theguardian

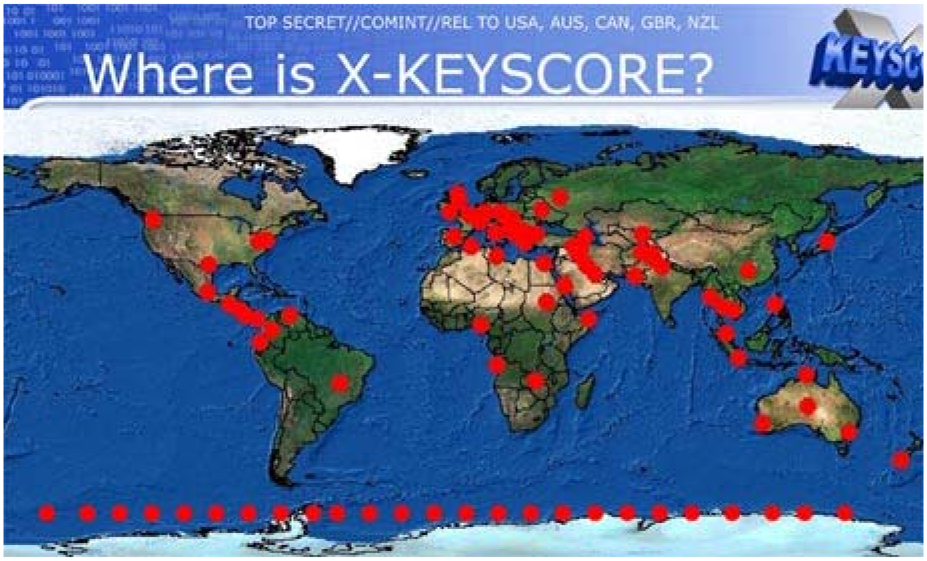
XKeyscore: NSA tool collects 'nearly everything a user does on the internet'

- XKeyscore gives 'widest-reaching' collection of online data
- NSA analysts require no prior authorization for searches
- Sweeps up emails, social media activity and browsing history
- NSA's XKeyscore program – read one of the presentations

Follow Glenn Greenwald by email BETA

Glenn Greenwald

theguardian.com, Wednesday 31 July 2013 08.56 EDT



One presentation claims the XKeyscore program covers 'nearly everything a typical user does on the internet'

A top secret National Security Agency program allows analysts to search with no prior authorization through vast databases containing emails, online chats and the browsing histories of millions of individuals, according to documents provided by whistleblower Edward Snowden.

The NSA boasts in training materials that the program, called XKeyscore, is its "widest-reaching" system for developing intelligence from the internet.

The latest revelations will add to the intense public and congressional debate around the extent of NSA surveillance programs. They come as senior intelligence officials testify to the Senate judiciary committee on Wednesday, releasing classified documents in response to the Guardian's earlier stories on bulk collection of phone records and Fisa surveillance court oversight.

The files shed light on one of Snowden's most controversial statements, made in his first video interview published by the Guardian on June 10.

"I, sitting at my desk," said Snowden, could "wiretap anyone, from you or your accountant, to a federal judge or even the president, if I had a personal email".

US officials vehemently denied this specific claim. Mike Rogers, the Republican chairman of the House intelligence committee, said of Snowden's assertion: "He's lying. It's impossible for him to do what he was saying he could do."

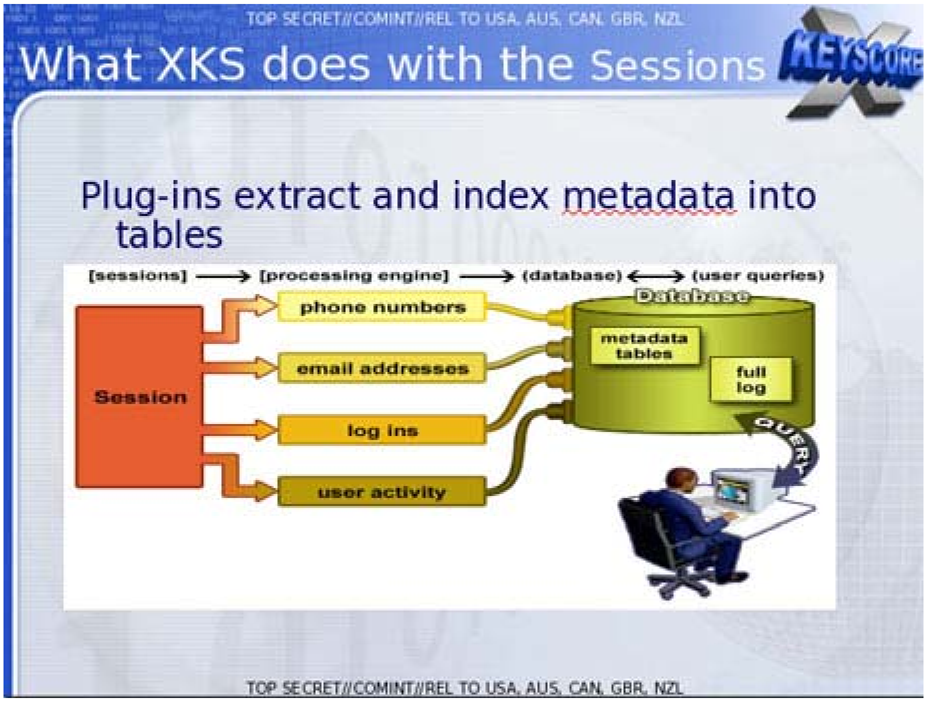
But training materials for XKeyscore detail how analysts can use it and other systems to mine enormous agency databases by filling in a simple on-screen form giving only a broad justification for the search. The request is not reviewed by a court or any NSA personnel before it is processed.

