The University of Texas at Austin
Center for Identity

The Danger of Putting Your Digital Life in One Place

Jennifer Brenner

2014

UT CID Report #1403

This UT CID research was supported in part by the following organizations:

LifeLock Guarantee Your Good Name
HID
TransUnion
Applied Fundamentals Consulting LLC
CSID
LexisNexis
SAFRAN MorphoTrust USA

Over ninety percent of all data in the world has been collected in the past few years alone.[1] As technology becomes increasingly integral to our lives, we create an incredible and unprecedented amount of personal data. There is more data than ever before and, consequently, more digital information about out lives.

Each of us creates and provides unique personal data on a daily basis. Every online behavior and action we take becomes an immediate data point. Each click, search, purchase, and "Like" comes together to create our digital identity. For instance, even the download of this white paper is now a data point in your profile.

Companies compile this huge collection of data to create user profiles. User profiles include information about our online behaviors, shopping preferences, privacy settings, location data, search history and more. Companies use these profiles to provide consumers with customized, efficient, and convenient shopping experiences. To that purpose, the voluminous collection of data offers clear benefits for users in terms of speed and convenience of online services. But with an increased collection of data, and a reliance on technology, comes unprecedented amounts of and access to personally identifiable information (PII). With the increased availability of personal data comes an increased vulnerability to the negative consequences of online data collection, including identity theft, fraud, and abuse.

What happens when you put your digital life all in one place?

Allured by promises of speed and convenience, consumers have started to use tools that keep their digital lives all in one place. Applications such as Google Wallet, Coin, and PayPal Beacon compile users' account information (name, address, phone number), credit card information, user behaviors, and shopping preferences in a (seemingly) convenient bundle. When people want to pay for an item or service, they need only to access a single account that houses all of their relevant PII. But what concerns should we have when housing our digital lives all in one place?

There has been a notable increase in the number of industry applications that compile consumer PII with the promise of a seamless and quick payment transaction. These applications assert that by having credit cards, user preferences, and shopping habits tightly knit in one place, shopping experiences can meet consumers' needs of speed, convenience, and customization.

[1] Science Daily, 2013: http://www.sciencedaily.com/releases/2013/05/130522085217.htm

Factors that Attract Consumers to Data Aggregation Tools

1. Time
2. Convenience
3. Customization

The convenience of data aggregators is admittedly alluring. The notion that a customer's name could be added automatically to a waitlist for a table at a restaurant, or that a pharmacy could populate prescriptions automatically upon the person walking through the door appears convenient and useful. As the consumer shops around the store, the prescriptions are filled, and by the time they make it to the pharmacy, everything is ready to go. In a world where people are busier than ever, any time saved can help improve a consumer's quality of life.

Convenience and speed are the primary motivators for consumer adoption of data aggregators such as Google Wallet, Coin, and PayPal Beacon.

However, convenience often faces a direct tradeoff with security. Frequently, the "easier" it is to conduct a payment transaction, the more likely it is to be unsafe. This is not always the case, but with convenience of transaction often comes a circumvention of necessary security precautions. For instance, third parties such as hackers may be able to access your information simply be being on the same wireless network during a seemingly secure transaction.[2] Identity thieves may even go so far as to set up their own wireless network at coffee shops under the name of the business (e.g., "S-Bucks Wi-Fi") in attempt to lure consumers to use their insecure Wi-Fi networks. The thief then records customers' transactions and behaviors and gains access to their financial information.

Why the Aggregation of Data Can Be Dangerous

The housing of multiple forms of PII in a single place changes the identity landscape. Thus it is important to examine the repercussions of aggregating personal information that people use to both identify and protect themselves.

How do data aggregators change the identity landscape?

In the past, a fraudster was required to piece together the digital life of a victim via illegal skip

[2] CSID, "When Good Technology Goes Bad: Evolution of Mobile Technology"

tracing sites and paid searches. They had to scour online databases and hack low-security accounts to "get to know" the intended victim. Now, in conjunction with social media sites, data aggregation tools give fraudsters access to pools of data about potential victims. Essentially, consumer identities have been wrapped in a package with a bow and left on a labeled shelf for any hacker to pick up and use as they desire. These services pair not only debit and credit card numbers, but also your hobbies, habits, online behaviors, and personal interests.

