Some Lessons Learned From OSINT Tool Development & Operations

Mike Geide
PUNCH Cyber Analytics Group

About

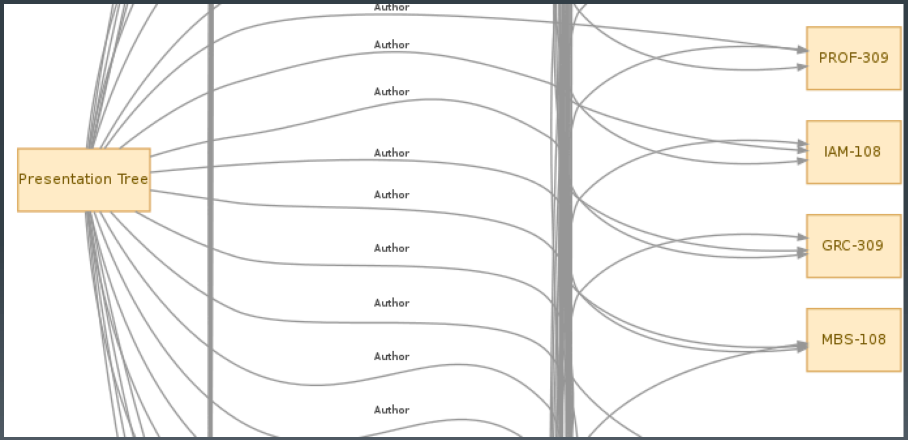
- Former life:
  - USG SOC/CSIRC Analyst (US-CERT and others)
  - Security researcher at Zscaler – data-rich secure web gateway cloud service provider
  - Author of Poortego and some other tools

- Present:
  - Co-founder and CTO at PUNCH Cyber Analytics Group

Mike Geide | mike@punchcyber.com | punchcyber.com

Poortego

- Why:
  - Intelligence tools largely commercial and some short-comings
  - Doing threat intelligence on budget
  - Sensed need/want for a FOSS equivalent to Maltego
- Developed a prototype
  - Ruby code - run stand-alone or as Metasploit plugin
  - Presented: SecTor, RSA Europe
  - Mentioned in Team Cymru Dragon Newsbytes, on CIF mailing list, etc.
- https://github.com/mgeide/poortego



Some things I’ve learned

- Getting contributors is hard… people are busy
  - Maybe more luck if was Python and/or web-based
- I got busy (and didn’t want to maintain code)
  - Start company, have 2nd kid, train for marathon, etc.
- Not going to compete with a commercial entity
  - Developer resources, QA, GUI/features, Documentation, Support
  - Maltego Tungsten is cool (see BlackHat 2013 preso)
- Intelligence components got a budget all the sudden
  - Threat intel, big data, APT, etc. buzz words

Poortego Future - TBD

- Death?
  - Instead purchase a Maltego license for $650
  - Write local transforms and leverage Python frameworks like Sploitego (https://github.com/allfro/sploitego)

- Or lots more coding?
  - Team-up with other developers if there is interest
  - Update code-base to support Python-based API/transforms
  - Make client/server web-based (nodejs?)
  - Leverage graph database? (neo4j)
  - Add in a smattering of features and cool graphics
  - Integrate, integrate, integrate …

Integration versus “yet another tool”

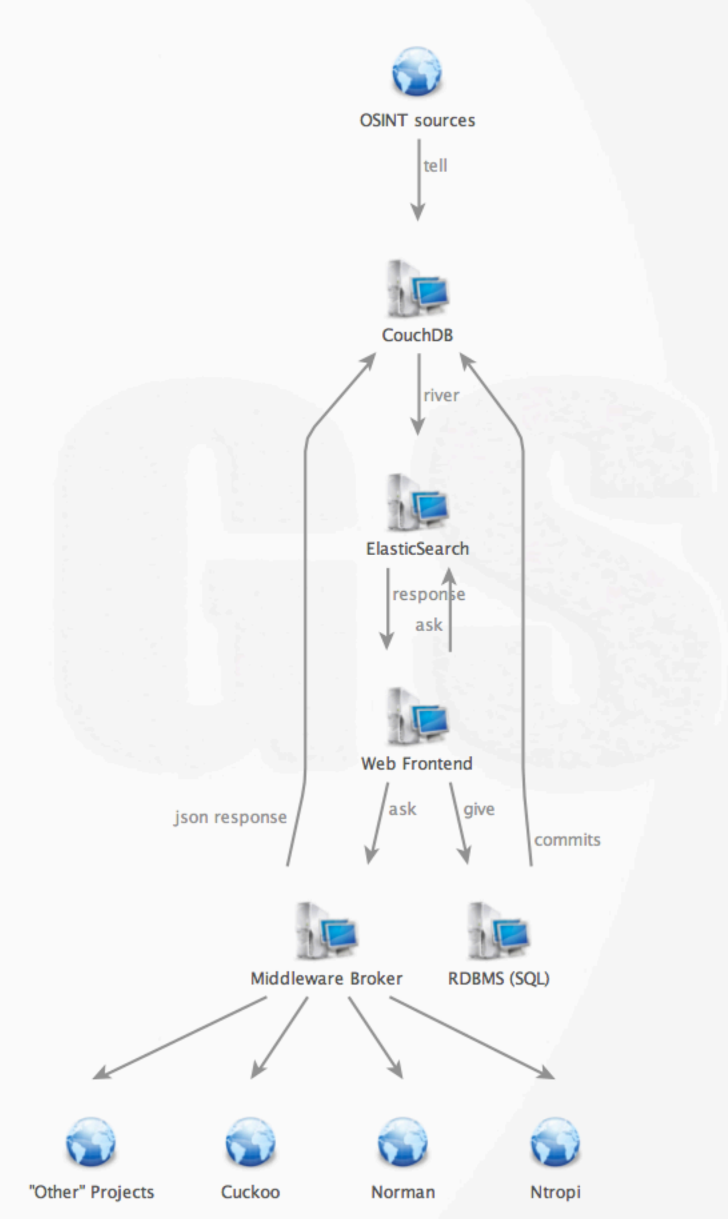
- Maltego machines will refresh/link off of what you know, but there is other data out there that you don’t know about
- More and more data feeds and tools that provide levels of intelligence are available – but they are not integrated
- Case study of how we handle in-house…