XKeyscore, the documents boast, is the NSA's "widest reaching" system developing intelligence from computer networks – what the agency calls Digital Network Intelligence (DNI). One presentation claims the program covers "nearly everything a typical user does on the internet", including the content of emails, websites visited and searches, as well as their metadata.

Analysts can also use XKeyscore and other NSA systems to obtain ongoing "real-time" interception of an individual's internet activity.

Under US law, the NSA is required to obtain an individualized Fisa warrant only if the target of their surveillance is a 'US person', though no such warrant is required for intercepting the communications of Americans with foreign targets. But XKeyscore provides the technological capability, if not the legal authority, to target even US persons for extensive electronic surveillance without a warrant provided that some identifying information, such as their email or IP address, is known to the analyst.

One training slide illustrates the digital activity constantly being collected by XKeyscore and the analyst's ability to query the databases at any time.



TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL

What XKS does with the Sessions

Plug-ins extract and index metadata into tables

TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL

The purpose of XKeyscore is to allow analysts to search the metadata as well as the content of emails and other internet activity, such as browser history, even when there is no known email account (a "selector" in NSA parlance) associated with the individual being targeted.

Analysts can also search by name, telephone number, IP address, keywords, the language in which the internet activity was conducted or the type of browser used.

One document notes that this is because "strong selection [search by email address] itself gives us only a very limited capability" because "a large amount of time spent on the web is performing actions that are anonymous."

The NSA documents assert that by 2008, 300 terrorists had been captured using intelligence from XKeyscore.

Analysts are warned that searching the full database for content will yield too many results to sift through. Instead they are advised to use the metadata also stored in the databases to narrow down what to review.

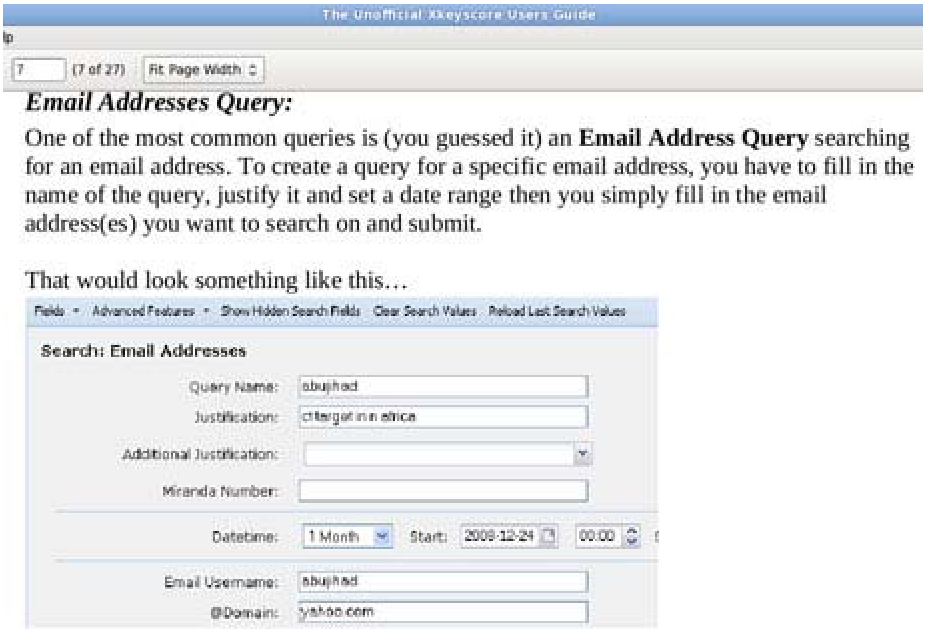
A slide entitled "plug-ins" in a December 2012 document describes the various fields of information that can be searched. It includes "every email address seen in a session by both username and domain", "every phone number seen in a session (eg address book entries or signature block)" and user activity – "the webmail and chat activity to include username, buddylist, machine specific cookies etc".

Email monitoring

In a second Guardian interview in June, Snowden elaborated on his statement about being able to read any individual's email if he had their email address. He said the claim was based in part on the email search capabilities of XKeyscore, which Snowden says he was authorized to use while working as a Booz Allen contractor for the NSA.

One top-secret document describes how the program "searches within bodies of emails, webpages and documents", including the "To, From, CC, BCC lines" and the 'Contact Us' pages on websites".

To search for emails, an analyst using XKS enters the individual's email address into a simple online search form, along with the "justification" for the search and the time period for which the emails are sought.


The Unofficial Xkeyscore Users Guide
lp
7
(7 of 27)
Fit Page Width
Fields
Advanced Features
Show Hidden Search Fields
Clear Search Values
Reload Last Search Values

Email Addresses Query:

One of the most common queries is (you guessed it) an **Email Address Query** searching for an email address. To create a query for a specific email address, you have to fill in the name of the query, justify it and set a date range then you simply fill in the email address(es) you want to search on and submit.

