APWG Adventures In Information Sharing:

Now and For the Future

Patrick Cain

Resident Research Fellow

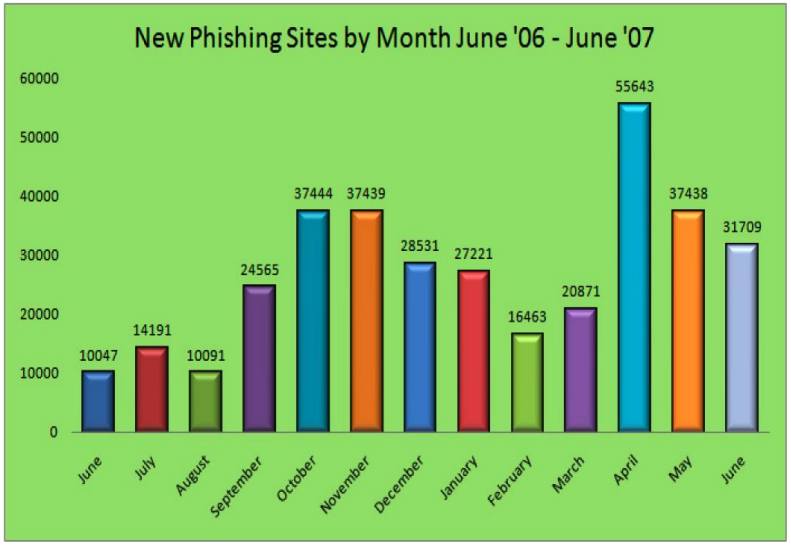
Agenda

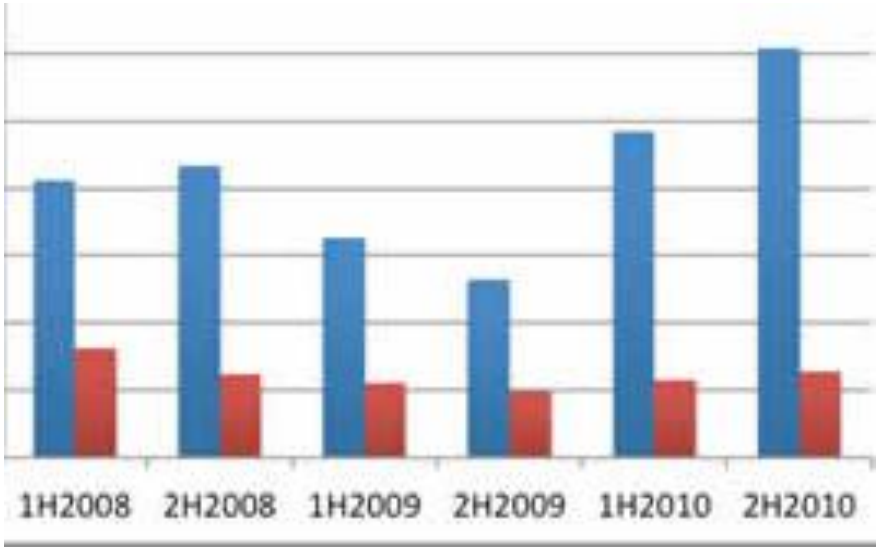
- The APWG
- Our Dilemma
- The Plan
- Current Environment
- Long-Term Goal
- Issues