“The Hub” Overview

- Intended to solve our “integration problem”
  - Leverage any/all intelligence projects (FOSS or commercial)
    - Let them do what they’re good at / intended to do
  - Integrate into workflow quickly and easily
    - Versus modifying each underlying project code
  - One screen for querying or adding data quickly
    - Query/caching service providing actionable output

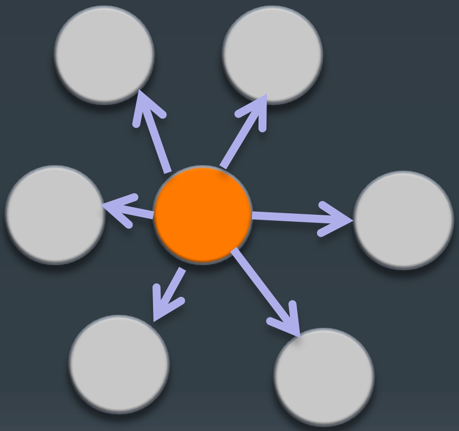
"The Hub" Architecture

- CouchDB – result data broker (JSON)
  - http://couchdb.apache.org/
- ElasticSearch / CouchDB Elastic River – index documents
  - http://www.elasticsearch.org/
  - https://github.com/elasticsearch/elasticsearch-river-couchdb
- Python and Bottle web framework



“The Hub” Integration

- In-house we leverage existing FOSS projects in the Hub:
    - MITRE’s CRITS –threat campaign tracking
        - https://github.com/MITRECND/
    - CIF instances – pulls of data feeds (e.g., ZeusTracker)
        - https://github.com/collectiveintel/cif-v1
    - Cuckoo – malware sandbox
        - https://github.com/cuckoobox/cuckoo
    - Moloch – pcap analysis and repository
        - https://github.com/aol/moloch
    - IOCextractor
        - https://github.com/stephenbrannon/IOCextractor
    - News-pet – RSS open-source feeds
        - https://code.google.com/p/news-pet/
    - Internal projects
        - “Ntropi” – internal domain and resolution tracking project
        - Norman Sandbox (commercial)
        - And Public/Private/Internal data sources (our “Stream”)

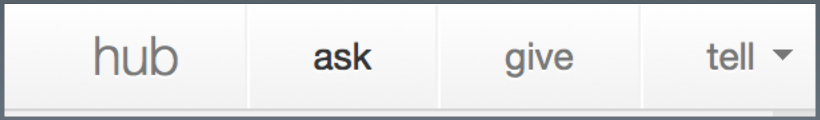


“The Hub” Simplicity

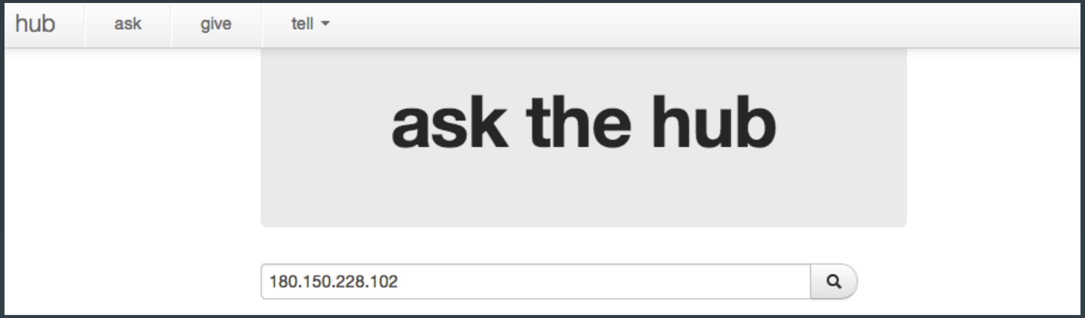
- Ask
  - Single query interface to all sources
    - Tags, email, hashes, domains, ips

- Give – directly provide data or a source of data
  - Allows importing of data from a URL, comma separated list, or plaintext file
    - If URL, and not already a source known to hub, it is added as “tell” source too

- Tell – tell the system where to go look on its own
  - Review/Add scheduled open source information gathering
    - Known sources, blogs, tweets, email lists, etc.