That would look something like this…

Search: Email Addresses

Query Name: abujihad

Justification: ct target in n africa

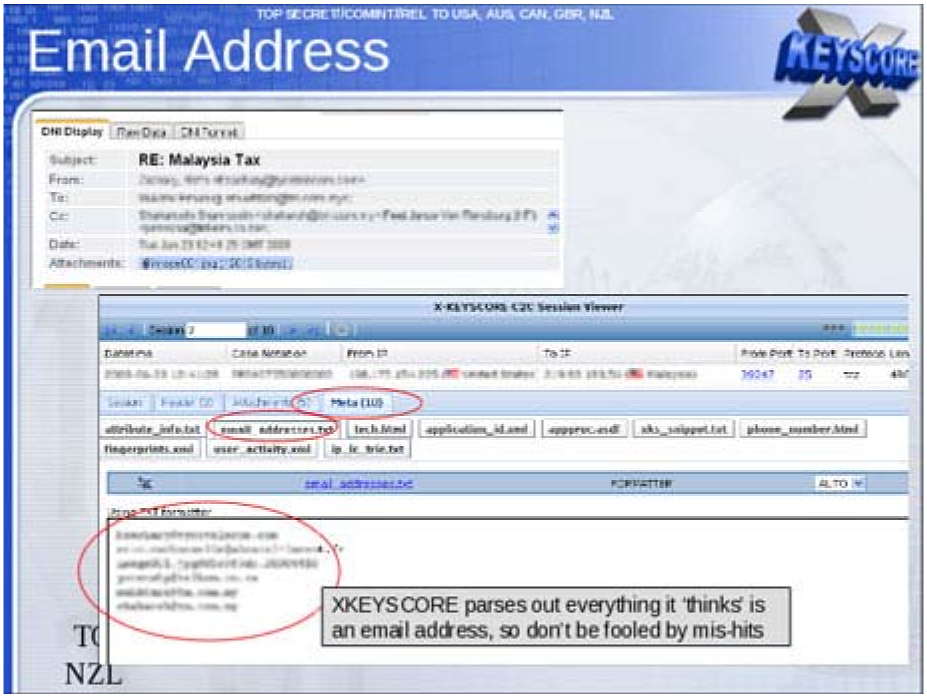
Additional Justification:

Miranda Number:

Datetime: 1 Month Start: 2008-12-24 00:00

Email Username: abujihad

@Domain: yahoo.com



Email Address

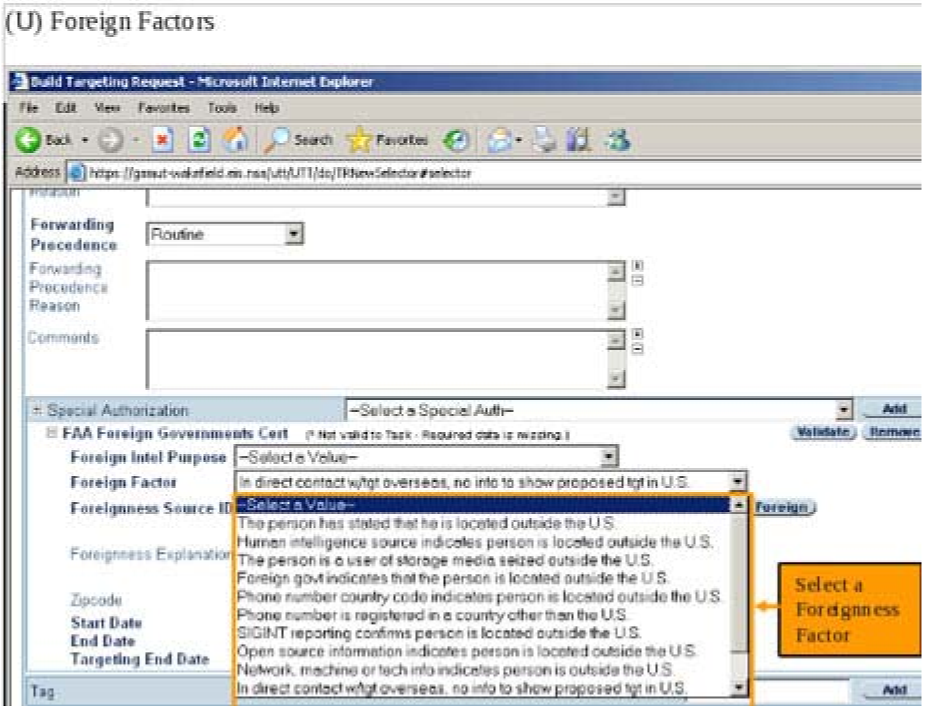
Subject: RE: Malaysia Tax
From: [illegible]
To: [illegible]
Cc: [illegible]
Date: [illegible]
Attachments: [illegible]

XKEYSCORE parses out everything it 'thinks' is an email address, so don't be fooled by mis-hits

The analyst then selects which of those returned emails they want to read by opening them in NSA reading software.

The system is similar to the way in which NSA analysts generally can intercept the communications of anyone they select, including, as one NSA document put it, "communications that transit the United States and communications that terminate in the United States".