But this compilation of personal data into a convenient bundle is harmless, right? Unfortunately, the answer is no. Personal data often acts as a second layer of defense on a consumer's most important personal accounts, in the form of security questions. Security questions are used as a form of authentication by banks, cable companies and wireless providers as an extra protective layer.[3] Typically, during the account creation process, security questions are selected and answered in addition to providing a password. This enables a consumer to authenticate him or herself in the event they forget the password for the account.

The following prompts are commonly utilized security questions:

- In what city was your first elementary school?
- What is your mother's maiden name?
- What is your pet's name?
- What is your favorite food?
- What is your favorite book?
- What is the name of the road you grew up on?
- What high school did you attend?

It is shocking to find how easily personal information can be obtained. For example, the information for the common security question, "In what city was your first elementary school?" can often be found on Facebook. Facebook's "About Me" section features a "hometown" descriptor where users name the city where they were born. And, many people attended elementary school in their hometown, making it easy to surmise the answer to the security question. Moreover, most of a person's former addresses can be found in public databases. Some Facebook users even list the schools they have attended on their profiles.

Questions such as "What is your mother's maiden name?" also often can be deduced via Facebook lineage or Ancestry.com, which permits users to establish a family tree. Some family

[3] Levin, Josh (2008-01-30). "In What City Did You Honeymoon? And other monstrously stupid bank security questions". Slate.

members even have their maiden names listed on their profiles, making the security answer even easier to find. In reality, most if not all of the security questions above could be discovered using a combination of public data, social media, and aggregated user profiles.

There are some security questions, however, that offer a greater level of protection. The strongest security questions are ones that are opinion-based or subjective in nature. For example, the questions "What is your favorite time of day?," "Who was your favorite teacher in elementary school?," or "What country would you most like to visit?" offer greater security because they require information not readily available online. Answers that can change over time, or that are opinion rather than fact, offer stronger security protection than concrete questions whose answers can be verified.

It is important for consumers to be mindful of the security questions they select. This simple precaution can prevent hackers from gaining access to their accounts.

The true danger, however, lies in what information can be discovered from analyzing the "subjective" data that surrounds consumer activities and the habitual decisions made on a regular basis. When housing multiple forms of PII in a single location, such as Google Wallet, questions inquiring about your opinions, such as your favorite restaurant may be easily discovered using Google Wallet information, where all of your loyalty cards are stored in one place.

Most people may not even realize how often they perform certain activities or how these activities are tied to their identity. The danger is not in a thief finding one single piece of your personal data, but in how individual pieces of personal information come together to create full profiles of potential victims. This type of synthesized profile is typically how fraudsters carry out large-scale theft that results in large quantities of money stolen.

Data aggregation also permits cross-site manipulations to occur. Cross-site manipulation is where a fraudster compromises multiple accounts using known, linkable vulnerabilities across the targeted platforms. This type of fraud is exemplified in Mat Honan's depiction, "How Apple and Amazon Security Flaws Led to My Epic Hacking."[4] In the article, writer and victim Mat Honan explains in grave detail, how easily identity was compromised through cross-site manipulation. He begins the article by explaining that, "in many ways, this was all my fault. My accounts were daisy-chained together." The author alludes here to the dangers of data