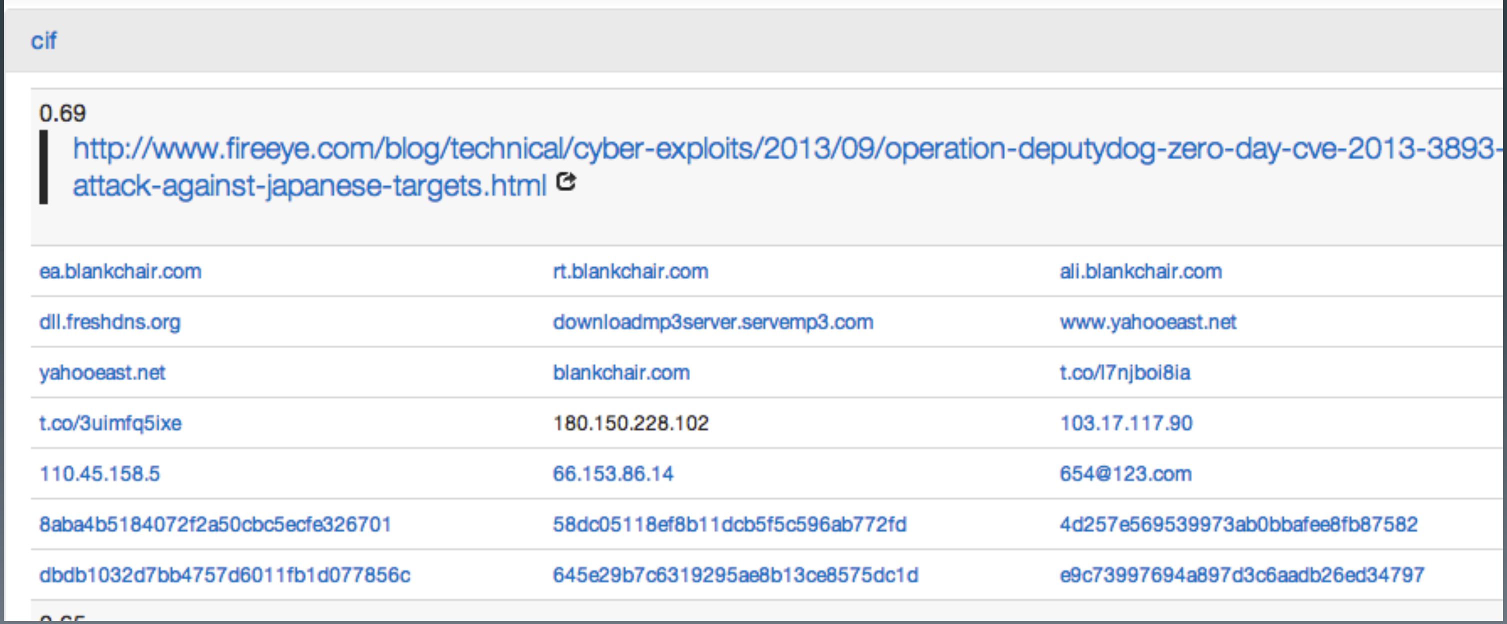
“The Hub” – Ask



ask the hub

180.150.228.102

IE 0Day example

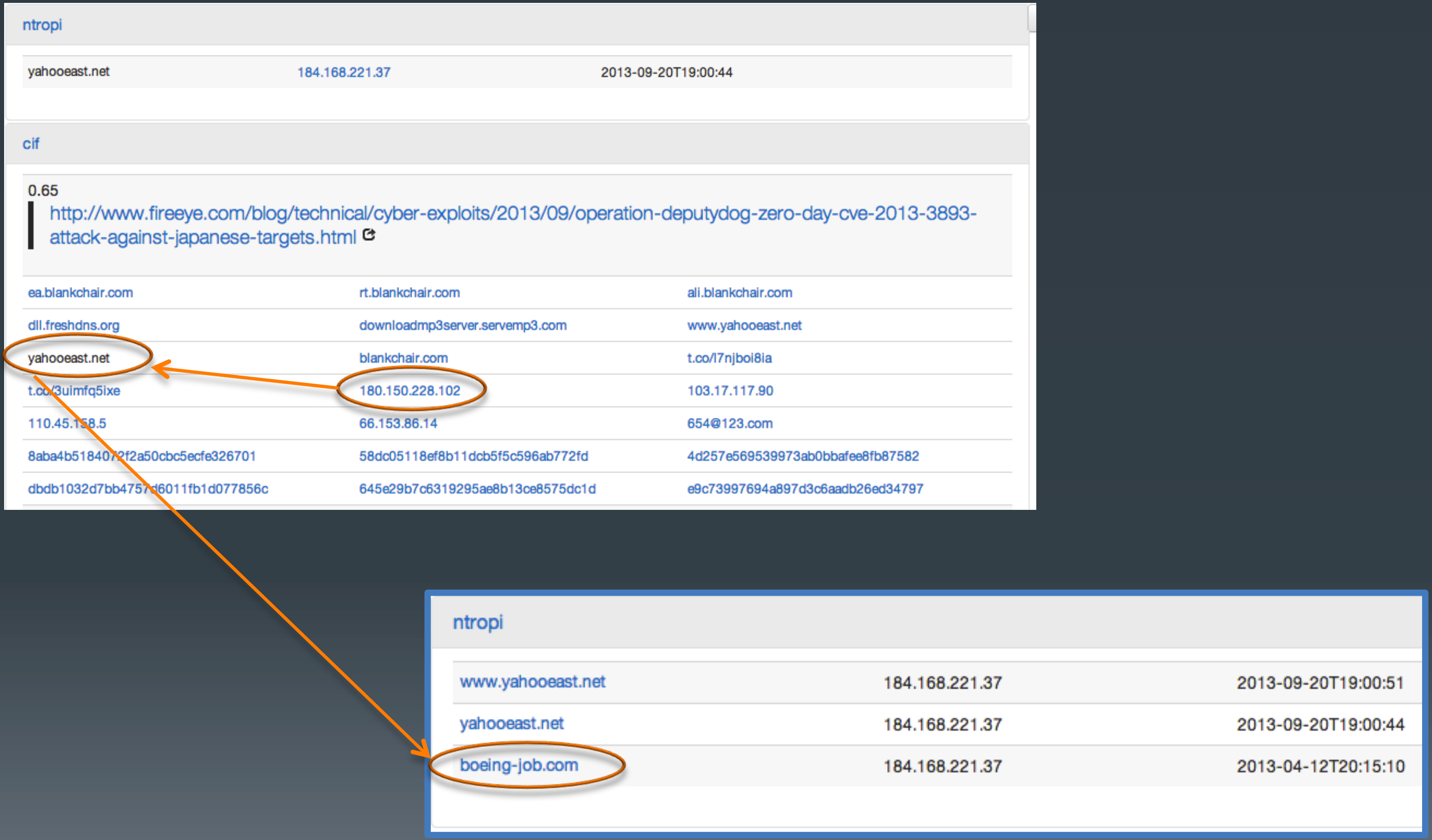


0.69

> http://www.fireeye.com/blog/technical/cyber-exploits/2013/09/operation-deputydog-zero-day-cve-2013-3893-attack-against-japanese-targets.html

ea.blankchair.com	rt.blankchair.com	ali.blankchair.com
dll.freshdns.org	downloadmp3server.servemp3.com	www.yahooeast.net
yahooeast.net	blankchair.com	t.co/I7njboi8ia
t.co/3uimfq5ixe	180.150.228.102	103.17.117.90