One document, a top secret 2010 guide describing the training received by NSA analysts for general surveillance under the Fisa Amendments Act of 2008, explains that analysts can begin surveillance on anyone by clicking a few simple pull-down menus designed to provide both legal and targeting justifications. Once options on the pull-down menus are selected, their target is marked for electronic surveillance and the analyst is able to review the content of their communications:



(U) Foreign Factors

Build Targeting Request - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites
Address https://gasut-wakefield.eis.nsa/utt/UT1/do/TRNewSelector#selector

Forwarding Precedence: Routine

Forwarding Precedence Reason

Comments

Special Authorization: --Select a Special Auth-- Add

FAA Foreign Governments Cert (* Not valid to Task - Required data is missing) Validate Remove

Foreign Intel Purpose: --Select a Value--

Foreign Factor: In direct contact w/tgt overseas, no info to show proposed tgt in U.S.

Foreignness Source ID: --Select a Value--

- The person has stated that he is located outside the U.S.
- Human intelligence source indicates person is located outside the U.S.
- The person is a user of storage media seized outside the U.S.
- Foreign govt indicates that the person is located outside the U.S.
- Phone number country code indicates person is located outside the U.S.
- Phone number is registered in a country other than the U.S.
- SIGINT reporting confirms person is located outside the U.S.
- Open source information indicates person is located outside the U.S.
- Network, machine or tech info indicates person is outside the U.S.
- In direct contact w/tgt overseas, no info to show proposed tgt in U.S.

Foreign

Foreignness Explanation

Zipcode

Start Date

End Date

Targeting End Date

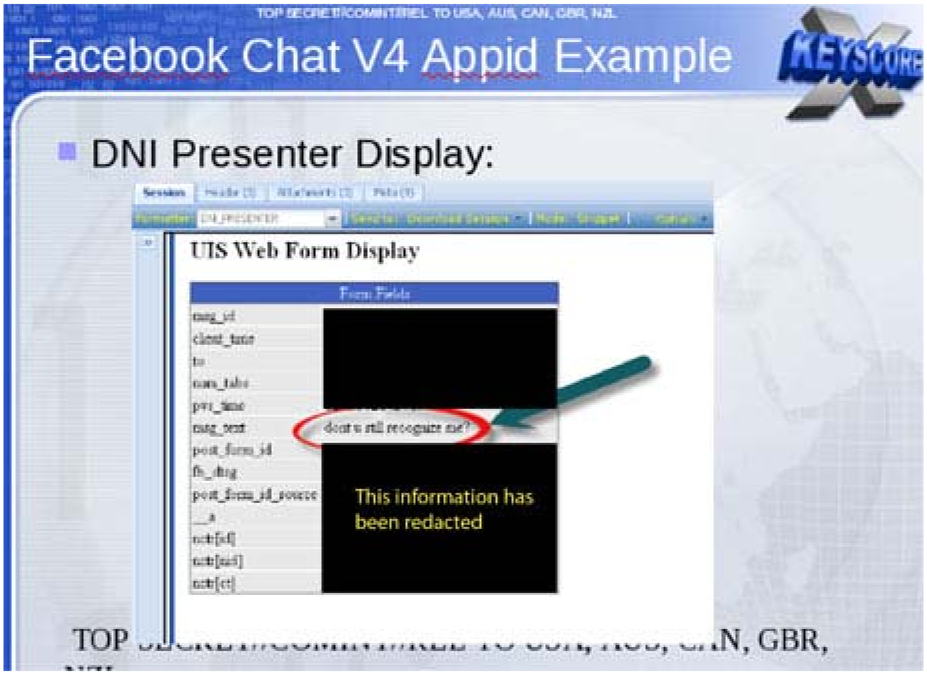
Tag Add

Select a Foreignness Factor

Chats, browsing history and other internet activity

Beyond emails, the XKeyscore system allows analysts to monitor a virtually unlimited array of other internet activities, including those within social media.

An NSA tool called DNI Presenter, used to read the content of stored emails, also enables an analyst using XKeyscore to read the content of Facebook chats or private messages.



TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL

Facebook Chat V4 Appid Example

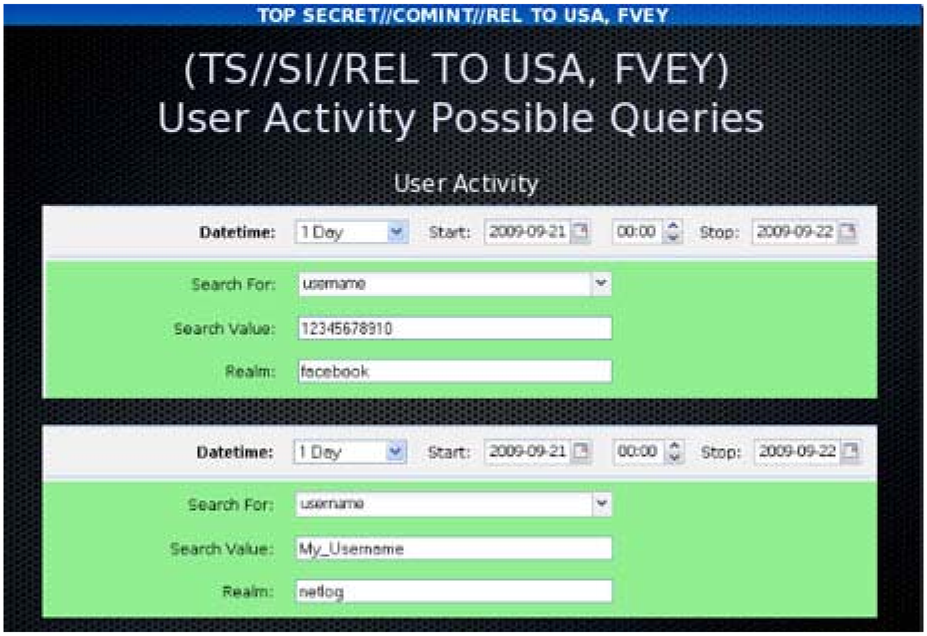
- DNI Presenter Display:

UIS Web Form Display

Form Fields	
msg_id	[redacted]
client_time	
to	
num_tabs	
pvs_time	
msg_text	dont u still recognize me?
post_form_id	[redacted] This information has been redacted
fb_dtsg	
post_form_id_source	
__a	
nctr[id]	
nctr[nid]	
nctr[ct]	

TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL

An analyst can monitor such Facebook chats by entering the Facebook user name and a date range into a simple search screen.



TOP SECRET//COMINT//REL TO USA, FVEY

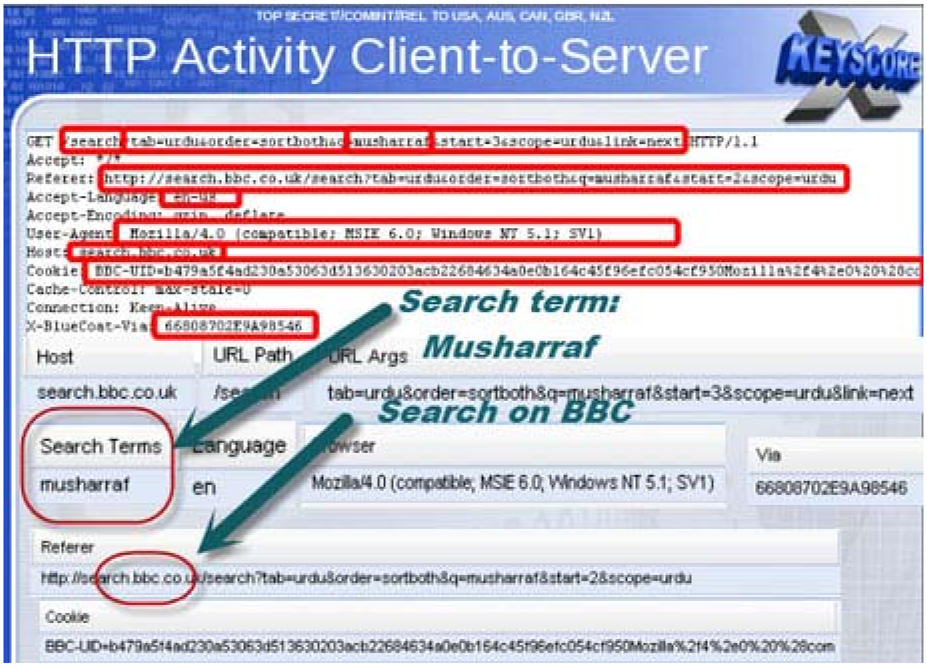
(TS//SI//REL TO USA, FVEY)
User Activity Possible Queries

User Activity

Datetime:	1 Day	Start:	2009-09-21	00:00	Stop:	2009-09-22
Search For:	username					
Search Value:	12345678910					
Realm:	facebook					

Datetime:	1 Day	Start:	2009-09-21	00:00	Stop:	2009-09-22
Search For:	username					
Search Value:	My_Username					
Realm:	netlog					

Analysts can search for internet browsing activities using a wide range of information, including search terms entered by the user or the websites viewed.



TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL

HTTP Activity Client-to-Server

```
GET /search?tab=urdu&order=sortboth&q=musharraf&start=3&scope=urdu&link=next HTTP/1.1
Accept: */*
Referer: http://search.bbc.co.uk/search?tab=urdu&order=sortboth&q=musharraf&start=2&scope=urdu
Accept-Language: en-us
Accept-Encoding: gzip, deflate
User-Agent: Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1)
Host: search.bbc.co.uk
Cookie: BBC-UID=b479a5f4ad230a53063d513630203acb22684634a0e0b164c45f96efc054cf950Mozilla%2f4%2e0%20%28co
Cache-Control: max-stale=0
Connection: Keep-Alive
X-BlueCoat-Via: 66808702E9A98546
```

Search term: Musharraf

Host	URL Path	URL Args
search.bbc.co.uk	/search	tab=urdu&order=sortboth&q=musharraf&start=3&scope=urdu&link=next

Search on BBC

Search Terms	Language	Browser	Via
musharraf	en	Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1)	66808702E9A98546

