pipl

IDENTITY DATA TIPS

Defense Against Synthetic Identity Fraud

Synthetic Identity Fraud

Synthetic identity fraud is the fastest growing type of financial crime, according to the Federal Reserve[1]. Account application processes that rely on static personally identifiable information (PII), such as Social Security numbers and credit file information, are the most vulnerable. Relying on static PII is risky, because too much of it has already been compromised.

A basic example of synthetic identity fraud is when perpetrators combine partial information from a real person with fake details, such as a combination of name, address, contact and other information. It's called "synthetic identity" because the result is a fraud-ready synthesis of real and falsified information.

Fraudsters seed and cultivate these fake identity records to apply for credit and other services, which they ultimately "bust out" at an opportune time. Additionally, there are credible estimates that 85 to 95 percent of applicants who were (later) identified as synthetic identities were not initially flagged as high risk by traditional digital-channel fraud models.

Some defensive measures, such as biometrics (e.g. fingerprints) are increasingly being used for fraud prevention and seamless commerce, but the problem when it comes to synthetic ID fraud is the lack of a 'real user baseline.' If you have no baseline of biometrics being associated with a synthetic ID account, perpetrators can fake those biometrics.

85%
of synthetic identities are not initially flagged as high risk by traditional digital channel fraud models.[2]

How an Advanced Identity Index Can Help

While one of the peskier strengths of synthetic IDs is that they are partially composed of very real attributes – there is a bright side. The right identity verification approach can flip this 'strength' into an effective array of countermeasures.

Machine Learning

One approach is to parse individual identity records and use machine learning to identify correlations that are more subtle within traffic data patterns. Phone numbers for instance are a common data point used by consumers and organizations during application, onboarding and other activities. This may require an upstream verification tactic that tracks phone numbers back to telecom providers to confirm it's being used and has historically been used, by the actual owner of that data.

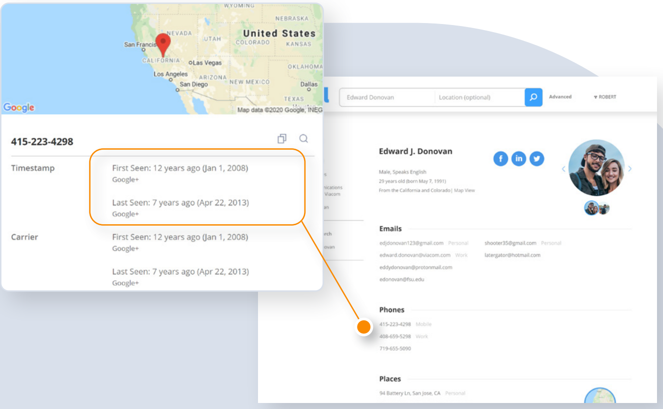
Identity Resolution

Knowing that some aspects of a synthetic ID may be legitimate, it is more important than ever to accurately calculate the relationship all attributes have to each other (or not) on a statistical basis. We think of this as "high-confidence identity resolution" and it factors into what we believe about overall coverage. In short, an effective identity index must have the ability to resolve or "match" the highest number of attributes with the greatest degree of confidence, across the broadest possible geographic dimension.

[1] https://www.federalreserve.gov/newsevents/pressreleases/other20190709a.htm
[2] https://www.idanalytics.com/wp-content/uploads/2018/11/Synthetic-Identity_Slipping-through-the-cracks_Executive-Summary.pdf

Signals of Synthetic Identities

With vast and accurate identity indexes, it is possible to drill down at scale on individual identity data attributes and corresponding metadata to probe for tell-tale signals of synthetic identity records. For this, there are several pointers to share for enabling fraud models or analysts to determine whether the identity, as a whole, is a real person, no matter their location.