110.45.158.5	66.153.86.14	654@123.com
8aba4b5184072f2a50cbc5ecfe326701	58dc05118ef8b11dcb5f5c596ab772fd	4d257e569539973ab0bbafee8fb87582
dbdb1032d7bb4757d6011fb1d077856c	645e29b7c6319295ae8b13ce8575dc1d	e9c73997694a897d3c6aadb26ed34797

“The Hub” – Ask / Pivot



ntropi

yahooeast.net	184.168.221.37	2013-09-20T19:00:44

cif

0.65

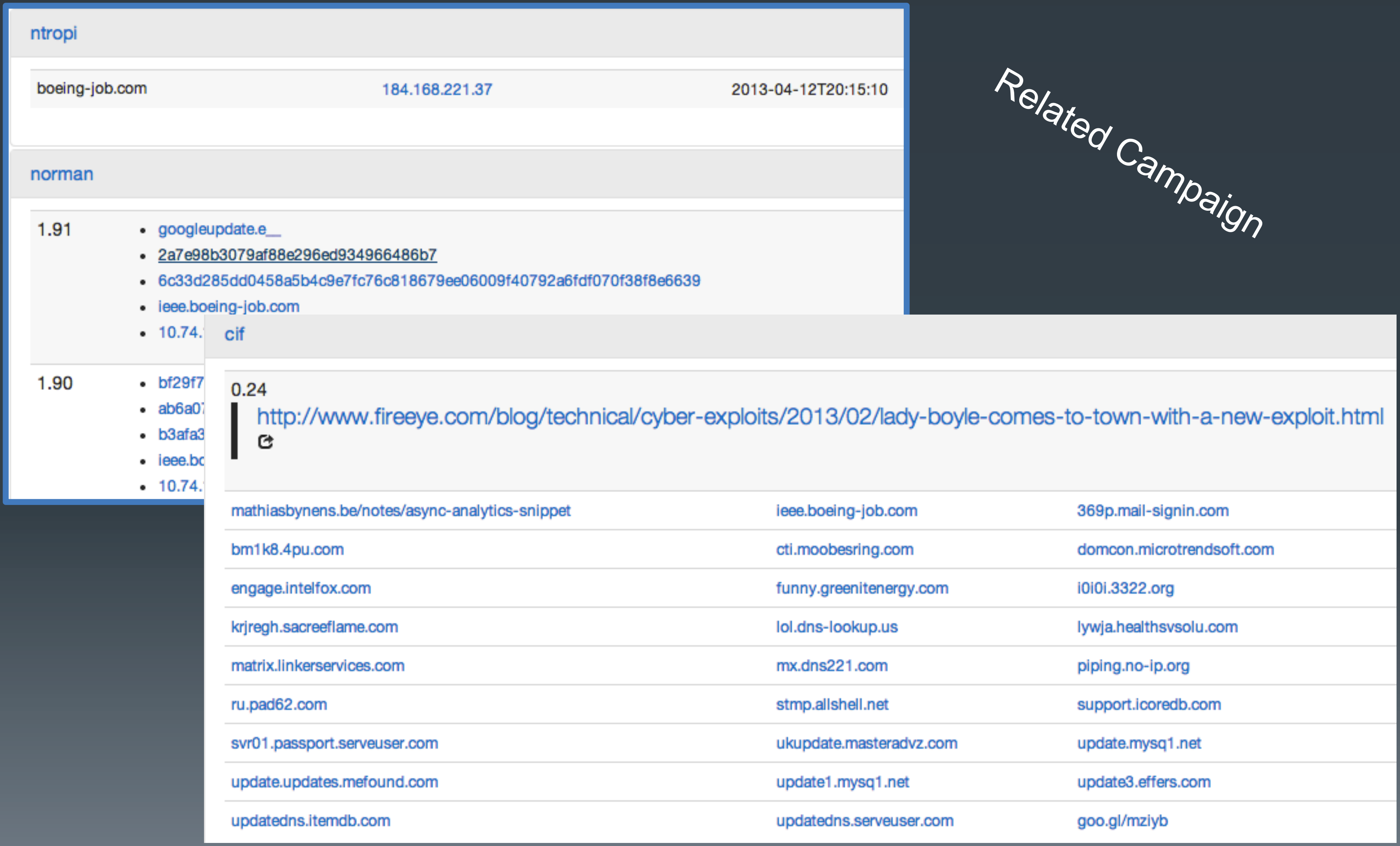
> http://www.fireeye.com/blog/technical/cyber-exploits/2013/09/operation-deputydog-zero-day-cve-2013-3893-attack-against-japanese-targets.html