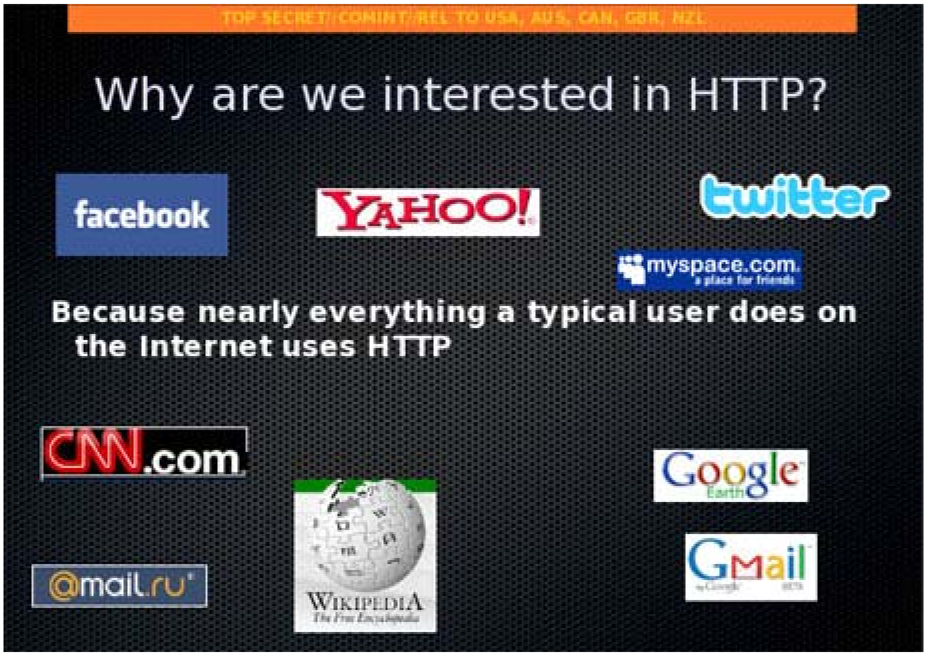
Referer

http://search.bbc.co.uk/search?tab=urdu&order=sortboth&q=musharraf&start=2&scope=urdu

Cookie

BBC-UID=b479a5f4ad230a53063d513630203acb22684634a0e0b164c45f96efc054cf950Mozilla%2f4%2e0%20%28com

As one slide indicates, the ability to search HTTP activity by keyword permits the analyst access to what the NSA calls "nearly everything a typical user does on the internet".



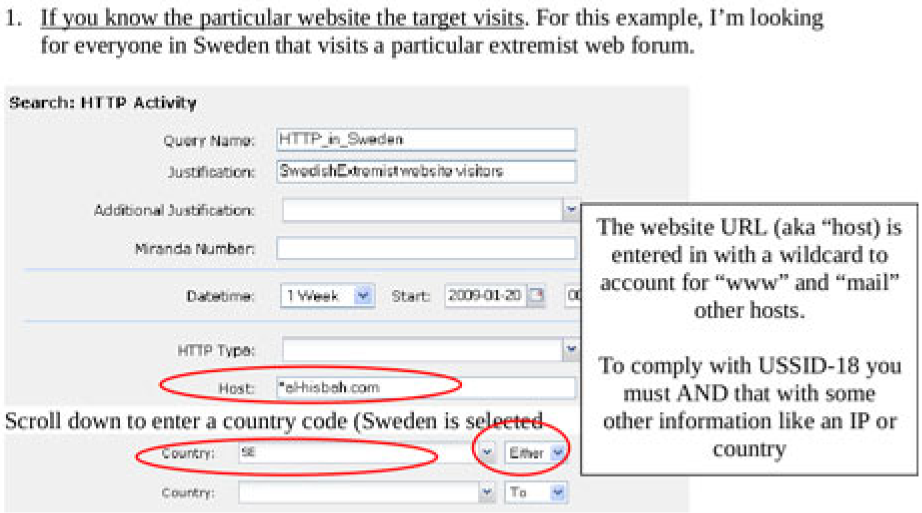
TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL

Why are we interested in HTTP?

Because nearly everything a typical user does on the Internet uses HTTP

facebook
YAHOO!
twitter
myspace.com a place for friends
CNN.com
Google Earth
Gmail
@mail.ru
WIKIPEDIA The Free Encyclopedia

The XKeyscore program also allows an analyst to learn the IP addresses of every person who visits any website the analyst specifies.



1. If you know the particular website the target visits. For this example, I'm looking for everyone in Sweden that visits a particular extremist web forum.

Search: HTTP Activity

Query Name:	HTTP_in_Sweden
Justification:	SwedishExtremistwebsite visitors
Additional Justification:	
Miranda Number:	
Datetime:	1 Week Start: 2009-01-20
HTTP Type:	
Host:	*alHisbah.com

Scroll down to enter a country code (Sweden is selected

Country:	SE	Either
Country:		To

The website URL (aka "host) is entered in with a wildcard to account for "www" and "mail" other hosts.

To comply with USSID-18 you must AND that with some other information like an IP or country

