- 3) Merge them together!

What you can do now

- Type an organization's name and see all of the anonymous edits that came from their local network.
 - ◎ Found 187,529 different orgs with at least 1 edit.
- See what organizations have edited a particular Wikipedia page.
- [Vote on the interesting stuff](#)

WikiScanner - The harvest

- Different % of anonymous edits by country
- Yes, the CIA does in fact edit Wikipedia [\[1\]](#) [\[2\]](#)
- FOIA lawsuit filed over Mike Huckabee white-washing . [\[1\]](#)
- Dutch princess white-washes connections to drug baron.
- Politicians do in fact hire staff to police their pages. [\[1\]](#)
 - So do corporations. A lot. [\[1\]](#)

Hacker's view of Wikipedia

- WP is fundamentally Darwinian
 - ⊙ People randomly inject crap into WP pages
 - ⊙ Crap people disagree with is removed
 - ⊙ What sticks around is what people agree with
- Disinformation techniques on WP is a preview of what is to come.

Complete openness good for Wikipedia?

- WP's openness stems from desire for Katamari Damacy-like growth
 - ⊙ They're encyclopedia builders, not philosophers.
 - ⊙ Locking down a page is function of (%useful injections to page, desire to increase pagelength)
- If WP locked down pages, its growth would nigh flat line.

Hacker's view of privacy

- Technology switches between concealing and revealing people.
 - Oscillating cat and mouse game
 - One has upper hand at any given moment.
- With massive amounts of data being accrued and indexed, revealing people seems stronger right now.

Social Networks

- People give out huge amounts of unnecessary info on social networks
 - ⊙ No one is sure why.
- However, no one has yet abused this on a large scale
 - ⊙ No one is sure why.
- Most think it's only time, but not sure what the attacker could do.

Nifty Tricks

● Wikipedia

- IP# disclosure of registered users 

- Detecting duplicate accounts owned by one person

● WikiScanner 2.0 – Automatic detection of vanity edits

- Approximate name match

- Trademark name match

- Location match to block coffee-shop attack

Nifty Tricks II


- Mining your webhistory
 - ◎ New use -- Auto-Blackmailer?
- DomainTools
 - ◎ Archive.org for WHOIS information (demo)
 - ◎ Reverse-lookups now possible! (demo)
- Visualizing Speeches
- Revealing Redacted PDFs (demo)
- Indexing MSWord Metadata

Nifty Tricks: Protecting yourself

- Tor – <http://tor.eff.org>
 - ◎ The finest IP# protector money can buy
 - ◎ Convenient that it's free.
 - ◎ Hosting websites within Tor
- Open WI-FI hot-spots
- Changing your MAC Address
 - ◎ Easiest: connect via a home wireless box and its web-interface lets you change it's MAC.



Policy Question

- What to do when users have an expectation of some privacy but most aren't private at all while a small number are completely anonymous?
- 

Questions?



Anonymous edit to the page "Tea Tree Oil"

Talk:Tea tree oil

From Wikipedia, the free encyclopedia

(Difference between revisions)

Revision as of 18:06, 16 December 2006 (edit)

[Siraj555](#) (Talk | contribs | block)

(→*Rename a good thing?*)

← Older edit

Revision as of 17:16, 21 December 2006 (edit)

(undo)

[12.152.168.248](#) (Talk | block)

(→*Rename a good thing?*)

Newer edit →

Line 170:

I agree with "Tea Tree Oil (Melaleuca Oil)" as well.
Anyone else to chime in? [\[\[User:Siraj88|Siraj88\]\]](#) 23:40,
14 December 2006 (UTC)

Line 170:

I agree with "Tea Tree Oil (Melaleuca Oil)" as well.
Anyone else to chime in? [\[\[User:Siraj88|Siraj88\]\]](#) 23:40,
14 December 2006 (UTC)
+
It's been a week. I'll change it from the
+ current "Melaleuca Tea tree oil" to "Tea Tree Oil
(Melaleuca Oil)" [\[\[User:12.152.168.248|12.152.168.248\]\]](#)
17:16, 21 December 2006 (UTC)

Credit: Durova

But then later...

Revision as of 17:16, 21 December 2006 (edit)

12.152.168.248 (Talk | block)

(→Rename a good thing?)

← Older edit

Revision as of 17:17, 21 December 2006 (edit)

(undo)

Siraj88 (Talk | contribs | block)

(→Rename a good thing?)

Newer edit →

Line 171:

I agree with "Tea Tree Oil (Melaleuca Oil)" as well.

Anyone else to chime in? [[User:Siraj88|Siraj88]] 23:40, 14 December 2006 (UTC)

It's been a week. I'll change it from the current "Melaleuca Tea tree oil" to "Tea Tree Oil (Melaleuca Oil)" [[User:12.152.168.248|12.152.168.248]] 17:16, 21 December 2006 (UTC)

Line 171:

I agree with "Tea Tree Oil (Melaleuca Oil)" as well.

Anyone else to chime in? [[User:Siraj88|Siraj88]] 23:40, 14 December 2006 (UTC)

It's been a week. I'll change it from the current "Melaleuca Tea tree oil" to "Tea Tree Oil (Melaleuca Oil)" [[User:Siraj88|Siraj88]] 17:17, 21 December 2006 (UTC)



Credit: Durova