ea.blankchair.com	rt.blankchair.com	ali.blankchair.com
dll.freshdns.org	downloadmp3server.servemp3.com	www.yahooeast.net
yahooeast.net	blankchair.com	t.co/l7njboi8ia
t.co/3uimfq5ixe	180.150.228.102	103.17.117.90
110.45.158.5	66.153.86.14	654@123.com
8aba4b5184072f2a50cbc5ecfe326701	58dc05118ef8b11dcb5f5c596ab772fd	4d257e569539973ab0bbafee8fb87582
dbdb1032d7bb4757d6011fb1d077856c	645e29b7c6319295ae8b13ce8575dc1d	e9c73997694a897d3c6aadb26ed34797

ntropi

www.yahooeast.net	184.168.221.37	2013-09-20T19:00:51
yahooeast.net	184.168.221.37	2013-09-20T19:00:44
boeing-job.com	184.168.221.37	2013-04-12T20:15:10

“The Hub” – Ask / Pivot (2)



ntropi

boeing-job.com	184.168.221.37	2013-04-12T20:15:10

norman

1.91
- googleupdate.e__
- 2a7e98b3079af88e296ed934966486b7
- 6c33d285dd0458a5b4c9e7fc76c818679ee06009f40792a6fdf070f38f8e6639
- ieee.boeing-job.com
- 10.74.

1.90
- bf29f7
- ab6a07
- b3afa3
- ieee.bo
- 10.74.

