The Washington Post

National Security

In NSA-intercepted data, those not targeted far outnumber the foreigners who are

Files provided by Snowden show extent to which ordinary Web users are caught in the net

By Barton Gellman, Julie Tate and Ashkan Soltani July 5, 2014

Ordinary Internet users, American and non-American alike, far outnumber legally targeted foreigners in the communications intercepted by the National Security Agency from U.S. digital networks, according to a four-month investigation by The Washington Post.

Nine of 10 account holders found in a large cache of intercepted conversations, which former NSA contractor Edward Snowden provided in full to The Post, were not the intended surveillance targets but were caught in a net the agency had cast for somebody else.

Many of them were Americans. Nearly half of the surveillance files, a strikingly high proportion, contained names, e-mail addresses or other details that the NSA marked as belonging to U.S. citizens or residents. NSA analysts masked, or "minimized," more than 65,000 such references to protect Americans' privacy, but The Post found nearly 900 additional e-mail addresses, unmasked in the files, that could be strongly linked to U.S.

citizens or U.S.residents.

(How 160,000 intercepted conversations led to The Post's latest NSA story)

The surveillance files highlight a policy dilemma that has been aired only abstractly in public. There are discoveries of considerable intelligence value in the intercepted messages — and collateral harm to privacy on a scale that the Obama administration has not been willing to address.

Among the most valuable contents — which The Post will not describe in detail, to avoid interfering with ongoing operations — are fresh revelations about a secret overseas nuclear project, double-dealing by an ostensible ally, a military calamity that befell an unfriendly power, and the identities of aggressive intruders into U.S. computer networks.

Months of tracking communications across more than 50 alias accounts, the files show, led directly to the 2011 capture in Abbottabad of Muhammad Tahir Shahzad, a Pakistan-based bomb builder, and Umar Patek, a suspect in a 2002 terrorist bombing on the Indonesian island of Bali. At the request of CIA officials, The Post is withholding other examples that officials said would compromise ongoing operations.

(Transcript: Q&A with Barton Gellman)

Many other files, described as useless by the analysts but nonetheless retained, have a startlingly intimate, even voyeuristic quality. They tell stories of love and heartbreak, illicit sexual liaisons, mental-health crises, political and religious conversions, financial anxieties and disappointed hopes. The daily lives of more than 10,000 account holders who were not targeted are catalogued and recorded nevertheless.

In order to allow time for analysis and outside reporting, neither Snowden nor The Post has disclosed until now that he obtained and shared the content of intercepted communications. The cache Snowden provided came from domestic NSA operations under the broad authority granted by Congress in 2008 with amendments to the Foreign Intelligence

Surveillance Act. FISA content is generally stored in closely controlled data repositories, and for more than a year, senior government officials have depicted it as beyond Snowden's reach.

The Post reviewed roughly 160,000 intercepted e-mail and instant-message conversations, some of them hundreds of pages long, and 7,900 documents taken from more than 11,000 online accounts.

The material spans President Obama's first term, from 2009 to 2012, a period of exponential growth for the NSA's domestic collection.

Taken together, the files offer an unprecedented vantage point on the changes wrought by Section 702 of the FISA amendments, which enabled the NSA to make freer use of methods that for 30 years had required probable cause and a warrant from a judge. One program, code-named PRISM, extracts content stored in user accounts at Yahoo, Microsoft, Facebook, Google and five other leading Internet companies. Another, known inside the NSA as Upstream, intercepts data on the move as it crosses the U.S. junctions of global voice and data networks.

No government oversight body, including the Justice Department, the Foreign Intelligence Surveillance Court, intelligence committees in Congress or the president's Privacy and Civil Liberties Oversight Board, has delved into a comparably large sample of what the NSA actually collects — not only from its targets but also from people who may cross a target's path.

Among the latter are medical records sent from one family member to another, résumés from job hunters and academic transcripts of schoolchildren. In one photo, a young girl in religious dress beams at a camera outside a mosque.

