

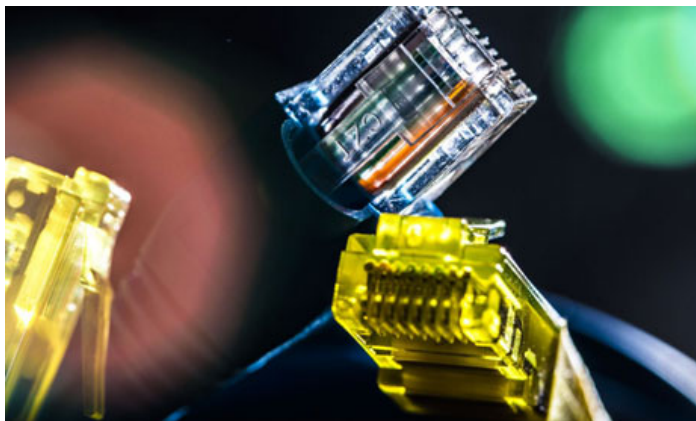
# NSA stores metadata of millions of web users for up to a year, secret files show

- Vast amounts of data kept in repository codenamed Marina
- Data retained regardless of whether person is NSA target
- Material used to build 'pattern-of-life' profiles of individuals
- [What is metadata? Find out with our interactive guide](#)

James Ball

theguardian.com, Monday 30 September 2013 17.35 BST

[Jump to comments \(...\)](#)



Any computer metadata picked up by NSA collection systems is routed to the Marina database, the guide explains. Photograph: Felix Clay

The National Security Agency is storing the online metadata of millions of internet users for up to a year, regardless of whether or not they are persons of interest to the agency, top secret documents reveal.

Metadata provides a record of almost anything a user does online, from browsing history – such as map searches and websites visited – to account details, email activity, and even some account passwords. This can be used to build a detailed picture of an individual's life.

The Obama administration has repeatedly stated that the NSA keeps only the content of messages and communications of people it is intentionally targeting – but internal documents reveal the agency retains vast amounts of metadata.

An introductory guide to digital network intelligence for NSA field agents, included in documents disclosed by former contractor Edward Snowden, describes the agency's metadata repository, codenamed Marina. Any computer metadata picked up by NSA collection systems is routed to the Marina database, the guide explains. Phone metadata is sent to a separate system.

"The Marina metadata application tracks a user's browser experience, gathers contact information/content and develops summaries of target," the analysts' guide explains. "This tool offers the ability to export the data in a variety of formats, as well as create various charts to assist in pattern-of-life development."

The guide goes on to explain Marina's unique capability: "Of the more distinguishing features, Marina has the ability to look back on the last 365 days' worth of DNI metadata seen by the Sigint collection system,

Share

Tweet this

Pint it

Email



Article history

## The NSA Files: Decoded



What the surveillance revelations mean for you

### World news

The NSA files · NSA · Surveillance · Privacy · United States · US politics

### Technology

Data protection · Internet

### Law

US constitution and civil liberties

### More news

#### More on this story



NSA chief admits agency tracked US cell phone locations in secret tests

Keith Alexander insists pilot programs were not used for analysis but that tracking location data 'may be a future requirement'

Former NSA chief: western intelligence agencies must be more transparent

Glenn Greenwald and Janine Gibson: 10 highlights from their Reddit AMA

## Today's best video



### European football papers review

James Richardson brings you all the continental reaction to the European Champions League last 16 draw



### Inside the mind of Lee Rigby's killer

Forensic psychiatrists and Michael Adebolajo's brother discuss what might

have motivated him to kill soldier Lee Rigby



### Seattle gunman wrestled to ground by bus passengers

CCTV footage of the moment a gun-wielding man is disarmed by passengers on a bus in Seattle



### The Guardian Film Show

The team discuss American Hustle, All is Lost, The Secret Life of Walter Mitty and

Anchorman 2 in our 100th episode

## GuardianWitness



regardless whether or not it was tasked for collection." [Emphasis in original.]

On Saturday, the [New York Times](#) reported that the NSA was using its metadata troves to build profiles of US citizens' social connections, associations and in some cases location, augmenting the material the agency collects with additional information bought in from the commercial sector, which is not subject to the same legal restrictions as other data.

The ability to look back on a full year's history for any individual whose data was collected – either deliberately or incidentally – offers the NSA the potential to find information on people who have later become targets. But it relies on storing the personal data of large numbers of internet users who are not, and never will be, of interest to the US intelligence community.

Marina aggregates NSA metadata from an array of sources, some targeted, others on a large scale. [Programs such as Prism](#) – which operates through legally compelled "partnerships" with major internet companies – allow the NSA to obtain content and metadata on thousands of targets without individual warrants.

The NSA also collects enormous quantities of metadata from the fibre-optic cables that make up the backbone of the internet. The agency has placed taps on undersea cables, and is given access to internet data through partnerships with American telecoms companies.

About 90% of the world's online communications cross the US, giving the NSA what it calls in classified documents a "home-field advantage" when it comes to intercepting information.

By confirming that all metadata "seen" by NSA collection systems is stored, the Marina document suggests such collections are not merely used to filter target information, but also to store data at scale.

A sign of how much information could be contained within the repository comes from a [document voluntarily disclosed by the NSA in August](#), in the wake of the first tranche of revelations from the Snowden documents.