The APWG

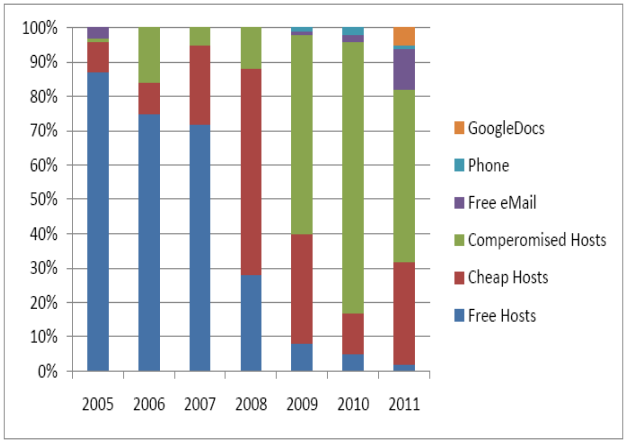
- Started In 2004
- Non-profit CA corporation
- ~3700 members, 25 researcher groups
  - National Bodies, CERTs, LEA == free
  - International Composition
- 1500 or more ‘clingeroners’
- Goal: solve problems, share experiences and data
- Be vendor, country, and * agnostic

We Publish Statistics





RANK	TLD	TLD Location	# Unique Phishing attacks 2H2010	Unique Domain Names used for phishing 2H2010	Domains in registry Oct 2010	Score: Phish per 10,000 domains 2H2010
1	.th	Thailand	125	65	51,438	12.6
2	.ir	Iran	295	169	175,600	9.6
3	.ma	Morocco	73	34	36,669	9.3
4	.ie	Ireland	112	96	151,023	6.4
5	.tk	Tokelau	2,533	2,429	4,030,709	6.0
6 (tie)	.kz	Kazakhstan	49	28	50,534	5.5
6 (tie)	.cc	Cocos Islands	4,963	55	100,000	5.5
7	.in	India	523	421	791,165	5.3
8	.my	Malaysia	68	55	108,211	5.1
9	.hu	Hungary	365	255	542,000	4.7



APWG

Detail from the 2H2010 Report

Rank	TLD	TLD Location	# Unique Phishing Attacks 2H2010	Unique Domain Names used for Phishing 2H2010	Domains in Registry 2010	Score: Phish per 10,000 domains
1	.th	Thailand	125	65	51,438	12.6
2	.ir	Iran	295	169	175,600	9.6
3	.ma	Morocco	73	34	36,669	9.3
4	.ie	Ireland	112	96	151,023	6.4
5	.tk	Tokelau	2,533	2,429	4,030,709	6.0
6 (tie)	.kz	Kazakhstan	49	28	50,534	5.5
6 (tie)	.cc	Cocos Islands	4,963	55	100,000	5.5
7	.in	India	523	421	791,165	5.3
8	.my	Malaysia	68	55	108,21	5.1
9	.hu	Hungary	365	255	542,000	4.7

Many Years as a Trend

Year	1H2008	2H2008	1H2009	2H2009	1H2010	2H2010
1	*Hong Kong*	Venezula	Peru	**Thailand**	**Thailand**	**Thailand**
2	**Thailand**	**Thailand**	**Thailand**	Korea	Korea	Iran
3	Belize	Belize	Belize	Ireland	Ireland	Morocco
4	Venezuela	**Soviet Union**	*Belgium*	*Belgium*	Poland	Ireland
5	*Chile*	Romania	Romania	Romania	Chile	Tokelau
6	Romania	*Chile*	Taiwan	Malaysia	Malaysia	Korea
7	**Liechtenstein**	Korea	Korea	.eu	Greece	**Cocos Islands**
8	.name	Vietnam	*Chile*	Iran	Romania	India
9	Taiwan	Russia	Ireland	Poland	Vietnam	Malaysia
10	Korea	Taiwan	Malaysia	Mexico	Czech Rep	Hungary

APWG

The ‘Big Plan’

- Don’t Identify today’s problem(s)!
- Don’t research the next big one!
- Plan for the generic future
  - When new threats arise, be ready to triage & correlate
- Act more like a data clearinghouse
  - Use the power of others for common good
    - Make it easy for investigators to get good data
    - Make it easy for parallel investigations
  - Assemble a data corpus for research and investigation
    - (real) Stats make the message! (e.g., global phishing report)
    - Trending allows for more saner decision making