Scores of pictures show infants and toddlers in bathtubs, on swings, sprawled on their backs and kissed by their mothers. In some photos, men show off their physiques. In others, women model lingerie, leaning suggestively into a webcam or striking risque poses in shorts and bikini tops.

“None of the hits that were received were relevant,” two Navy cryptologic technicians write in one of many summaries of nonproductive surveillance. “No additional information,” writes a civilian analyst. Another makes fun of a suspected kidnapper, newly arrived in Syria before the current civil war, who begs for employment as a janitor and makes wide-eyed observations about the state of undress displayed by women on local beaches.

By law, the NSA may “target” only foreign nationals located overseas unless it obtains a warrant based on probable cause from a special surveillance court. For collection under PRISM and Upstream rules, analysts must state a reasonable belief that the target has information of value about a foreign government, a terrorist organization or the spread of nonconventional weapons.

Most of the people caught up in those programs are not the targets and would not lawfully qualify as such. “Incidental collection” of third-party communications is inevitable in many forms of surveillance, but in other contexts the U.S. government works harder to limit and discard irrelevant data. In criminal wiretaps, for example, the FBI is supposed to stop listening to a call if a suspect’s wife or child is using the phone.

There are many ways to be swept up incidentally in surveillance aimed at a valid foreign target. Some of those in the Snowden archive were monitored because they interacted directly with a target, but others had more-tenuous links.

If a target entered an online chat room, the NSA collected the words and identities of every person who posted there, regardless of subject, as well as every person who simply “lurked,” reading passively what other people wrote.

“1 target, 38 others on there,” one analyst wrote. She collected data on them all.

In other cases, the NSA designated as its target the Internet protocol, or

IP, address of a computer server used by hundreds of people.

The NSA treats all content intercepted incidentally from third parties as permissible to retain, store, search and distribute to its government customers. Raj De, the agency's general counsel, has testified that the NSA does not generally attempt to remove irrelevant personal content, because it is difficult for one analyst to know what might become relevant to another.

The Obama administration declines to discuss the scale of incidental collection. The NSA, backed by Director of National Intelligence James R. Clapper Jr., has asserted that it is unable to make any estimate, even in classified form, of the number of Americans swept in. It is not obvious why the NSA could not offer at least a partial count, given that its analysts routinely pick out "U.S. persons" and mask their identities, in most cases, before distributing intelligence reports.

If Snowden's sample is representative, the population under scrutiny in the PRISM and Upstream programs is far larger than the government has suggested. In a June 26 "transparency report," the Office of the Director of National Intelligence disclosed that 89,138 people were targets of last year's collection under FISA Section 702. At the 9-to-1 ratio of incidental collection in Snowden's sample, the office's figure would correspond to nearly 900,000 accounts, targeted or not, under surveillance.

'He didn't get this data'

U.S. intelligence officials declined to confirm or deny in general terms the authenticity of the intercepted content provided by Snowden, but they made off-the-record requests to withhold specific details that they said would alert the targets of ongoing surveillance. Some officials, who declined to be quoted by name, described Snowden's handling of the sensitive files as reckless.

In an interview, Snowden said "primary documents" offered the only path to a concrete debate about the costs and benefits of Section 702 surveillance. He did not favor public release of the full archive, he said, but

he did not think a reporter could understand the programs "without being able to review some of that surveillance, both the justified and unjustified."

"While people may disagree about where to draw the line on publication, I know that you and The Post have enough sense of civic duty to consult with the government to ensure that the reporting on and handling of this material causes no harm," he said.

In Snowden's view, the PRISM and Upstream programs have "crossed the line of proportionality."

"Even if one could conceivably justify the initial, inadvertent interception of baby pictures and love letters of innocent bystanders," he added, "their continued storage in government databases is both troubling and dangerous. Who knows how that information will be used in the future?"

For close to a year, NSA and other government officials have appeared to deny, in congressional testimony and public statements, that Snowden had any access to the material.

As recently as May, shortly after he retired as NSA director, Gen. Keith Alexander denied that Snowden could have passed FISA content to journalists.