415-223-4298

Timestamp
First Seen: 12 years ago (Jan 1, 2008)
Google+
Last Seen: 7 years ago (Apr 22, 2013)
Google+

Carrier
First Seen: 12 years ago (Jan 1, 2008)
Google+
Last Seen: 7 years ago (Apr 22, 2013)
Google+

Edward Donovan | Location (optional) | Advanced | ROBERT

Edward J. Donovan
Male, Speaks English
29 years old (born May 7, 1991)
From the California and Colorado | Map View

Emails
edjdonovan123@gmail.com Personal
shooter35@gmail.com Personal
edward.donovan@viacom.com Work
latergator@hotmail.com
eddydonovan@protonmail.com
edonovan@fsu.edu

Phones
415-223-4298 Mobile
408-659-5298 Work
719-655-5090

Places
94 Battery Ln, San Jose, CA Personal

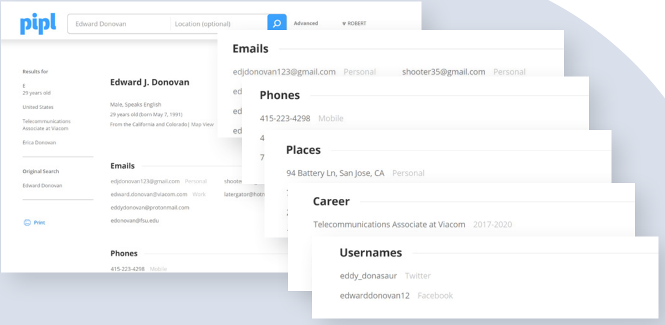
Timestamps

The right identity data index for this purpose should include a robust and consistent hash function for delineating *first seen/last seen* dates for each attribute contained within identity records.

🚩 Strong symmetrical pattern of *first/last seen* timestamps across a majority of attributes is indicative of a synthetic ID.

⚠ Recent origination dates of attributes, such as emails, social media and others, could indicate a *pop-up* identity.

✓ An asymmetrical pattern, like the profile shown here, of *first/last seen* timestamps is indicative of a legitimate online identity evolved over a 'normal' period of time.



pipl | Edward Donovan | Location (optional) | Advanced | ROBERT

Results for
E
29 years old
United States
Telecommunications Associate at Viacom
Erica Donovan

Original Search
Edward Donovan

Print

Edward J. Donovan
Male, Speaks English
29 years old (born May 7, 1991)
From the California and Colorado | Map View

Emails
edjdonovan123@gmail.com Personal
shooter35@gmail.com Personal
edward.donovan@viacom.com Work
eddydonovan@protonmail.com
edonovan@fsu.edu

Phones
415-223-4298 Mobile

Places
94 Battery Ln, San Jose, CA Personal

Career
Telecommunications Associate at Viacom 2017-2020

Usernames
eddy_donasaur Twitter
edwarddonovan12 Facebook

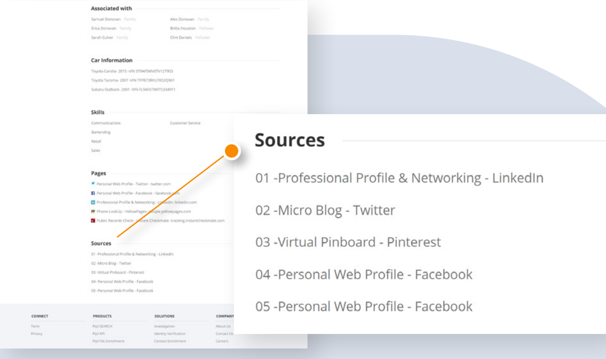
Counts (Density of Attributes)

Content rich profiles are needed to make this kind of calculation effective. The right identity data index must contain a high degree of coverage of many different attributes (depth/breadth).

🚩 Extreme sparsity of attributes can be indicative of a synthetic ID.

⚠ Sparsity of attributes with minimal to medium source-to-source corroboration could indicate 'work-in-progress' synthetic identity records.

✓ High attribute match rates (especially when combined with asymmetrical timestamps) is indicative of a legitimate identity record.



