Glenn Greenwald on security and liberty The NSA files

• This article is more than 10 years old

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



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

Glenn Greenwald

Wed 31 Jul 2013 08.56 EDT

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."

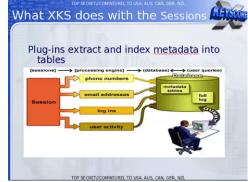
But training materials for XKeyscore detail how analysts can use it and other systems to mine enormous agency databases by filling in a simple on-screen form giving only a broad justification for the search. The request is not reviewed by a court or any NSA personnel before it is processed.

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.



🗅 KS1 Photograph: Guardian

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.

7 (7. of 27) Rt Page Width

Email Addresses Query:

One of the most common queries is (you guessed ii) 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...

arch: Email Addresses		
Quary Name:	abuj hod	
Justification:	ctitarget in in africa	
Additional Justification:		*
Miranda Number:	6	
Datebine:	1 Month 😁 Start: 2009-12-24 📑	00:00 0 1
Email Username:	abujhad	1
@Domain:	yahoo com	



🗅 KS3edit2 Photograph: Guardian

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)	For	reign	Facto	rs	
30	ald Ta	rgeting	Request -	Microsi	oft Internet Explorer
260	164	Mass	Instates	Teole	Maks

Procedence Procedence Procedence Procedence Reason Comments	2			
Precedence Forwarding Precedence Reason	4		-1	
Precedence Reason			-1	
Reason			1 1 1	
Commente		ţ		
Comments		Į.		
			83	
- Special Authorization	-Select a Speci	si Auth-	_	
E FAA Foreign Governments	Cert (* Not valid to Tagk - Require	i pristwi si ekto be		Validate Rem
Foreign Intel Purpose -S	elect a Value-			
Foreign Factor In c	firect contact witgt overseas, n	o into to show propose	id tot in U.S.	1
Foreignness Source ID	elect à Value-			Foreign
Foreignness Explanation The	e person has stated that he is i man intelligence source indice e person is a user of storage n reign govt indicates that the pe	les person is localed o redia seized outside the	whice the U.S. e U.S.	
Zacode	one number country code indic	ates person is located	outside the U.S.	Select a
Start Date Sig End Date Op	one number is registered in a c SINT reporting confirms person en source information indicate	is located outside the I s person is located outs	US. side the US -	For dignness Factor
Ne	Work, machine or tech info indi firect contact what overseas, n			

🗱 KS4 Photograph: Guardian

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.

Facebo	ook Chat V4 Appid Example	ALTING
DNI	Presenter Display:	
Ser.	an fealait dhanna dh' Palath - an 1949:Dein - Danas ann an taona a' Channa an Dhan a' La	
	UIS Web Form Display	
TOP J.		GBR,

🗅 KS55edit Photograph: Guardian

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

(TS//S User A								s
		Us	ier Ac	tivity				
Datetime:	1 Dey	*	Start:	2009-09-21	3	00:00 🗘	Stop:	2009-09-22
Search For:	usemame				۲			
Search Value:	123456789	10						
Realm:	facebook							
Datetime:	1 Dey	~	Start:	2009-09-21	3	00:00 🗘	Stop:	2009-09-22
Search For:	usemane				~			
Search Value:	My_Usem	sme						
Realm:	netlog							

🗅 KS6 Photograph: Guardian

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

HTTP Activity Client-to-Server	(13 Sade
Tel: Enter: Relandanceder-settloffe, Bushares, Starts-Staroge-up dui takewer Intti-/ Protect: Relandanceder-settloffe, Bushares, Starts-Staroge-up dui takewer Intti-/ Protect: Relandanceder-settloffe, Bushares, Starts-Staroge-up dui takewer Intti-/ Protect: Relandanceder-settloffe, Bushares, Starts-Staroge-up dui takewer Intti-/ Relandanceder-settloffe, Bushares, Starts-Star	24scope≈urdu
search bbc.co.uk /search Search Terms banguage www.	scope=urdu&link=next Via
musharrat en Mozilial 4.0 (compatible; MSE 6.0; Windows NT 5.1; SV1) Referer	66808702E9A98546
http://liserich.bbc.co.//search?tab-urduℴ=contboth&q=mushamaf&start=28scope=urdu Coole BBC-U/D-b47965f4ast20as506505351350203acb22684634a0e0te164-65965erc054c1950MozileV/214%2	10%20%28com
KS7 Photograph: Guardian	

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".



🖾 KS8 Photograph: Guardian

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

earch: HTTP Activity		
Query Name:	HTTP_is_Sweden	
Justification:	SwedishExtremist website visitors	
Additional Justification: Niranda Number: Datetime:	1 Week 💌 Start: 2009-01-25 📑 00	The website URL (aka "host) is entered in with a wildcard to account for "www" and "mail" other hosts.
HTTP Type:	*elflisbah.com	To comply with USSID-18 you must AND that with some

KS9 Photograph: Guardian

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.

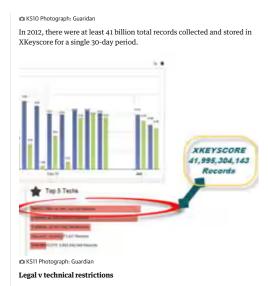
William Binney, a former NSA mathematician, said last year that the agency had "assembled on the order of 20th transactions about US citizens with other US citizens", an estimate, he said, that "only was involving phone calls and emails". A 2010 Washington Post article reported that "every day, collection 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.





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.