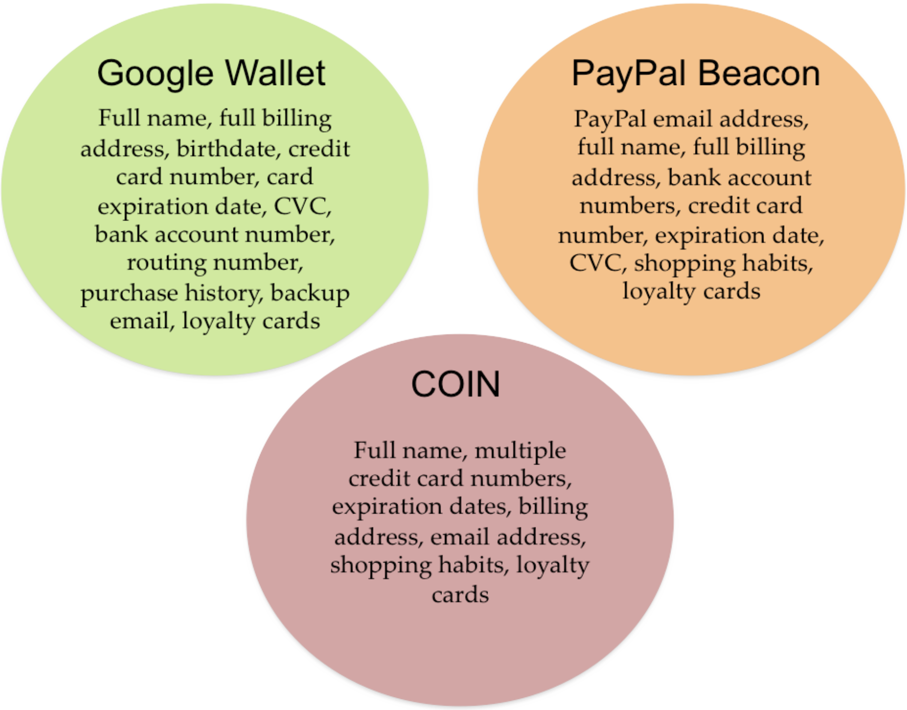
[4] Honan, Mat. "How Apple and Amazon Security Flaws Led to My Epic Hacking" | Gadget Lab | Wired.com. Gadget Lab http://www.wired.com/gadgetlab/2012/08/apple-amazon-mat-honan-hacking/

aggregation. In Matt Honan's case, the fraudster did not have access to a complete profile of the victim, but the victim had much of their personal information unnecessarily connected. Once the fraudster infiltrated one area, he was able to hop from one piece of information to the next by manipulating vulnerabilities in both the Amazon and Apple platforms.

Although there are several measures the victim could have taken to protect himself, the illustrative point is that when the attacker needed to gain access to his Apple Account, all that was needed were the last four digits of a credit card listed on the account. The fraudster knew that the same last four digits of a credit card are readily displayed in an Amazon account. The hacker gained access to the victim's Amazon account first, viewed the last four digits of his credit card on file, and then used that information to authenticate himself and reset the password on the Apple account. If an attacker has access to pools of consumer data, it becomes significantly easier for inherent platform vulnerabilities to be manipulated, enabling this type of access.

Consumers do not have control over these inherent software vulnerabilities nor are they always aware they exist. It is important to minimize the "daisy chaining" of accounts whenever possible to make it more difficult and less rewarding for a hacker to steal personal data. Figure 1 depicts all of the consumer data that is collected by the data aggregators Google Wallet, Coin, and Paypal Beacon.

Sampling of Data Collected by Aggregator



Google Wallet

Full name, full billing address, birthdate, credit card number, card expiration date, CVC, bank account number, routing number, purchase history, backup email, loyalty cards

PayPal Beacon

PayPal email address, full name, full billing address, bank account numbers, credit card number, expiration date, CVC, shopping habits, loyalty cards

COIN

Full name, multiple credit card numbers, expiration dates, billing address, email address, shopping habits, loyalty cards

Moving Forward

A consumer's level of trust in a provider ultimately determines whether they will use their data aggregation service. But in the aftermath of large-scale security breaches of Target, Neiman Marcus, eBay, and others, users are justified in being skeptical of trusting companies' security processes. Users are susceptible to identity theft anytime they access sensitive data over a network, making these issues important to ponder and processes important to improve.

Consumers and businesses alike must make concerted efforts to protect the privacy and security of their own and their customers' PII. Debating the ease of convenience against the rigors of security is an ongoing and important debate in the field of identity and data aggregation. Increased attention to security processes will yield benefits for all involved. In the meantime, consumers can help protect themselves by avoiding the seeming convenience of piling all of their PII into a single basket.

The University of Texas at Austin
Center for Identity

For more information on Center for Identity research, resources and information, visit **identity.utexas.edu.**