Associated with

Samuel Donovan Family	Alex Donovan Family
Erica Donovan Family	Britta Houston Follower
Sarah Culver Family	Clint Daniels Follower

Car Information

Toyota Carolla- 2015 -VIN ST9AF5MV0TV127903
Toyota Tacoma- 2007 -VIN TFP873RXU7832Q961
Subaru Outback- 2001 -VIN FL9AF6TW0TS334911

Skills

Communications
Bartending
Retail
Sales
Customer Service

Pages

Personal Web Profile - Twitter - twitter.com
Personal Web Profile - Facebook - facebook.com
Professional Profile & Networking - LinkedIn- linkedin.com
Phone LookUp - YellowPages- people.yellowpages.com
Public Records Check - Instant Checkmate- tracking.instantcheckmate.com

Sources

01 -Professional Profile & Networking - LinkedIn
02 -Micro Blog - Twitter
03 -Virtual Pinboard - Pinterest
04 -Personal Web Profile - Facebook
05 -Personal Web Profile - Facebook

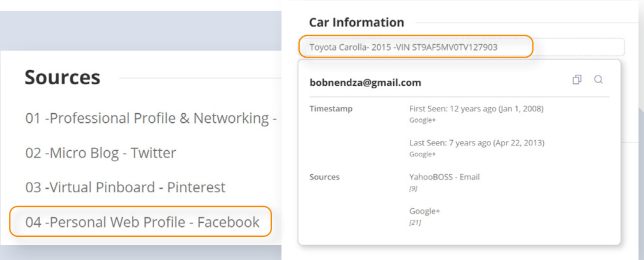
Number of Sources

Identity indexes that aggregate many data sets from a comprehensive source network is essential for connecting enough characteristics to render an insightful and rich digital footprint. It's also important for minimizing false positives amongst 'thin-file' applicants.

⚑ Extreme sparsity of linked sources is indicative of a synthetic ID, as whole records can be built from a single social media profile.

⚠ A minimal or medium number of sources could indicate synthetic ID fraud, but can also indicate 'thin file' applicants.

✓ A high number of sources (especially when combined with asymmetrical timestamps) is indicative of a legitimate identity record.



Sources

01 -Professional Profile & Networking -
02 -Micro Blog - Twitter
03 -Virtual Pinboard - Pinterest
04 -Personal Web Profile - Facebook

Car Information

Toyota Carolla- 2015 -VIN ST9AF5MV0TV127903

bobnendza@gmail.com

Timestamp	First Seen: 12 years ago (Jan 1, 2008) Google+
	Last Seen: 7 years ago (Apr 22, 2013) Google+
Sources	YahooBOSS - Email [9]
	Google+ [21]

Corroboration Between Online & Offline Sources

It's easy for perpetrators to create identity records in highly available online data systems like social media accounts. Accurately matching online attributes with offline records, such as street address, VIN or other 'real-world' features, adds another layer of intelligence to crackdown on synthetic IDs and reduces 'thin-file' false positives.

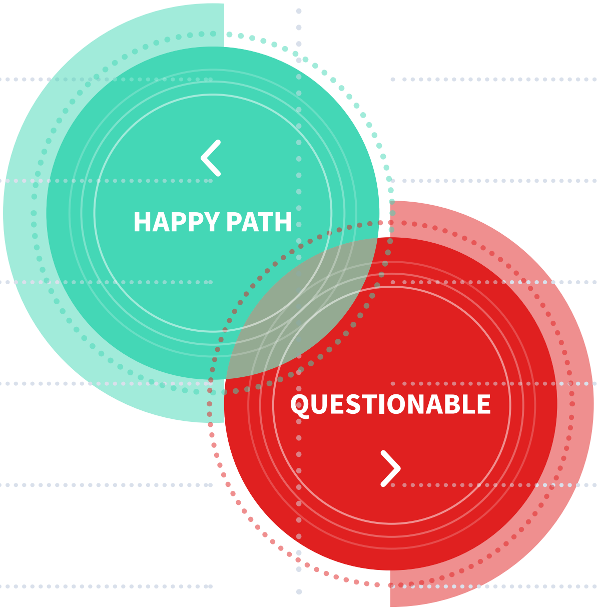
⚑ A complete absence of on/offline matching could be indicative of a synthetic ID, especially if a street address history is missing.

