Abstract—Identity theft is perhaps the defining crime of the information age. Identity theft threatens various demographics, but some age groups, e.g., senior citizens, are particularly vulnerable. In this paper, we study how identity theft puts different personally identifiable information (PII) assets at risk of exposure, and how this risk changes throughout one’s lifecycle. We categorize PII assets, introducing a fourth novel category, measure their exposure risk using the Identity Theft Assessment and Prediction (ITAP) repository of over 3,000 identity theft cases, and track the risk change throughout an individual’s lifecycle. We introduce the concept of PII Balance Sheets™, and finally, we present a free publicly available Android app that demonstrates our research results. This app not only educates individuals and highlights their vulnerable identity assets, but also is useful when they make the decision of whether or not to share their personally identifiable information.

I. INTRODUCTION

Identity theft is reportedly one of the fastest growing crimes in the U.S. According to the Bureau of Justice Statistics of the U.S. Department of Justice, only in 2014, about 7% of persons age 16 or older in the U.S. were victims of identity theft. These victims suffered from financial loss as well as emotional distress: 64% of identity theft victims reported a direct financial loss with a median of $300 and a mean of $1,349. Furthermore, one in ten identity theft victims reported severe distress due to the crime [1]. When it comes to identity theft, some age groups of victims are particularly vulnerable. For example, the number of elderly victims of identity theft in the U.S. increased from 2.1 million in 2012 to 2.6 million in 2014, showing a 24% increase, while the number of identity theft victims in all other age groups measured did not significantly change [1]. Furthermore, people use various personally identifiable information (PII) throughout their lifecycle to identify themselves, putting different PII at risk at each age. For instance, while a Social Security Number is used to identify a person at any age, one’s spouse name is not applicable as PII for a child and a phone number is less frequently used as PII for a child.

While it is important to study how the risk of exposure for different PII assets changes throughout a typical lifecycle. In this paper, we make the following contributions:

1) We categorize personally identifiable information (Section II).

2) We propose a fourth PII category—what you do—in addition to the three commonly used categories—what you are, what you have, and what you know (Section III).

3) We identify how the risk of exposure for different PII assets changes throughout an individual’s lifetime, using the Identity Theft Assessment and Prediction (ITAP) repository of over 3,000 identity theft cases (Section III).

4) We present the concept of “PII Balance Sheets™” (Section III).

5) We introduce a tool—Risk Kit Android app—that presents this information in a free, interactive, and easy to use manner (Section IV).

II. PII CATEGORIES

In order to study and present the risks of exposure for PII assets and their trend of change throughout a person’s lifetime, a clear concise categorization of PII assets is helpful. Traditionally, user authentication has been based on three factors: what you are (e.g., a fingerprint), what you have (e.g., a drivers license), and what you know (e.g., a password). Similarly, PII assets have been categorized in the same three categories. Researchers have suggested multiple options to be used as the fourth authentication factor and consequently as a fourth category of PII. For instance, Brainard et al suggested somebody you know [2], and Choi and Zage used where you are [3]. In light of recent developments of pervasive and mobile computing, we introduce a new authentication factor and PII category: what you do. Examples of what you do include GPS location, mobile phone usage, and online purchase patterns.

Here we define the four authentication factors and use them to categorize PII assets.

1) What you are: Something you are born with, difficult or impossible to take away. For example, fingerprints, iris scan, gender, and eye color.

2) What you have: Something assigned to you by other organizations, such as credentials and numbers. For example, birth certificate, employee ID, mobile phone number, and Social Security Number.

3) What you know: Information known privately to you, often created by you. For example, date of birth, first pet, mother’s maiden name, and password.
4) What you do: Behaviors influenced by your location and actions. For example, air travel patterns, GPS location, Internet browsing patterns, mobile phone usage patterns, and online purchase patterns.

III. PII Risk of Exposure for Specific Age Groups

We identify a list of PII assets, categorize them under the four categories, and predict the trajectory of exposure risk throughout a typical lifecycle of the PII owner. To obtain a list of important PII in each category and also the risk of exposure and its changes, we use the Identity Threat Assessment and Prediction (ITAP) project at the Center for Identity at the University of Texas at Austin.

A. ITAP

ITAP is a risk assessment tool that increases fundamental understanding of a criminal’s identity theft and fraud processes and patterns. ITAP aggregates data on identity theft from multiple sources (e.g., law enforcements, fraud cases, and news stories) to model and analyze identity vulnerabilities, the value of identity attributes, and their risk of exposure. At the Center for Identity, a team of modelers carefully analyzes identity theft and fraud news stories on a daily basis and models this information using the ITAP schema. For each case of identity theft and/or fraud, ITAP collects and analyzes tools used by criminals, type of information exploited, demographics of victims, etc. In doing so, ITAP captures and analyzes a structured computational model of identity theft and fraud processes and outcomes [4], [5].