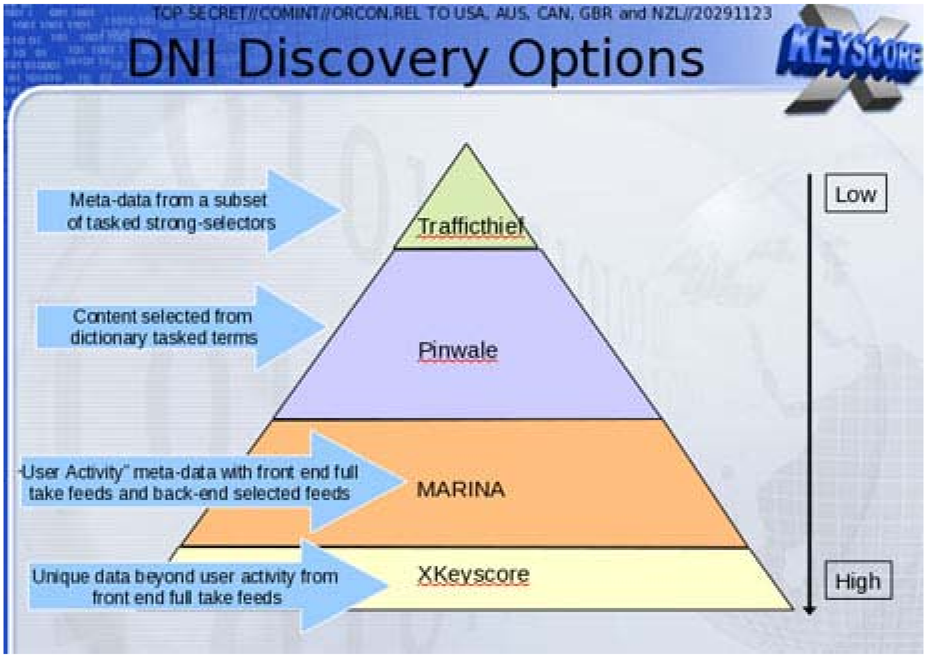
The quantity of communications accessible through programs such as XKeyscore is staggeringly large. One NSA report from 2007 estimated that there were 850bn "call events" collected and stored in the NSA databases, and close to 150bn internet records. Each day, the document says, 1-2bn records were added.

William Binney, a former NSA mathematician, said last year that the agency had "assembled on the order of 20tn transactions about US citizens with other US citizens", an estimate, he said, that "only was involving phone calls and emails". A 2010 Washington Post article reported that "every day, collection systems at the [NSA] intercept and store 1.7bn emails, phone calls and other type of communications."

The XKeyscore system is continuously collecting so much internet data that it can be stored only for short periods of time. Content remains on the system for only three to five days, while metadata is stored for 30 days. One document explains: "At some sites, the amount of data we receive per day (20+ terabytes) can only be stored for as little as 24 hours."

To solve this problem, the NSA has created a multi-tiered system that allows analysts to store "interesting" content in other databases, such as one named Pinwale which can store material for up to five years.

It is the databases of XKeyscore, one document shows, that now contain the greatest amount of communications data collected by the NSA.



TOP SECRET//COMINT//ORCON,REL TO USA, AUS, CAN, GBR and NZL//20291123

DNI Discovery Options

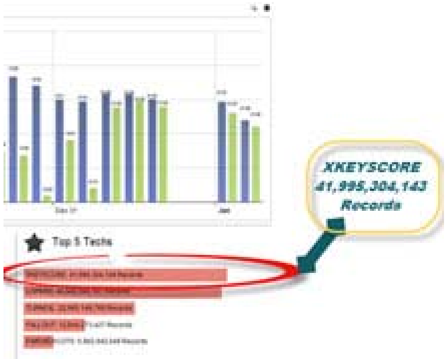
Meta-data from a subset of tasked strong-selectors

Content selected from dictionary tasked terms

"User Activity" meta-data with front end full take feeds and back-end selected feeds

Unique data beyond user activity from front end full take feeds

In 2012, there were at least 41 billion total records collected and stored in XKeyscore for a single 30-day period.



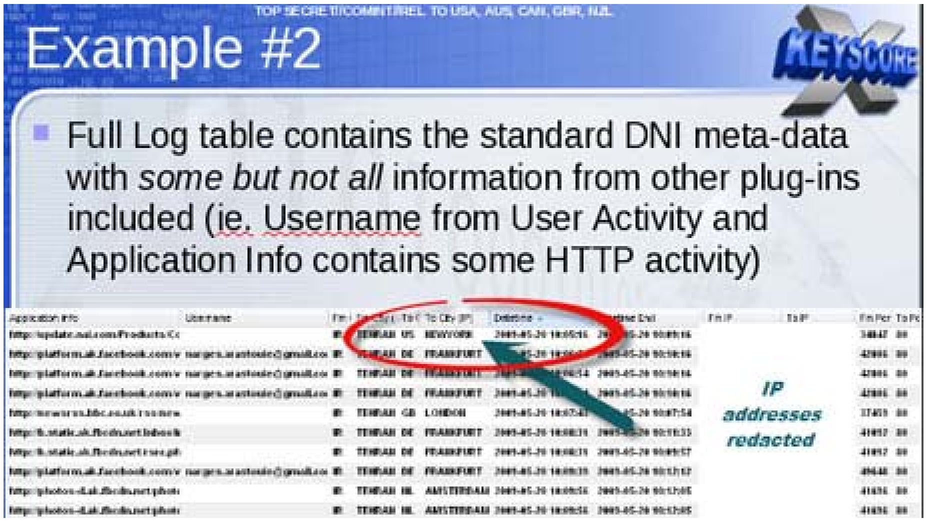
Legal v technical restrictions

While the Fisa Amendments Act of 2008 requires an individualized warrant for the targeting of US persons, NSA analysts are permitted to intercept the communications of such individuals without a warrant if they are in contact with one of the NSA's foreign targets.

The ACLU's deputy legal director, Jameel Jaffer, told the Guardian last month that national security officials expressly said that a primary purpose of the new law was to enable them to collect large amounts of Americans' communications without individualized warrants.