cif

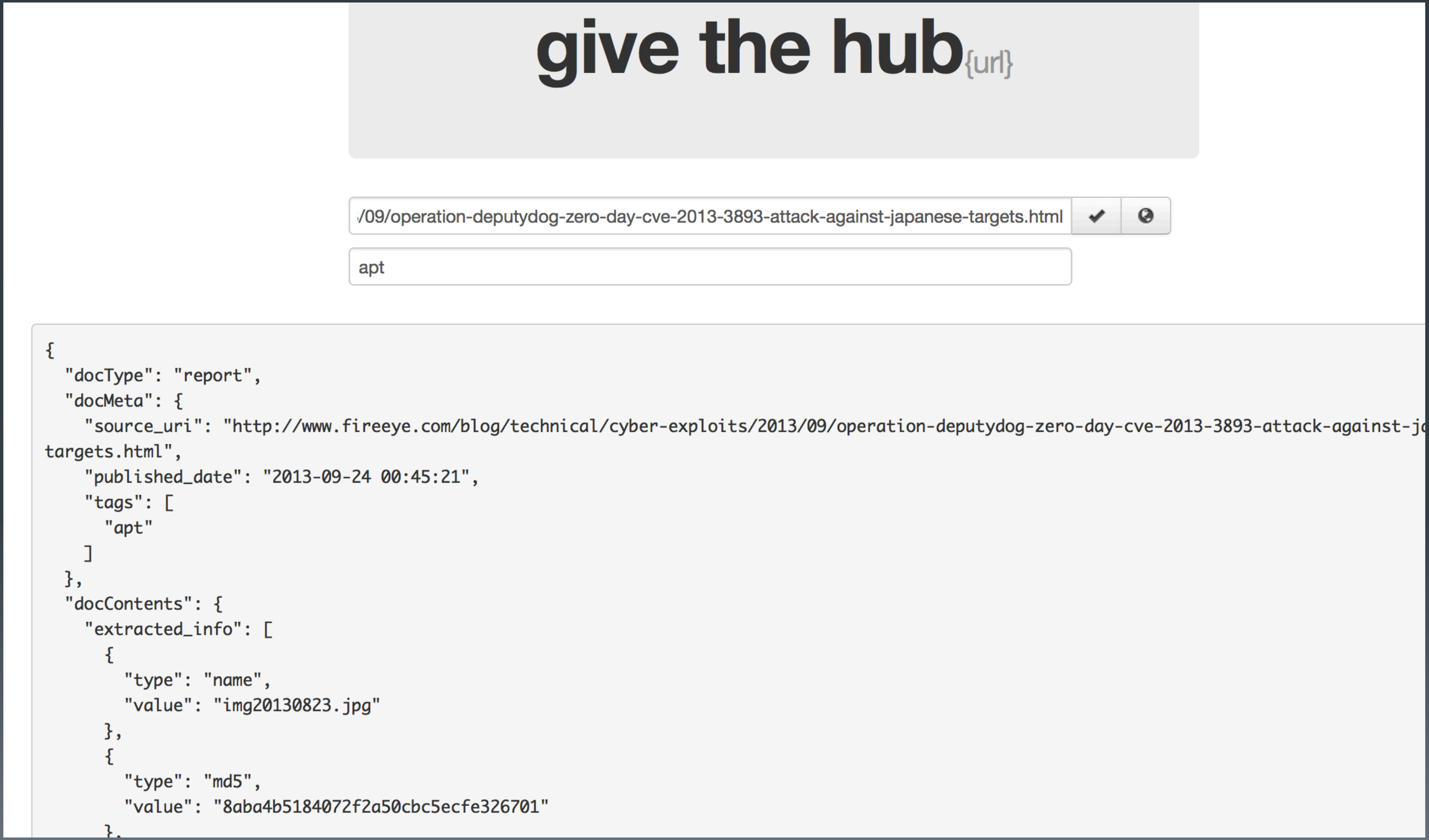
0.24

http://www.fireeye.com/blog/technical/cyber-exploits/2013/02/lady-boyle-comes-to-town-with-a-new-exploit.html

mathiasbynens.be/notes/async-analytics-snippet	ieee.boeing-job.com	369p.mail-signin.com
bm1k8.4pu.com	cti.moobesring.com	domcon.microtrendsoft.com
engage.intelfox.com	funny.greenitenergy.com	i0i0i.3322.org
krjregh.sacreeflame.com	lol.dns-lookup.us	lywja.healthsvsolu.com
matrix.linkerservices.com	mx.dns221.com	piping.no-ip.org
ru.pad62.com	stmp.allshell.net	support.icoredb.com
svr01.passport.serveuser.com	ukupdate.masteradvz.com	update.mysq1.net
update.updates.mefound.com	update1.mysq1.net	update3.effers.com
updatedns.itemdb.com	updatedns.serveuser.com	goo.gl/mziyb