⚠ A minimal degree of on/offline corroboration may be a sign of a 'work-in-progress' synthetic identity record. It may also be a 'thin-file' applicant.

✓ A high degree of source corroboration (especially when combined with asymmetrical timestamps) is indicative of a legitimate identity record.

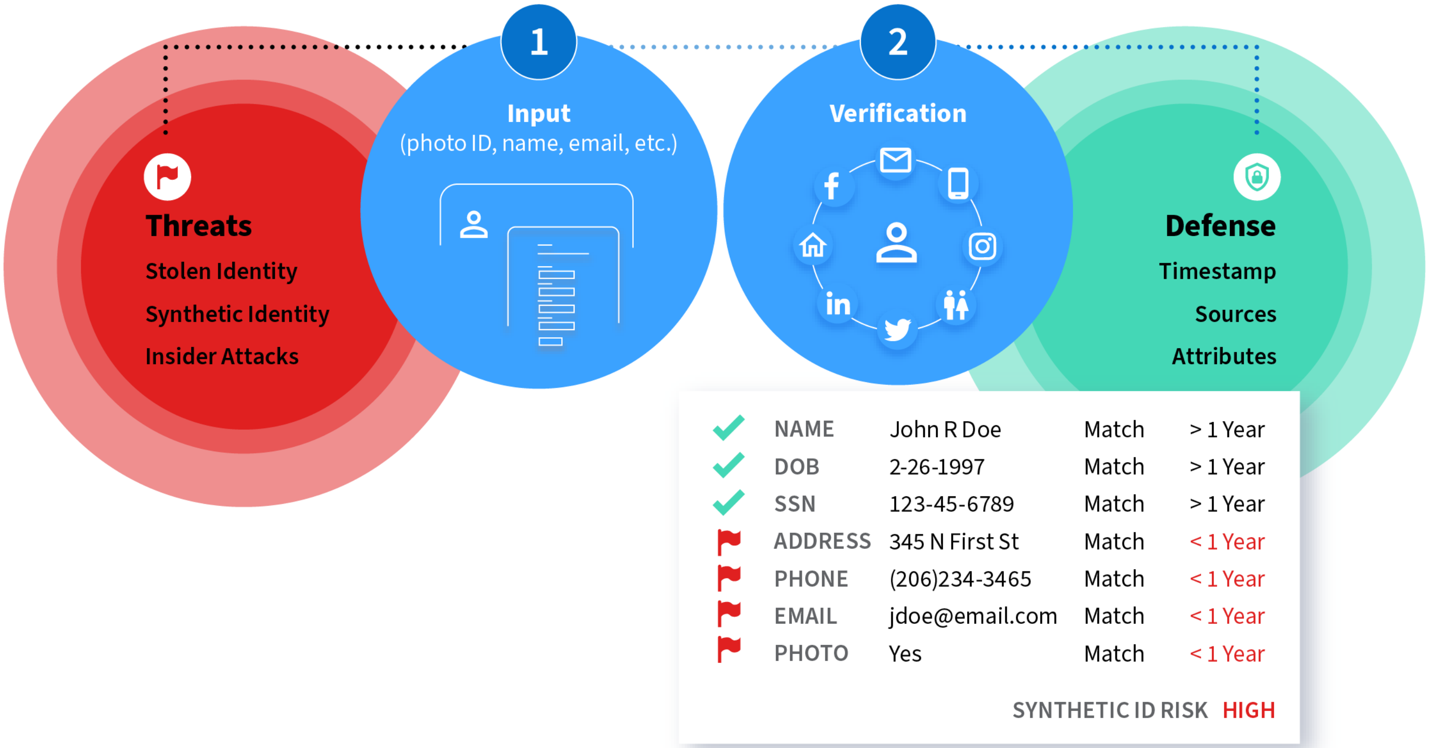
Create your own Happy Identity Path

- Asymmetry - first and last seen dates
- Strong blend of online and offline records
- Profile richness - number of data fields and sources
- Many matched associations with other people (with many sources)
- Multiple physical addresses with asymmetrical timestamps
- Multiple Vehicle Identification Numbers