APWG

Our Historical Dilemma

- Many people do ‘investigations’
  - Some won’t share with everybody
    - Laws, privacy, lawyers, will, etc
- We only get one crack at some data
  - It’s screwed → we’re screwed
  - Force submitter to supply some fields
- Many of our submitters/pullers have “no time/money/brains for tools”
- VOLUME: It can’t be done with a human

APWG

The Early Plan

- Convince people to send us phishing URLs
  - There's a lot of 'em
  - There's no standard way to do this
- Processing (extract URL, verify, etc) takes time

- !
- We need a standard format and process
- Get people to trust the process

APWG

The ‘We Know Better Now’ Plan

- Picked XML as a data format; selected IODEF as a message format
  - Wrote some extensions for phishing (IETF RFC5901)
- Take XML in; push XML out; (store XML)
  - Everybody can read XML (unlike ASN.1)
  - En/De-coding tools are pretty much free
  - ‘Security’ comes as part of XML Security Suite
- We’ll make the tools that people need
- BE FLEXIBLE – crime is evolving

APWG

As We Evolve…

- IODEF/XML is good
  - Language tags for all text elements
  - Easy to craft new EventData elements
    - Don’t need to mod the standard for new data types
  - Lots of OTS tools (XFORMS, XQUERY, XSLT, etc)

- IODEF/XML is not so good
  - Overhead (not good for packet capture!)
  - Not all data types are defined

APWG

An Unexpected Surprise

- If you correlate enough data you detect patterns (and the bad guys)
- How do we exchange data efficiently?
  - No human involved, automated
  - Share in-process investigative data
- Have the submitter do most of the processing
  - Make the process easy and cheap.
- Ooooh. We could do this with other data, too.

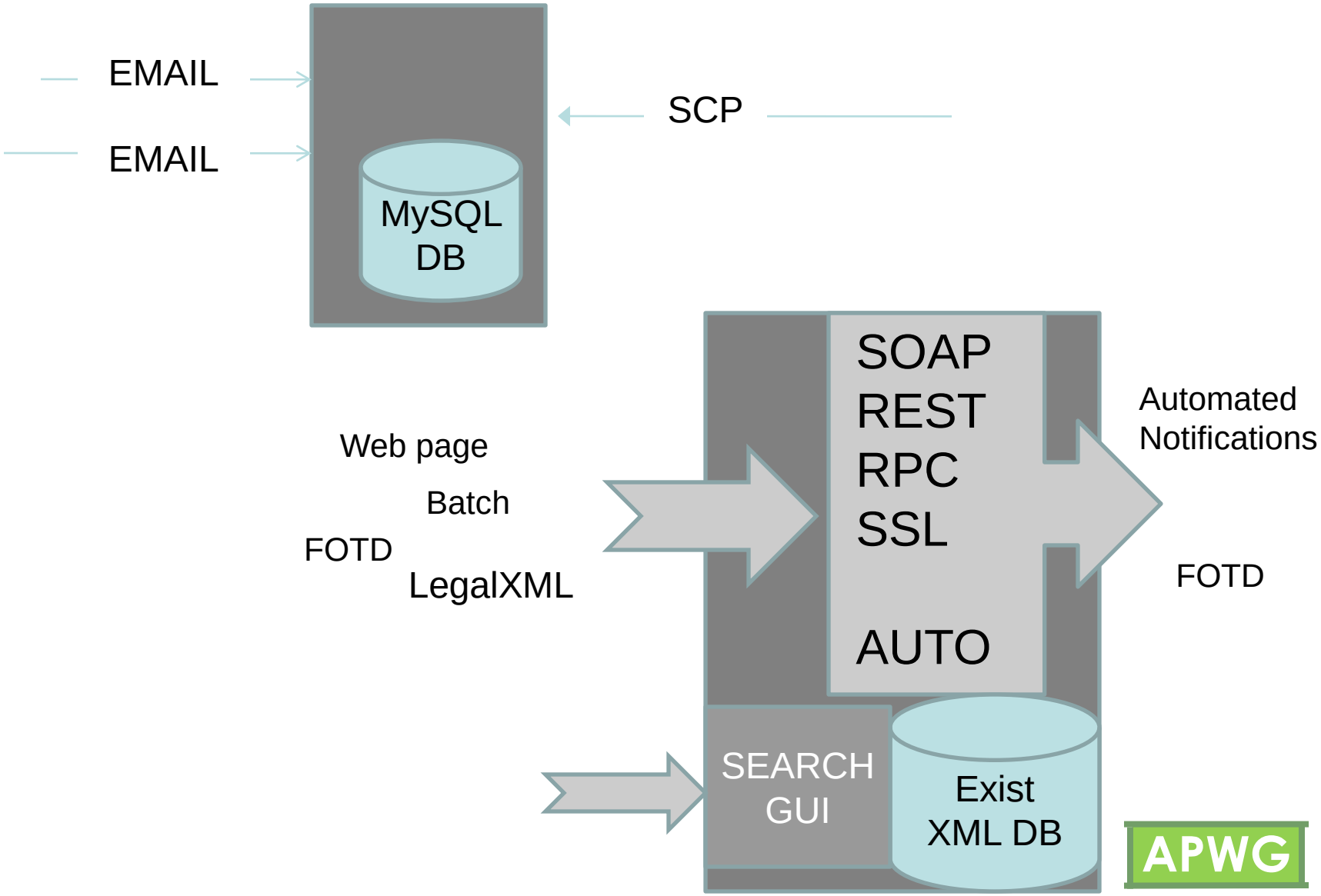
- Now people just show up with data.

