

Wikiwatching

Virgil Griffith

Disruptive Technologist

<http://virgil.gr>

virgil@caltech.edu

[Video]

When you edit Wikipedia...

You can edit leaving your username.

OR

You can edit anonymously, using your computer's IP address instead of a username.

BUT

Sometimes IP addresses can be traced back (it's tricky though).

How WikiScanner 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!

WikiScanner 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.

Other Interesting Work

- [#Hits for page](#)
- [Page edit statistics](#)
- [Coloring text by Trustworthiness](#)
- [Vispedia](#)

Only tip of the iceberg

- WikiScanner is charming, but easy to circumvent.
 - Too much data to sift through
 - Can be foiled by having an account, editing from home, etc.
- We can do better.

WikiWatcher = ++WikiScanner




New Work

Tools

1. [WikiScanner 2.0](#) (coi+pmcu)
Better, Faster, Stronger
2. [Wikiganda](#)
Finds edit wars between real-world organizations
3. ? (stay tuned)

Exhibits

1. [Poor Man's CheckUser](#) 
2. [Sockpuppetry](#)
3. [Beaver Scope](#)
4. "Diebold" (stay tuned)

Disruptive Datamining

- Security and Datamining get along great
 - It's friggin' easy!
 - You should be doing it.
- Take an example of a vulnerability someone found
 - ...and do it to the whole Internet.

Examples

- [Planet Sony](#)
- [Visualizing Speeches](#)
- WHOIS Scanner
 - “show me all domains registered to a @microsoft.com address”
(DomainTools.com has the data.)
- Spider for...
 - everything explicitly disallowed by robots.txt
 - salacious .doc metadata
 - uncensorable PDFs (demo)

Handy Tools

- Free data-mining tools
- General Architecture for Text Engineering (GATE)
- Text Similarity – python `difflib.get_close_matches()`
- MySQL / Python / Ruby

Special Thanks

- Rishi Chandy, Daniel Erenrich, Sonal Gupta
- Mani Chandy / www.infospheres.caltech.edu
- Durova
- Aaron Swartz / aaronsw.com
- Greg Randall / gregr.org
- Wikimedia Foundation

Questions ?

<http://virgil.gr>

<http://wikiwatcher.com>

virgil@caltech.edu

Poor Man's CheckUser

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)

+

But Later...

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

12.152.168.248 (Talk | block)

(→*Rename a good thing?*)

← Older 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)

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:Siraj88|Siraj88]] 17:17, 21
December 2006 (UTC)

Back