"The government doesn't need to 'target' Americans in order to collect huge volumes of their communications," said Jaffer. "The government inevitably sweeps up the communications of many Americans" when targeting foreign nationals for surveillance.

An example is provided by one XKeyscore document showing an NSA target in Tehran communicating with people in Frankfurt, Amsterdam and New York.



TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL

Example #2

- Full Log table contains the standard DNI meta-data with *some but not all* information from other plug-ins included (ie. Username from User Activity and Application Info contains some HTTP activity)

Application Info	Username	Fm	[illegible]	To C	To City (IP)	Datetime	[illegible]e End	Fm IP	To IP	Fm Por	To Po
http://update.aol.com/Products.C		IP	TEHRAN	US	NEWYORK	2009-05-20 18:05:16	[illegible]-20 00:09:16			34847	80
http://platform.ak.facebook.com/v	narges.arastouie@gmail.co	IP	[illegible]	DE	FRANKFURT	[illegible]	2009-05-20 00:10:16			42806	80
http://platform.ak.facebook.com/v	narges.arastouie@gmail.co	IP	TEHRAN	DE	[illegible]	[illegible]:54	2009-05-20 00:10:16			42806	80
http://platform.ak.facebook.com/v	narges.arastouie@gmail.co	IP	TEHRAN	DE	FRANKFURT	2009-05-20 [illegible]	2009-05-20 00:10:16			42806	80
http://newsrss.bbc.co.uk/rss/new		IP	TEHRAN	GB	LONDON	2009-05-20 18:07:54	[illegible]-20 00:07:54			37459	80
http://b.static.ak.fbcdn.net/rsrc		IP	TEHRAN	DE	FRANKFURT	2009-05-20 18:08:31	[illegible] 00:11:33			41857	80
http://b.static.ak.fbcdn.net/rsrc.ph		IP	TEHRAN	DE	FRANKFURT	2009-05-20 18:08:31	2009-05-20 00:09:57			41857	80
http://platform.ak.facebook.com/v	narges.arastouie@gmail.co	IP	TEHRAN	DE	FRANKFURT	2009-05-20 18:09:35	2009-05-20 00:17:57			49646	80
http://photos-d.ak.fbcdn.net/phot		IP	TEHRAN	NL	AMSTERDAM	2009-05-20 18:09:56	2009-05-20 00:12:05			41636	80
http://photos-d.ak.fbcdn.net/phot		IP	TEHRAN	NL	AMSTERDAM	2009-05-20 18:09:56	2009-05-20 00:12:05			41636	80

IP addresses redacted

In recent years, the NSA has attempted to segregate exclusively domestic US communications in separate databases. But even NSA documents acknowledge that such efforts are imperfect, as even purely domestic communications can travel on foreign systems, and NSA tools are sometimes unable to identify the national origins of communications.

Moreover, all communications between Americans and someone on foreign soil are included in the same databases as foreign-to-foreign communications, making them readily searchable without warrants.

Some searches conducted by NSA analysts are periodically reviewed by their supervisors within the NSA. "It's very rare to be questioned on our searches," Snowden told the Guardian in June, "and even when we are, it's usually along the lines of: 'let's bulk up the justification'."

In a letter this week to senator Ron Wyden, director of national intelligence James

Clapper acknowledged that NSA analysts have exceeded even legal limits as interpreted by the NSA in domestic surveillance.

Acknowledging what he called "a number of compliance problems", Clapper attributed them to "human error" or "highly sophisticated technology issues" rather than "bad faith".

However, Wyden said on the Senate floor on Tuesday: "These violations are more serious than those stated by the intelligence community, and are troubling."

In a statement to the Guardian, the NSA said: "NSA's activities are focused and specifically deployed against – and only against – legitimate foreign intelligence targets in response to requirements that our leaders need for information necessary to protect our nation and its interests.

"XKeyscore is used as a part of NSA's lawful foreign signals intelligence collection system.

"Allegations of widespread, unchecked analyst access to NSA collection data are simply not true. Access to XKeyscore, as well as all of NSA's analytic tools, is limited to only those personnel who require access for their assigned tasks ... In addition, there are multiple technical, manual and supervisory checks and balances within the system to prevent deliberate misuse from occurring."

"Every search by an NSA analyst is fully auditable, to ensure that they are proper and within the law.

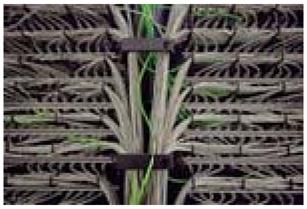
"These types of programs allow us to collect the information that enables us to perform our missions successfully – to defend the nation and to protect US and allied troops abroad."

More from the Guardian

What's this?

Why have so many liberals

US should leave Edward



Teaching the NSA files and

Mafia boss on the run since

been silent about NSA spying?
02 Aug 2013

Snowden alone
06 Aug 2013

data security – news and resources round up
04 Aug 2013

1994 arrested in London
08 Aug 2013

More from around the web

What's this?

What am I paying for in the price of a gallon of gasoline?
(ExxonMobil Perspectives Blog)

O.J. Simpson wins parole in robbery case
(Sports Illustrated)

17 Things You Should Always Buy New
(Daily Finance)

Case in Point: Pitched roofs or math — what matters for solar?
(The Washington Post)

;