Related Campaign

“The Hub” - Give



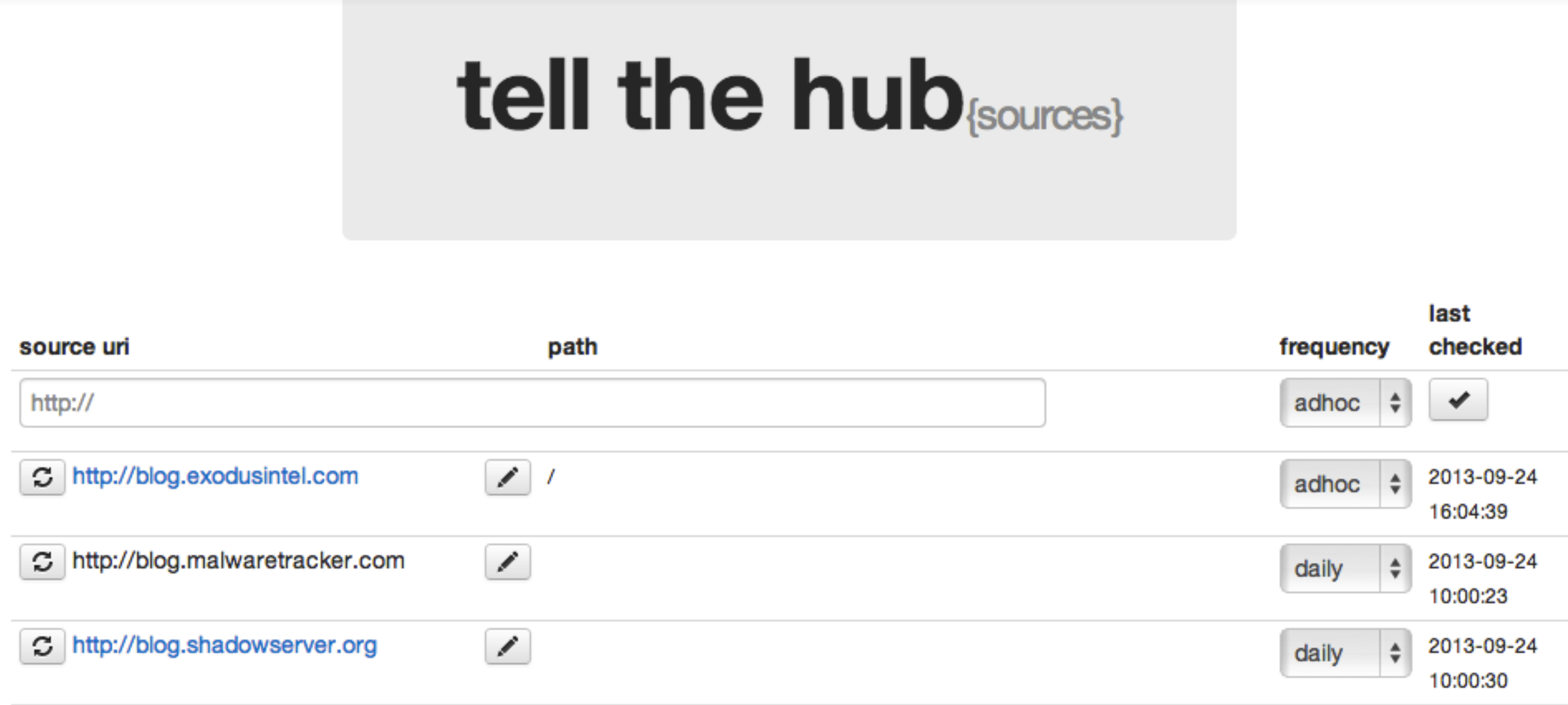
give the hub{url}

/09/operation-deputydog-zero-day-cve-2013-3893-attack-against-japanese-targets.html

apt

```
{
  "docType": "report",
  "docMeta": {
    "source_uri": "http://www.fireeye.com/blog/technical/cyber-exploits/2013/09/operation-deputydog-zero-day-cve-2013-3893-attack-against-j
targets.html",
    "published_date": "2013-09-24 00:45:21",
    "tags": [
      "apt"
    ]
  },
  "docContents": {
    "extracted_info": [
      {
        "type": "name",
        "value": "img20130823.jpg"
      },
      {
        "type": "md5",
        "value": "8aba4b5184072f2a50cbc5ecfe326701"
      },
```

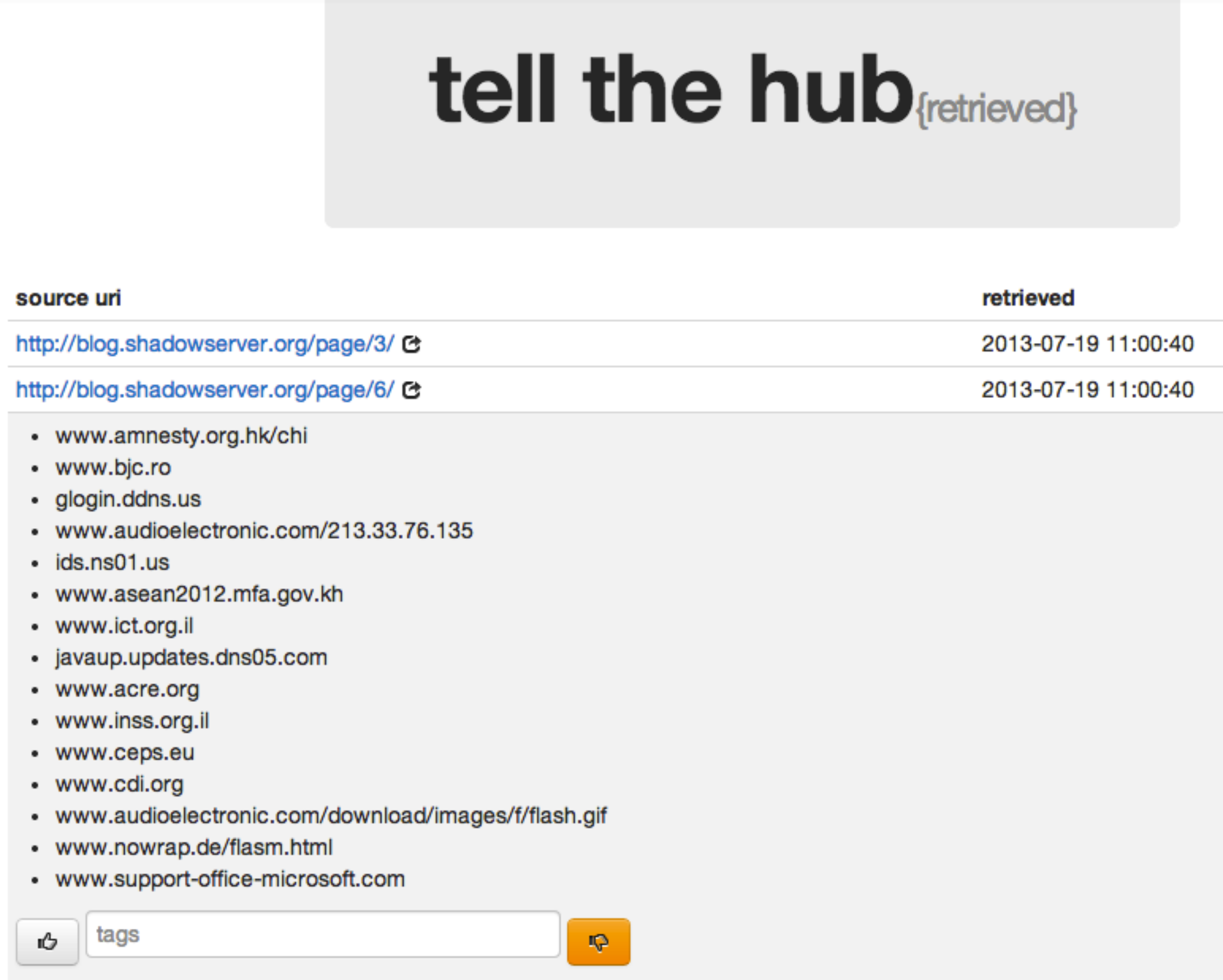
“The Hub” – Tell



tell the hub{sources}

source uri	path	frequency	last checked
http://		adhoc	✔
http://blog.exodusintel.com	/	adhoc	2013-09-24 16:04:39
http://blog.malwaretracker.com		daily	2013-09-24 10:00:23
http://blog.shadowserver.org		daily	2013-09-24 10:00:30