"He didn't get this data," Alexander told a New Yorker reporter. "They didn't touch —"

"The operational data?" the reporter asked.

"They didn't touch the FISA data," Alexander replied. He added, "That database, he didn't have access to."

Robert S. Litt, the general counsel for the Office of the Director of National Intelligence, said in a prepared statement that Alexander and other officials were speaking only about "raw" intelligence, the term for intercepted content that has not yet been evaluated, stamped with

classification markings or minimized to mask U.S. identities.

“We have talked about the very strict controls on raw traffic, the training that people have to have, the technological lockdowns on access,” Litt said. “Nothing that you have given us indicates that Snowden was able to circumvent that in any way.”

In the interview, Snowden said he did not need to circumvent those controls, because his final position as a contractor for Booz Allen at the NSA’s Hawaii operations center gave him “unusually broad, unescorted access to raw SIGINT [signals intelligence] under a special ‘Dual Authorities’ role,” a reference to Section 702 for domestic collection and Executive Order 12333 for collection overseas. Those credentials, he said, allowed him to search stored content — and “task” new collection — without prior approval of his search terms.

“If I had wanted to pull a copy of a judge’s or a senator’s e-mail, all I had to do was enter that selector into XKEYSCORE,” one of the NSA’s main query systems, he said.

The NSA has released an e-mail exchange acknowledging that Snowden took the required training classes for access to those systems.

‘Minimized U.S. president’

At one level, the NSA shows scrupulous care in protecting the privacy of U.S. nationals and, by policy, those of its four closest intelligence allies — Britain, Australia, Canada and New Zealand.

More than 1,000 distinct “minimization” terms appear in the files, attempting to mask the identities of “possible,” “potential” and “probable” U.S. persons, along with the names of U.S. beverage companies, universities, fast-food chains and Web-mail hosts.

Some of them border on the absurd, using titles that could apply to only one man. A “minimized U.S. president-elect” begins to appear in the files in early 2009, and references to the current “minimized U.S. president”

appear 1,227 times in the following four years.

Even so, unmasked identities remain in the NSA's files, and the agency's policy is to hold on to "incidentally" collected U.S. content, even if it does not appear to contain foreign intelligence.

In one exchange captured in the files, a young American asks a Pakistani friend in late 2009 what he thinks of the war in Afghanistan. The Pakistani replies that it is a religious struggle against 44 enemy states.

Startled, the American says "they, ah, they arent heavily participating . . . its like . . . in a football game, the other team is the enemy, not the other teams waterboy and cheerleaders."

"No," the Pakistani shoots back. "The ther teams water boy is also an enemy. it is law of our religion."

"haha, sorry thats kind of funny," the American replies.

When NSA and allied analysts really want to target an account, their concern for U.S. privacy diminishes. The rationales they use to judge foreignness sometimes stretch legal rules or well-known technical facts to the breaking point.

In their classified internal communications, colleagues and supervisors often remind the analysts that PRISM and Upstream collection have a "lower threshold for foreignness 'standard of proof' " than a traditional surveillance warrant from a FISA judge, requiring only a "reasonable belief" and not probable cause.

One analyst rests her claim that a target is foreign on the fact that his e-mails are written in a foreign language, a quality shared by tens of millions of Americans. Others are allowed to presume that anyone on the chat "buddy list" of a known foreign national is also foreign.

In many other cases, analysts seek and obtain approval to treat an account as "foreign" if someone connects to it from a computer address that seems

to be overseas. “The best foreignness explanations have the selector being accessed via a foreign IP address,” an NSA supervisor instructs an allied analyst in Australia.

Apart from the fact that tens of millions of Americans live and travel overseas, additional millions use simple tools called proxies to redirect their data traffic around the world, for business or pleasure. World Cup fans this month have been using a browser extension called Hola to watch live-streamed games that are unavailable from their own countries. The same trick is routinely used by Americans who want to watch BBC video. The NSA also relies routinely on locations embedded in Yahoo tracking cookies, which are widely regarded by online advertisers as unreliable.