The [seven-page document](#), titled "The National Security Agency: Missions, Authorities, Oversight and Partnerships", says the agency "touches" 1.6% of daily internet traffic – an estimate which is not believed to include large-scale internet taps operated by GCHQ, the NSA's UK counterpart.

The document cites figures from a major tech provider that the internet carries 1.826 petabytes of information per day. One petabyte, [according to tech website Gizmodo](#), is equivalent to over 13 years of HDTV video.

"In its foreign intelligence mission, NSA touches about 1.6% of that," the document states. "However, of the 1.6% of the data, only 0.025% is actually selected for review.

"The net effect is that NSA analysts look at 0.00004% of the world's traffic in conducting their mission – that's less than one part in a million."

However, critics were skeptical of the reassurances, because large quantities of internet data is represented by music and video sharing, or large file transfers – content which is easy to identify and dismiss without entering it into systems. Therefore, the NSA could be picking up a much larger percentage of internet traffic that contains communications and browsing activity.

Journalism professor and internet commentator [Jeff Jarvis](#) noted: "[By] very rough, beer-soaked-napkin numbers, the NSA's 1.6% of net traffic would be half of the communication on the net. That's one helluva lot of 'touching'."

Much of the NSA's data collection is carried out under section 702 of the Fisa Amendments Act. This provision allows for the collection of data without individual warrants of communications, where at least one end of the conversation, or data exchange, involves a non-American located outside the US at the time of collection.

The NSA is required to "minimize" the data of US persons, but is permitted to keep US communications where it is not technically possible to remove them, and also to keep and use any "inadvertently" obtained US communications if they contain intelligence material, evidence of a crime, or if they are encrypted.



#### GuardianWitness Awards 2014

Upload your contributions to GuardianWitness before 20th January to be entered into our reader awards



#### Extreme weather in the UK

The latest storms have brought power cuts, transport problems and flood warnings across parts of the UK, share your videos and images and help shape our coverage



#### Recipe swap: smoothies and juices

Share your recipes for a delicious vitamin boost after the over-indulgences of the Christmas and New Year holiday

#### Send us your assignment ideas

Do you have an idea for an assignment you think should run on GuardianWitness? [Let us know](#).

## THE NSA FILES



**Live:** Follow NSA-related developments as controversy over leaks continues to make headlines

### On World news

Most viewed

Latest

#### Last 24 hours



1. Mikhail Khodorkovsky 'exhausted but happy to be free' after Putin's pardon

2. Ugandan MPs rush through draconian laws against homosexuality

3. Nelson Mandela 'received weapons training from Mossad agents in 1962'

4. Barack Obama rejects claims 2013 has been worst year of his presidency

5. \$10m NSA contract with security firm RSA led to encryption 'back door'

More most viewed

#### Last 24 hours



1. Chinese hotelier among four feared dead in France helicopter crash

2. Eyewitness: Hula Valley, Israel

3. Iraq suicide bombers kill at least 18 military officers

4. Obama: now is not the time for new Iran sanctions - video

5. NSA surveillance issue is being taken 'very seriously', says Obama - video

All today's stories

The Guardian has also revealed the existence of a so-called "backdoor search loophole", a 2011 rule change that allows NSA analysts to search for the names of US citizens, under certain circumstances, in mass-data repositories collected under section 702.

According to the New York Times, NSA analysts were told that metadata could be used "without regard to the nationality or location of the communicants", and that Americans' social contacts could be traced by the agency, providing there was some foreign intelligence justification for doing so.

The Guardian approached the NSA with four specific questions about the use of metadata, including a request for the rationale behind storing 365 days' worth of untargeted data, and an estimate of the quantity of US citizens' metadata stored in its repositories.

But the NSA did not address any of these questions in its response, providing instead a statement focusing on its foreign intelligence activities.

"NSA is a foreign intelligence agency," the statement said. "NSA's foreign intelligence activities are conducted pursuant to procedures approved by the US attorney general and the secretary of defense, and, where applicable, the foreign intelligence surveillance (Fisa) court, to protect the privacy interests of Americans.

"These interests must be addressed in the collection, retention, and dissemination of any information. Moreover, all queries of lawfully collected data must be conducted for a foreign intelligence purpose."

It continued: "We know there is a false perception out there that NSA listens to the phone calls and reads the email of everyday Americans, aiming to unlawfully monitor or profile US citizens. It's just not the case.

"NSA's activities are directed against foreign intelligence targets in response to requirements from US leaders in order to protect the nation and its interests from threats such as terrorism and the proliferation of weapons of mass destruction."

[Share](#) [Tweet this](#) [Email](#)

## Comments

[Click here to join the discussion.](#)

We can't load the discussion on [theguardian.com](#) because you don't have [JavaScript enabled](#).

[Share](#)

[Tweet this](#)

## guardianbookshop

### This week's bestsellers



1. **Stoner**  
by John L. Williams  
£7.19

2. **Stephen Ward Was Innocent, OK**  
by Geoffrey Robertson £9.74

3. **Guardian Quick Crosswords 5 & 6**  
£8.00

4. **Tales from the Secret Footballer**  
£7.99

5. **Ammonites and Leaping Fish**  
by Penelope Lively £10.99

Search the Guardian bookshop

## guardianjobs

Find the latest jobs in your sector:

[Arts & heritage](#)

[Health](#)

[Charities](#)

[Marketing & PR](#)

[Education](#)

[Media](#)

[Environment](#)

[Sales](#)

[Government](#)

[Senior executive](#)

[Graduate](#)

[Social care](#)

[Browse all jobs](#)