Justin Harvey Senate Testimony

Senate Judiciary Subcommittee on Privacy, Technology, Law

November 3, 2015

Good afternoon Chairman Flake, Ranking Member Franken and distinguished members of the subcommittee. Thank you for inviting me to speak with you today. My name is Justin Harvey and I am the Chief Security Officer for Fidelis Cybersecurity. Fidelis is a global company that provides products and consulting services to many of the largest companies in the United States and around the world as they work to prevent and respond to cyber-threats.

We operate on the front lines every day to protect some of the world's most sensitive data by helping organization detect, investigate and stop advanced cyber attacks. Over the years, Fidelis consulting teams have responded to some of the largest breaches on record. These include incidents where attackers were insiders – meaning employees – and cases where the breaches originated externally from organized crime syndicates, state-sponsored groups, so-called "hacktivist" or some other outside attacker.

I have been active in the information security field since 1994. That's when the Internet service provider I worked at was breached and we had to learn very quickly how to recover and restore service. Since those early days, we have all seen a tremendous transformation in the security field. There is an increasing amount of sophistication, both in terms of the threats that have

evolved and the strategies and tools that modern companies use to defend themselves. I am here today to talk about what are known as "data brokers" and the related cybersecurity challenges we are now facing.

In my testimony today, I would actually like to use the term "big data brokers" instead of the term "data brokers." The reason for this is that all companies collect and exchange more and more information from us as the information age progresses. However, "big data brokers" are collecting and analyzing our personal information at a pace and scale that is so staggering, it deserves special attention. Last month, for instance, The Wall Street Journal reported that one such broker had gathered an estimated 1,500 data points on as many as 700 million customers. Let me put this into perspective: Just a few years ago, a typical company may have collected an average of 40 data points on a consumer, and these data points were most often used for established purposes like determining if a transaction was fraudulent.

By contrast, today's "big data brokers" don't just store our personally identifiable sensitive information – things like email address, name, address, birthdate, social security number, maiden name, credit history, employer, and in some cases legal judgments – for routine purposes. These "big data brokers" go further into our behaviors and preferences, collecting information and insights that you might never have thought possible such as:

- Our buying habits gleaned from products and services we purchase both online and in brick and mortar stores. Every time you use your loyalty card at retail chains, they're

collecting this data, as well information about all of the things we purchased – or even considered purchasing – online

- Our health information, sometimes including real-time telemetry about our activity and sleep patterns from FitBits and other wearable fitness trackers
- Our leisure activities
- Socioeconomic status
- Our online viewing habits, including every movie we've watched and where we paused, forwarded, rewound or played scenes again
- In some cases, yes, who our favorite Saturday Night Live character is, don't worry Senator Franken, regardless of what the data shows you're still my favorite.

These are just a few of the advanced **TYPES** of data sets that are collected. Where big data brokers add extra value is in the "joining the data", or taking two or more distinct tables of data and piecing them together with a unique identifier. The power of joining datasets is really where true insights are derived. These are the kind of insights companies are paying for. For instance:

- By being able to join on "address", you get household members
- By joining on birth records, you get family relationships
- By joining on your web history, purchase history and certain social media keywords, you get what products you may be interested in buying
- And so on

There are many legitimate and responsible ways to leverage sensitive information in this fashion, but breaches and other misuse of data can cause great damage. We need to harness

the benefits of big data without opening the door to abuse. One major priority here is to strike the right balance between security and access. All organizations, not just "big data brokers", struggle between locking down their sensitive data and making it available for use and analysis.

It is the age old security problem that has been around forever: Too much security disrupts productivity, and too little is reckless. As I said, it's a balance; and striking that balance is not always an exact science. Let's consider a case in point around encryption. As long as the personal information is not a credit card number or something related to a person's health, there is usually not a business mandate to encrypt that data. In fact, some companies see encryption as a headache since it slows down accessing the data, and we have not yet seen widespread adoption of encryption to secure data "at-rest", meaning sitting in a database on a system, and "in-transit", meaning data traversing a network.

Companies clearly need to embrace encryption more as a best practice. However, encryption is NOT a silver bullet. It's only one of many components in a strong information security program. Ultimately, encryption is only as good as the weakest link in that larger data security ecosystem. And unfortunately, that weakest link is often the human factor.

From clicking on a malicious link or opening up an unknown email attachment, to being tricked into giving up a password, people still represent the largest vulnerability we have in cybersecurity. It is a persistent problem that also happens to be very frustrating – just last

week we heard CIA Director John Brennan express his own “outrage” at the hack of his personal email account through social engineering his mail provider to turn over his credentials.

While we’re always on guard system-wide for weaknesses and vulnerabilities, the most common vector of attack is still: the people. And I should mention that there’s another kind of people challenge as well: one that involves our industry’s workforce. The cybersecurity field happens to be suffering from a huge information security skills shortage. Companies and the government simply cannot hire, or retain, enough information security professionals. I would hope to enlist your support in finding ways to attract more graduates and professionals to the challenging and rewarding work of safeguarding data and protecting against cyber-threats. Programs like 1nterrupt in Boston need funding, they are working with high school teenagers to train the next generation of cybersecurity professionals.