In an ordinary FISA surveillance application, the judge grants a warrant and requires a fresh review of probable cause — and the content of collected surveillance — every 90 days. When renewal fails, NSA and allied analysts sometimes switch to the more lenient standards of PRISM and Upstream.

“These selectors were previously under FISA warrant but the warrants have expired,” one analyst writes, requesting that surveillance resume under the looser standards of Section 702. The request was granted.

‘I don’t like people knowing’

She was 29 and shattered by divorce, converting to Islam in search of comfort and love. He was three years younger, rugged and restless. His parents had fled Kabul and raised him in Australia, but he dreamed of returning to Afghanistan.

One day when she was sick in bed, he brought her tea. Their faith forbade what happened next, and later she recalled it with shame.

“what we did was evil and cursed and may allah swt MOST merciful forgive us for giving in to our nafs [desires]”

Still, a romance grew. They fought. They spoke of marriage. They fought

again.

All of this was in the files because, around the same time, he went looking for the Taliban.

He found an e-mail address on its English-language Web site and wrote repeatedly, professing loyalty to the one true faith, offering to "come help my brothers" and join the fight against the unbelievers.

On May 30, 2012, without a word to her, he boarded a plane to begin a journey to Kandahar. He left word that he would not see her again.

If that had been the end of it, there would not be more than 800 pages of anguished correspondence between them in the archives of the NSA and its counterpart, the Australian Signals Directorate.

He had made himself a target. She was the collateral damage, placed under a microscope as she tried to adjust to the loss.

Three weeks after he landed in Kandahar, she found him on Facebook.

"Im putting all my pride aside just to say that i will miss you dearly and your the only person that i really allowed myself to get close to after losing my ex husband, my dad and my brother.. Im glad it was so easy for you to move on and put what we had aside and for me well Im just soo happy i met you. You will always remain in my heart. I know you left for a purpose it hurts like hell sometimes not because Im needy but because i wish i could have been with you."

His replies were cool, then insulting, and gradually became demanding. He would marry her but there were conditions. She must submit to his will, move in with his parents and wait for him in Australia. She must hand him control of her Facebook account — he did not approve of the photos posted there.

She refused. He insisted:

“look in islam husband doesnt touch girl financial earnigs unless she agrees but as far as privacy goes there is no room….i need to have all ur details everything u do its what im supposed to know that will guide u whether its right or wrong got it”

Later, she came to understand the irony of her reply:

“I don’t like people knowing my private life.”

Months of negotiations followed, with each of them declaring an end to the romance a dozen times or more. He claimed he had found someone else and planned to marry that day, then admitted it was a lie. She responded:

“No more games. You come home. You won’t last with an afghan girl.”

She begged him to give up his dangerous path. Finally, in September, she broke off contact for good, informing him that she was engaged to another man.

“When you come back they will send you to jail,” she warned.

They almost did.

In interviews with The Post, conducted by telephone and Facebook, she said he flew home to Australia last summer, after failing to find members of the Taliban who would take him seriously. Australian National Police met him at the airport and questioned him in custody. They questioned her, too, politely, in her home. They showed her transcripts of their failed romance. When a Post reporter called, she already knew what the two governments had collected about her.

Eventually, she said, Australian authorities decided not to charge her failed suitor with a crime. Police spokeswoman Emilie Lovatt declined to comment on the case.

Looking back, the young woman said she understands why her intimate

correspondence was recorded and parsed by men and women she did not know.

“Do I feel violated?” she asked. “Yes. I’m not against the fact that my privacy was violated in this instance, because he was stupid. He wasn’t thinking straight. I don’t agree with what he was doing.”

What she does not understand, she said, is why after all this time, with the case long closed and her own job with the Australian government secure, the NSA does not discard what it no longer needs.

Jennifer Jenkins and Carol D. Leonnig contributed to this report.

Barton Gellman writes for the national staff. He has contributed to three Pulitzer Prizes for The Washington Post, most recently the 2014 Pulitzer Prize for Public Service. Follow @bartongellman