Full Log table co				10 1 N		
Full Log lable co	ontains	the s	standar	d DNI	meta-da	ata
with some but n	ot all in	form	ation fr	om oth	or plug	inc
						IIIS
included (ie, Us	ername	e fron	User	Activity	/ and	
						-
Application Info	contair	ns so	me HT	TP act	ivity)	
Application info Upertaine	THI CALLER TO	C To City (P)	Deletine -	ADM EVI	Init Init	Enter Tark
	PROFESSION OF	FC TO City (IP) 5. HEVYYORK	Deletine - Joan.at.jo 18:45:46	19 15-29 10:49:16	FOF DP	Tenther Tarks
Intge Tupilate mal.com/Products.Cc	Res B SHEAR DE	FT TO CRY (P) B REVEYORS	Delatine - Joanual Je Incides 85-Je Incides	2010 EVI 2010 E5-20 5049116 2005-05-20 5021816	rnir fair	
Intge Tupilate mal.com/Products.Cc			15-20 HERE	2014 EVE 2014 S.20 States 2015 45:20 States 2015 45:20 States 2015 45:20 States		34842 00
http://update.nai.com/Products.Cc http://platform.ak.facebook.com/v_nargers.arastonie@gmai	Aco & TENDAR DE	ITANTURI I	15-20 HERE	200 0253 0259516 200 02520 0259516 200545220 0259516 200545220 0259616 200545220 0259616	IP	34842 00
Miller sopelate nai com Products. Cc Miller platformals facebook zemis narges arastonie () gmai Miller platformals facebook zemis narges arastonie () gmai Miller platformals facebook zemis narges arastonie () gmai	Aco & TENDAR DE	ITANTURI I	HEARING THE REAL	2005-05-20 50:10:16 2005-05-20 50:10:16 2005-05-20 50:10:16		34842 00 42985 00 42985 00 42985 00
Impringstate nations Products Co Impripatifermula Jacobook comv marges anastonie 2 gma Impripatifermula Jacobook comv marges anastonie 2 gma Impripatifermula Jacobook comv marges anastonie 2 gma Impripatifermula Jacobook com v marges anastonie 2 gma Imprime warenable couch is to imm	Aco & TENDAR DE	ITANTURI I	5-30 10-00 2000 - 2000 - 2000 2000 - 2000 - 2000 2000 - 2000 - 2000 - 2000 2000 - 2000 - 2000 - 2000	2005-05-20 50:10:16 2005-05-20 50:10:16 2005-05-20 50:10:16	IP addresse	34842 00 42885 00 42885 00 42885 00 42885 00 12855 00
Impi inplane naiseon Produits. Ce Impi inplatema di Jacobodi Jenov nai gestar antonico) gran Impi inplatforma di Jacobodi Jenov na gesta attonico) gran Impi inplatforma di Jacobodi Jenov na gesta attonico) gran Impi in provinsi attonico di Impi in provinsi attonico) gran Impi in provinsi attonico di Impi in provinsi Impi in producta di Technologi Impi in provinsi Impi in provinsi Impi in provinsi Impi Impi Impi Impi Impi Impi Impi Impi	Aco & TENDAR DE	ITANTURI I	1538 18384 2005 1990 1990 19 2005 4539 10 2005 4539 100024 2005 4539 100021	1985-85-28 98:19816 2985-85-28 98:19816 2985-85-28 98:19816 26-29 98:19154	IP	34842 00 42885 00 42885 00 42885 00 42885 00 12855 00
http://splate.nsi.com/Fradacts.Cc http:/platform.uk.facebook.com/r narges.iarantonie@gmai http:/platform.uk.facebook.com/r narges.arantonie@gmai	ALCO III TEMPAH DE ALCO III TEMPAH DE III TEMPAH DE III TEMPAH DE III TEMPAH DE	ITANTURI I	5-20 10200 2005-05-20 10 2005-05-20 10202-0 2005-05-20 102062-0 2005-05-20 102062-0 2005-05-20 102062-0	1995-05-20 50:10:16 2995-05-20 50:10:16 2995-05-20 50:10:16 15-20 50:07:54 2995-05-20 50:11:53	IP addresse	34842 00 42885 00 42885 00 22865 00 2285 17455 00 41052 00
Imp: spade nation Products Cr Imp: platemaa Jacobo Aron v narges arasteled; grad Imp: platemaa Jacobo Aron v narges arasteled; grad Imp: platesaa Jacobo Aron v narges arasteled; grad Imp: b.ataka Aroba Aron v narges arasteled; grad Imp: b.ataka Aroba Aroba Iones Imp: b.ataka Aroba Iones Iones Imp: b.ataka Iones Iones Iones Imp: b.ataka Iones	ALCO III TEMPAH DE ALCO III TEMPAH DE III TEMPAH DE III TEMPAH DE III TEMPAH DE	ITANTURI I	25-26 18205-14 2085-45-20 18 2085-45-20 18202-40 2085-45-20 18206218 2085-45-20 18206218 2085-45-20 18206218 2085-45-20 18206219	1995-45-29 50:16:16 2995-45-29 50:16:16 2995-45-29 50:16:16 2995-45-29 50:16:16 1-29 50:07.54 2995-26-29 50:16:15 2995-45-29 50:06:57	IP addresse	3442 00 42965 00 42965 00 42965 00 42965 00 42965 00 41952 00 41952 00

🛱 KS12 Photograph: Guardian

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 on this story

BT and Vodafone among telecoms companies passing details to GCHQ GCHQ accused of selling its services after revelations of funding by NSA Flexible laws and weak oversight give GCHQ room for manoeuvre





Most viewed