We utilized 3,074 ITAP news stories that report on specific identity theft and fraud cases. The information gathered includes the PII assets stolen or misused, time line, location, demographics of victims and performers, resources used, monetary loss incurred, emotional distress to victims, and counter measures taken, among other information. Based on the 3,074 identity theft and fraud cases, ITAP identified 197 PII assets that were stolen, fraudulently used or misused.

B. Risk of Exposure Calculation

We estimate the risk (i.e., probability) of exposure for any PII asset $i$ as follows:

$$risk(i) = \frac{\#\text{cases}(i)}{\#\text{cases}} \tag{1}$$

where $\#\text{cases}(i)$ is the number of theft cases in which this PII was exposed to or misused by the criminal and $\#\text{cases}$ is the total number of theft cases in ITAP.

Furthermore, we calculate the risk of exposure for a PII asset, given the victim belongs to a specific age group. Utilizing the victim demographics data of ITAP, we use conditional probability to focus on any given age range of victims. For any age range, it suffices to filter the cases to those related to that age range of victims and use the same formula as above. Thus, for any PII asset $i$, and age range $a$ (e.g., children, adults, senior citizens)

$$risk(i|a) = \frac{\#\text{cases}(i,a)}{\#\text{cases}(a)} \tag{2}$$

where $\#\text{cases}(i,a)$ is the number of cases that involve the theft or misuse of PII $i$ from a victim in the age range $a$ and $\#\text{cases}(a)$ is the total number of cases with victims in the age range $a$.

C. Value Calculation

We define the dollar value of any PII $i$ as the average dollar value lost in cases that $i$ was misused by identity criminals. Hence:

$$value(i) = \frac{\sum_{c \in \text{cases}(i)} v_c}{\#\text{cases}(i)} \tag{3}$$

were $v_c$ is the dollar value loss incurred in a given case $c$.

We calculate the value of PII $i$ for a given age range $a$ by filtering to the cases that involve victims in that age range. Therefore:

$$value(i, a) = \frac{\sum_{c \in \text{cases}(i,a)} v_c}{\#\text{cases}(i,a)} \tag{4}$$

D. PII Balance Sheets

We borrow the concept of a balance sheet—a statement of the assets, liabilities, and capital of a business or organization—to elaborate on the risk of exposure and value of PII assets. Considering PII as identity assets, the value of PII is its asset value, and the risk of exposure can be viewed as a liability—the probability that criminals steal, fabricate or misuse this PII type.

Figure [1] shows the commonly top ranked value and top ranked risks across the life cycle. Under each of the six age ranges, the top 12 valuable PII assets (out of 197) are identified. In addition, for each age rage, the top 12 PII assets (out of 197) that have the highest risk of exposure are listed too. Interestingly, Name, Social Security Number (SSN), and Date of Birth consistently have the highest risk of misuse throughout the life cycle. These PII assets are in fact the only three ones that a fraudsters needs to perform many identity crimes, e.g., filing false tax return in one’s name.
PII Balance Sheets

PII is an asset. The liability and value of most PII assets changes over your lifetime and beyond. In fact, these values are changing for all individuals all of the time, based on ever-changing identity threats.

Leveraging data and analytics from the Identity Threat Assessment and Prediction (ITAP) and the Identity Ecosystem Maps projects at the Center for Identity, the "PII Balance Sheets™" below illustrate the liability and value of identity assets over time.

<table>
<thead>
<tr>
<th>COMMON TOP RANKED LIABILITY (RISKS) ACROSS THE LIFECYCLE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ASSET VALUE</strong> Gain to criminal who steals, fabricates, or abuses this PII type</td>
</tr>
<tr>
<td><strong>LIABILITY</strong> Probability criminal will steal, fabricate, or abuse this PII type</td>
</tr>
</tbody>
</table>

---

**0-5 YEARS**

1. Credit Card Information
2. Medical Record Number
3. Last Appointment Date
4. Treatment Information
5. Debit Card Information
6. Driver’s License
7. Email Address
8. PII of Children
9. Phone Number
10. Date of Birth
11. Doctor’s Name
12. Birth Certificate(s)

**1. Credit Card Information**
2. Medical Record Number
3. Last Appointment Date
4. Treatment Information
5. Debit Card Information
6. Driver’s License
7. Email Address
8. PII of Children
9. Phone Number
10. Date of Birth
11. Doctor’s Name
12. Birth Certificate(s)