APWG

The APWG Current World

- We take in raw phish lures, URLs via email, (IODEF XML reports over SOAP), (XFORMS UI) (phish email addresses) (vishing numbers)
- Data can also be submitted via a web page
- We take in … (other stuff)
- Output UBL as CSV (and as IODEF XML)
- (Search repository -> output as HTML, IODEF)

Current Test Environment



Long-Term Goal

- Make schemas for different ‘types of data’
  - IODEF EventData XML blobs
    - Even if not used w/ IODEF, they can be useful
    - During development, we call them APWG standards
  - We’re moving towards ‘eCrime’ reporting
    - Can we make the data ‘actionable’? Understandable by LEO?
- Use standard transport, email, etc, vehicles
  - No new protocols; no multi-century developments
- Get other people to buy into the ideas
  - Pretty successful so far

APWG

Issues

- As we slop data, there's more to agree on…
- How to convey policy info
  - Restriction markings
    - How to mark: Share with LEO?
    - How to mark: Share with Friends?
  - Generally accepted impact definitions
    - The attack 'method'
    - The 'impact' of the attack
  - How to mark: Know but no Touchee!
- LEO guidance on data to put in a report
- Watch ITU-related and other similar efforts

APWG

How to convey policy/sharing info

- Restriction markings
  - How to mark: Share with LEO? Friends? Public?
  - How to show: Know but no Touch!
- Can this data be shared with law enforcement?
  - 0 - Do NOT share this data with Law Enforcement
  - 1 - Share this data with Law Enforcement if an investigation is open
  - 2 - Feel free to share this data with Law Enforcement
  - 3 - I have previously shared this data with Law Enforcement

APWG

A silly example

- Sharing with the ‘Public’:
  - 0 - Do not share
  - 1 - Summary data may be shared
  - 2 - Details may be shared
  - 3 – Too late. (already shared)

An example…

- How can this data be shared within the APWG/xxx?
  - 0 - For recipient use only
  - 1 – Recipient(s) should NOT share details of this data outside of members
  - 2 - Recipient(s) may share with their internal group
  - 4 – Summary data may be shared with other trusted security types
  - 6 - Data details may be shared with other trusted security types
  - 9 – Data has no sharing restrictions

Other Items to Specify

- Generally accepted impact definitions
- Common attack method definitions
  - Can we use CAPEC?
- LEO guidance on data to put in a report
- Watch ITU-related and other efforts

APWG

Getting the LEAs attention… ☺

- The goal it to catch the bad guy
- How do we get countries to devote resources to eCrime?
- How do we get LEA's attention?
  - We need the minister of justice's attention
- How do we get Justice's attention?
  - Define risks to their environment
  - Use statistics for education
  - Sounds like a paper.. ☺ (Has it been done before?)