The more we can improve our cybersecurity processes and workforce, the better we’ll be able to stop data from falling into the wrong hands. It is an ongoing struggle, and I’d like to share with you two axioms in that battle that we often use in our company and tell our customers:

The first is my own, and it reads: “Plan for the worst: If the information is stored online, it has a significant chance of getting leaked at some point in time.” No dataset is completely secure, no matter how much security or encryption is placed upon it. We must face and prepare for the high chance that it will be stolen and possibly leaked. Sometimes the fallout is massive – as with the 21.5 million cyber-attack victims in the Office of Personnel Management’s database.

Sometimes there is potential embarrassment – as with the hacking of customer records at Ashley Madison, a commercial website designed to enable extramarital affairs. These are just two datasets that have caused a measure of public or personal havoc that we might not have thought possible before.

The second axiom is one I attribute to a colleague named Will Irace. He says: "If determined attackers want to get in, they're going to get in." He's right: No matter how much money is spent on people, process and technology: every single organization – whether it be a military, company, government or non-profit entity – is vulnerable in some fashion.

Both these axioms point to one unfortunate truth: **Breaches are inevitable.** And when it happens to big data brokers, the scope, scale and sophistication of sensitive data and analytics they possess make the stakes even higher than they otherwise would be.

In my opinion, the most obvious risk is widespread fraud. Should these datasets get leaked to the general public, or sold to the highest bidder, this could have cascading effects on our economy. Imagine a breach where every American's name, social security number, address, email, phone number and mother's maiden name was leaked to the Internet. The size of this data-set is not outside the realm of possibility.

To put this into perspective, the Ashley Madison breach, which involved far more data-fields than I just described, had information on 36 million people contained in only 10 gigabytes of

stolen data. Given that the last US census reported a little over 320 million Americans, a whole country's worth of personal data could therefore be compressed into about 100 gigabytes. That is a little under twice the size of the computer game Battlefield 4, which our kids play today, or ten seasons of Saturday Night Live in high definition. That amount of data could fit on a USB thumb drive or mobile phone.

A breach of this size would have a lasting effect on our nation's economy and national security as the government and corporations rushed to implement stop-gap measures to respond to the leak. Consumers would also be harmed, as they would have to essentially change every single password and reestablish their own secure place on the Internet. Identity recovery, in other words, could be as troubling as identity theft. I'm reminded of the famous caption from a 1993 New Yorker cartoon that has since become a maxim of the digital age: "On the Internet, nobody knows you're a dog". Well, in the aftermath of a massive breach, you may find that on the Internet, nobody knows you're you.

Finally, the information that "big data brokers" have collected and generated represents some of the richest "metadata" about our citizens in the world. Metadata is that all-important "data about data" that helps us understand the context and usability of the information we possess. Unlike Internet surveillance, which can be foiled by encryption and legislation, brokers have gotten their data first-hand from companies collecting that information from users. This provides unrivaled clarity and visibility into the meaning and usefulness of personal data.

Why is this important? I believe the volume, detail and richness of information that "big data brokers" possess makes them a prime target for breaches – especially state sponsored cyberespionage. Nation-states likely see brokers as a "one-stop-shop" for intelligence on US citizens. No need to breach ten, fifty or a hundred other corporations in the US to get data when they can compromise just one or two brokers who have already done the work of integrating data from many other sources and optimizing its usefulness. Since I forgot my tin foil hat at home, I won't begin to discuss the possibility of US intelligence agencies using this metadata to find threats.

Nation-states that have access to this rich metadata can easily track the habits of US citizens for nefarious purposes. This could include shadowing a target of interest – say a government employee – to discover online and real world habits in order to gain access to Internet accounts that may hold classified or sensitive information. Or perhaps an espionage recruiter might sift through telltale behavioral patterns and sentiment analyses to identify possible candidates who share political leanings or – as the Ashley Madison breach reminds us – social vulnerabilities are ripe for exploitation or coercion.

I realize that my testimony today is probably drawing an uncomfortable picture of the challenges we face, and we do indeed need to accept the "when, not if" reality that breaches will happen. Fortunately, the same huge information ecosystem that makes cyberattacks on big data brokers so tempting also provides us many tools to track and fight those incursions.

Attackers can be relentless and clever, but so can we. We're increasingly able to "listen" to data in real time and deploy sophisticated analytics to recognize immediately when a breach is taking place, follow the attackers' trail and freeze them in their tracks. As an industry, we can even mitigate future attacks by examining breach patterns that reap insights and visualizations about where and how new incursions may unfold.

My point is: We are very much in the game! It takes commitment, ingenuity and – as I mentioned earlier – a robust cybersecurity workforce, but we are getting better and better at fighting cyberattacks, and so are the "big data brokers". While we may not be able to stop all breaches, our ability to leverage big data for insights about these attacks means we're rarely in the dark and we are making progress all the time.

Thank you Senators and I look forward to your questions on the matter.