---

**10-18 YEARS**

1. AMC (BN) Number(s)
2. Magnetic Stripe(s)
3. Direct Patient Number
4. Debit Card Information
5. Credit Card Information
6. Medical Information
7. Social Security Number (SSN)
8. Date of Birth
9. Driver License Number
10. Birth Certificate(s)
11. PII of Children
12. Email Address

**1. AMC (BN) Number(s)**
2. Magnetic Stripe(s)
3. Direct Patient Number
4. Debit Card Information
5. Credit Card Information
6. Medical Information
7. Social Security Number (SSN)
8. Date of Birth
9. Driver License Number
10. Birth Certificate(s)
11. PII of Children
12. Email Address

---

**18-30 YEARS**

1. AMC (BN) Number(s)
2. Magnetic Stripe(s)
3. Direct Patient Number
4. Debit Card Information
5. Credit Card Information
6. Social Security Number (SSN)
7. Date of Birth
8. Driver License Number
9. Birth Certificate(s)
10. PII of Children
11. Email Address

**1. AMC (BN) Number(s)**
2. Magnetic Stripe(s)
3. Direct Patient Number
4. Debit Card Information
5. Credit Card Information
6. Social Security Number (SSN)
7. Date of Birth
8. Driver License Number
9. Birth Certificate(s)
10. PII of Children
11. Email Address

---

**30-50 YEARS**

1. AMC (BN) Number(s)
2. Magnetic Stripe(s)
3. Direct Patient Number
4. Debit Card Information
5. Credit Card Information
6. Social Security Number (SSN)
7. Date of Birth
8. Driver License Number
9. Birth Certificate(s)
10. PII of Children
11. Email Address

**1. AMC (BN) Number(s)**
2. Magnetic Stripe(s)
3. Direct Patient Number
4. Debit Card Information
5. Credit Card Information
6. Social Security Number (SSN)
7. Date of Birth
8. Driver License Number
9. Birth Certificate(s)
10. PII of Children
11. Email Address

---

**50+ YEARS**

1. AMC (BN) Number(s)
2. Magnetic Stripe(s)
3. Direct Patient Number
4. Debit Card Information
5. Credit Card Information
6. Social Security Number (SSN)
7. Date of Birth
8. Driver License Number
9. Birth Certificate(s)
10. PII of Children
11. Email Address

**1. AMC (BN) Number(s)**
2. Magnetic Stripe(s)
3. Direct Patient Number
4. Debit Card Information
5. Credit Card Information
6. Social Security Number (SSN)
7. Date of Birth
8. Driver License Number
9. Birth Certificate(s)
10. PII of Children
11. Email Address

---

**DECEASED**

1. AMC (BN) Number(s)
2. Magnetic Stripe(s)
3. Direct Patient Number
4. Debit Card Information
5. Credit Card Information
6. Social Security Number (SSN)
7. Date of Birth
8. Driver License Number
9. Birth Certificate(s)
10. PII of Children
11. Email Address

**1. AMC (BN) Number(s)**
2. Magnetic Stripe(s)
3. Direct Patient Number
4. Debit Card Information
5. Credit Card Information
6. Social Security Number (SSN)
7. Date of Birth
8. Driver License Number
9. Birth Certificate(s)
10. PII of Children
11. Email Address

---

**Fig. 1.** Top Ranked Stolen, Fraudulently Used or Abused PII Assets Based on Value and Risk of Exposure (Liability).
IV. RISK KIT APP

We incorporate the results of this research in an Android app called IDWise Risk Kit. Risk Kit is a valuable educational resource, but most importantly it is an available source of information at the palm of users’ hands when they are making the decision of whether or not to share their PII in a given situation. If for their specific age range, that piece of PII is at high risk of exposure, it means that identity thieves have frequently stolen, fraudulently used or misused such PII from victims that share the same age range. Therefore, it is not advisable to share the PII unless proper safeguards are in place. The IDWise Risk Kit was prepared as part of the IDWise outreach and education project at the Center for Identity at the University of Texas at Austin. It is freely available at the Center’s web page at https://identity.utexas.edu/risk-kit-app.

A. Risk Kit App Content

Since it is not practical to list all the 197 PII assets in the app, Risk Kit aggregates all the PII assets that ranked in the top 12 most valuable or top 12 most risky for any age range as shown in Figure 1. It lists the PII assets under the four categories of Section III. The app shows the risk of exposure of PII assets as calculated on the ITAP data using the formulas of Section III-B. It further shows how the risk changes over each of the following six age ranges:

- Age 0-5 years
- Age 5-18 years
- Age 18-30 years
- Age 30-50 years
- Age 50 and older
- After death

To facilitate the representation, the risk of exposure for each PII asset is indicated by a color:

- Green: low risk of exposure
- Yellow: medium risk of exposure
- Red: high risk of exposure
- Gray: not applicable

If the risk is less than a preset LOW threshold it is deemed low, more than LOW threshold but less than MEDIUM threshold is medium, and more than MEDIUM threshold is high. Currently LOW = 0.02 and MEDIUM = 0.05. The average of risks for the current data shown in the app is 0.02, therefore, by setting MEDIUM = 0.05, we assign a high risk to any risk over the average.