“The Hub” – Tell (History)



tell the hub{retrieved}

source uri	retrieved
http://blog.shadowserver.org/page/3/	2013-07-19 11:00:40
http://blog.shadowserver.org/page/6/	2013-07-19 11:00:40

- www.amnesty.org.hk/chi
- www.bjc.ro
- glogin.ddns.us
- www.audioelectronic.com/213.33.76.135
- ids.ns01.us
- www.asean2012.mfa.gov.kh
- www.ict.org.il
- javaup.updates.dns05.com
- www.acre.org
- www.inss.org.il
- www.ceps.eu
- www.cdi.org
- www.audioelectronic.com/download/images/f/flash.gif
- www.nowrap.de/flasm.html
- www.support-office-microsoft.com

Conclusions

- Lots of projects, data, etc.
  - “Big Data” is hot right now

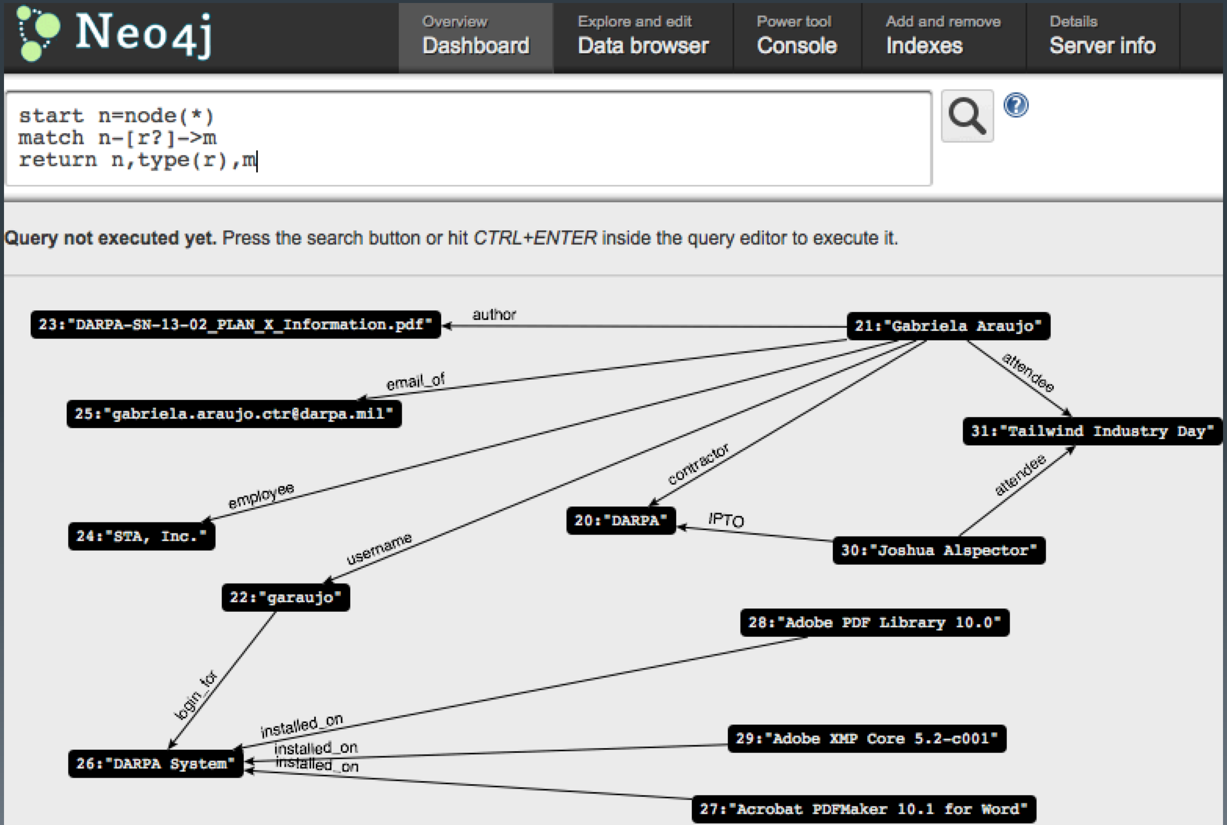
- Spend more time integrating and doing your analysis than developing
  - Unless you enjoy maintaining code ;)

- “The Hub” style integration is effective and simple to use
  - Let the complexities reside in any of the other integrated projects

And yes, PUNCH can help set this up for you
(shameless plug)

Mike Geide | mike@punchcyber.com | punchcyber.com

Sidebar – GraphDB (neo4j)



```
start n=node(*)
match n-[r?]->m
return n,type(r),m
```

Query not executed yet. Press the search button or hit *CTRL+ENTER* inside the query editor to execute it.