APWG

A Diversion

- Interaction with the UN eCrime Commission convinced me that some organizations, companies, and member-states will never report any type of specific eCrime statistics.
- This is bad
  - Stats help countries prioritize response
  - Stats help us plan response actions
  - Our stats won't help (non-country specific)
- It will get worse
  - APT, night dragon, cheese slider, etc

APWG

Ignoring our current stats…

Can we slide some stats to a new model?

- Define the risks to an organization from the internet
  - Kind of like what ISO/IEC 27032 may do
- Refine some (general) threats from those risks
- Identify threat-specific malicious behaviour

- Report stats as ‘threats and risks’ based.

So how could this be useful?

- We volunteered to write an “Internet Threat Assessment” to help our treaty partners understand the risks and educate their justice ministries.
- APWG effort to develop an Internet eCrime Taxonomy

- This is live research; views welcome
  - ‘Live’ as in still changing

APWG

The Top-Level Risks

- Financial Loss
- Data Misuse
  - Proprietary
  - Personal
- Content Controls
  - Content Restrictions
  - Access to Prohibited Content
- Business Interference
- Loss of Network Control
- Distribution of Prohibited Speech
- Loss of Privacy
- (Reputation)

Digging into the Risks/Threats

- Financial Loss
  - Fraudulent transactions
  - Improper Credential Use
  - Laundering Activities
  - Extortion
- Proprietary Data Misuse
  - Possession
  - Corruption, Deletion
  - Misuse
  - Cyber Stalking
- Personal Data Misuse
  - Possession
  - Alteration
  - Misuse/Trafficing?
  - Falsification
- (Controlling Content)
- Access to Prohibited Content
  - Illegal porn
  - Pirated artistic works
- Distribution of Prohibited Speech
  - Hate speech
  - Death threats
  - Cyber-bullying
- Business Interference
  - DOS
- Loss of Network Control
  - Network Service Unavail – (DOS)
  - Network Compromised
- Loss of Privacy
  - Data Aggregation

Risks vs Participants

Risk	Company	Government	Person	Alien
Financial Loss	✓	✓	✓	
Data Misuse	✓	✓		
Proprietary	✓	✓		
Personal	✓	?	✓	
Controlling Content				
Access to Prohibited Content	✓	✓	✓	
Restrictions	✓	✓	✓	
Distribution of Prohibited Speech	✓	✓	✓	
Business Interference	✓	✓		
Loss of Network Control	✓	✓		
Personal Data Misuse		✓	✓	
Loss of Privacy	✓	✓	✓	

A Better View

- That ‘view’ requires user education
- A better view may be listing the crime (as defined in current laws) and generating:
  - How this crime is done ‘Internetly’
  - How it relates to the current policing and justice models

  - We’ll have to re-educate the techies, but fall more in line with normal justice/policing terminology

Current Work

- More schemas
  - Recovered Credentials, botz, cyber bullying
  - How do we share ‘computer misuse’ data?
- Work on “The Internet Threat Assessment”
  - Figure out how to measure eCrime
- Deal with ‘International issues’

APWG

Our next steps

- Deal with the issues; find new ones
  - The APWG way is find a problem; craft a solution; try solution; declare defeat; and modify solution. Try again.
  - We've learned a LOT trying to share data.
- Finish ongoing development
  - Finish our toolsets
- Run an eCrime IODEF Pilot this fall (maybe).
  - Multi-country, multi-language, multi-grief
  - Can we report and understand set scenarios?

APWG

Thank you

Pat Cain
Resident Research Fellow
APWG
pcain@apwg.org