The same information displayed by the app is shown in Figure 2. In this figure PII assets are broken into four groups which represent what you are, what you have, what you know, and what you do categories in order. Note that PII assets are not equally distributed between the four categories, because, in practice, some PII assets are more frequently stolen, fraudulently used or misused. The fact that only two PII assets of the what you do category and only four of the what you are category ranked in the commonly stolen or misused PII assets shows that either these types of PII are not frequently used, or they are harder to steal, fraudulently use, or misuse.

B. Download and Installation

In order to install the Risk Kit app on an Android phone or tablet:

1) Allow installation of apps from sources other than the Google Play Store by going to Settings/Security and checking Unknown Sources.
2) Navigate to the web page for the IDWise Risk Kit app, https://identity.utexas.edu/risk-kit-app and click the link to download.
3) Once the download is complete, click on the downloaded apk file (for example from the notification bar) and click Install. The app will now be installed.
4) Remember to un-check Unknown Sources which was checked in step 1 after the installation is complete.
Fig. 2. Identity Risk throughout the Lifecycle. The Groups of PII Assets are What You Are, What You Have, What You Know, and What You Do in Order.
C. How to Use

In the very first launch of the app, Risk Kit shows hints on how to use it (Figure 3). The same information can be found under Menu/About the App (Figure 4).

Within the app, the user may select an age range using the slider as shown in Figure 5. The app displays PII assets and their risk of exposure for that specific age group in the above-mentioned four categories (Figure 6). The app and its help menus are self-explanatory. You can go to the Menu by touching the IDWise Risk Kit icon on the upper left. You can get out of any screen (including menu pages, list of assets, etc.) by pushing the hardware Back button.

V. Related Work

There exists a body of research that studies demographics of identity theft victims. The research is mainly conducted by Federal and State agencies, private organizations, or academic institutions.

From Federal and State agencies, studies by U.S. department of Justice [1] regularly report the distribution of identity theft victims including their age. Their latest report found a 24% in the number of elderly victims of identity theft.

Private organizations such as Javelin [6] have published comprehensive studies and case studies on the subject of identity theft and fraud.
In the academia, Allison et al. [7] found the typical victim of identity theft and fraud to be male and white. They reported an average age of 41 for the victims, however did not report the actual distribution. Copes et al. [8] used the National Public Survey on White Collar Crime to separately investigate existing credit card fraud, new credit card fraud, and existing bank account fraud. They reported statistics on age groups of victims and compared demographic patterns of identity theft victims to their representation in the general U.S. population. Reyns [9] performed an in depth empirical examination of identity theft in the United Kingdom from a victimization perspective. Using Routine Activity Theory, he found male older individuals, and those with higher incomes more likely to experience identity theft. Others have performed similar studies based on Routine Activity Theory in various jurisdictions (e.g., Pratt et al. in the U.S. [10]).

These studies, however, typically report statistics on the demographics of identity theft and fraud victims and have never focused on investigating the identity theft for a particular victim age range in the way we did. In particular, these studies report the frequency of victimization for a particular group but not the specific PII as liabilities/risk levels and asset value for respective age groups.

VI. Conclusion and Future Work

In this paper, we studied the risk of exposure for various PII assets and their trajectory of risk change over the PII owner’s lifecycle. We introduced a novel category of PII, namely what you do in addition to the widely used categories what you are, what you have, and what you know. We calculated the risk (i.e., probability) of exposure and the value of PII using data from the UT CID Identity Threat Assessment and Prediction Project model of over 3,000 identity theft and fraud cases. We introduced the concept of “PII Balance Sheets™”. Finally, we developed an Android mobile application, namely IDWise Risk Kit, that incorporates the findings of this research. We hope that the Risk Kit app can become a widely used tool to guide individuals whether or not to share a particular PII at a given situation.

We envision multiple paths as future work:
1) Performing regular updates of the PII risks of exposure as our database of identity theft cases grows.
2) Implementing Risk Kit as an iOS app for iPhones and other Apple devices.
3) Including the dollar value of each of the PII assets in the app to estimate the monetary loss if the PII is stolen, fraudulently used or abused for that age range.
4) Extending the dimensions of specializing the presented information beyond age range by including gender, citizenship, etc.

ACKNOWLEDGMENT

The authors would like to thank Jeremy Gordon for designing the user interface of Risk Kit.

REFERENCES