- Addresses near large international airports or shipping areas
- Multiple applicants with same address or phone number
- Credit file depth is inconsistent with customer profile
- Using secured credit to build history
- Social Security Numbers issued after 2011
- Multiple accounts from one IP address
- Multiple authorized users on the same account

Synthetic Identity & Application Fraud



Threats

Stolen Identity

Synthetic Identity

Insider Attacks

1

Input

(photo ID, name, email, etc.)

2

Verification

Defense

Timestamp

Sources

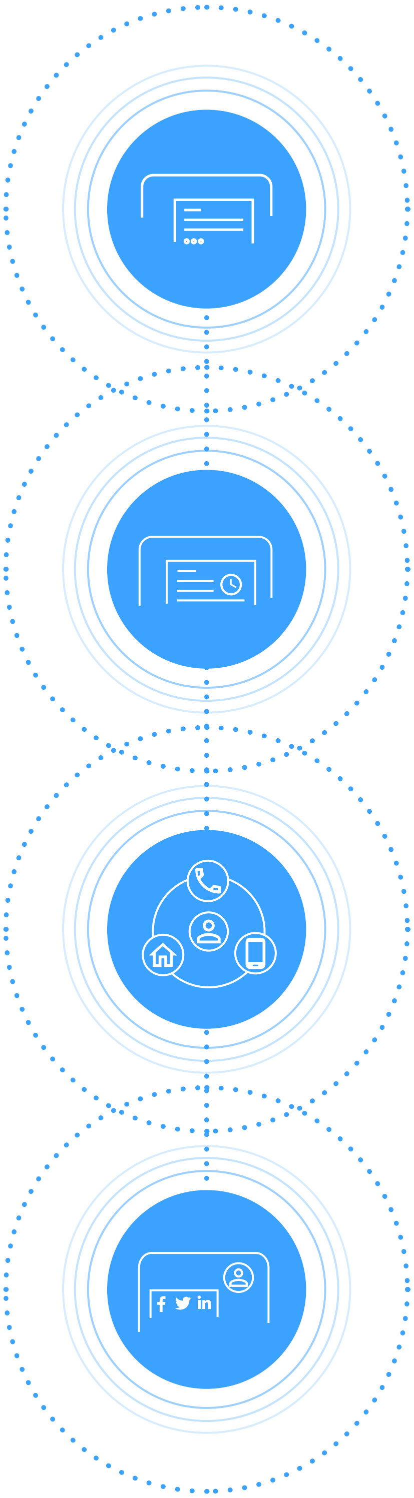
Attributes

✓	NAME	John R Doe	Match	> 1 Year
✓	DOB	2-26-1997	Match	> 1 Year
✓	SSN	123-45-6789	Match	> 1 Year
⚑	ADDRESS	345 N First St	Match	< 1 Year
⚑	PHONE	(206)234-3465	Match	< 1 Year
⚑	EMAIL	jdoe@email.com	Match	< 1 Year
⚑	PHOTO	Yes	Match	< 1 Year

SYNTHETIC ID RISK **HIGH**

Additional Identity Intelligence

The Pipl response returns a complete online/offline historical footprint of an identity, which contains additional layers of information that may be useful. The following data points are worth considering as you design a model to make a determination about an identity:



Counts

The quantity of data returned may correlate to confirming a real identity.
Data Field Counts — The Pipl API response contains a section called “available data” indicating a summary for the number of data fields associated with the person (e.g. the number of emails, phones, addresses, etc.). A real person typically has multiple emails, phones, or addresses over the course of their lifetime.

Number of sources — The identity of a person is created from various public source records. A real person typically appears in many public sources.

Time Stamps

Pipl transparently displays first/last seen timestamp information for each identity element.
Timestamps can be an accurate indicator of synthetic versus real identities, as real people typically appear in public records over a long period of time, proportionate to their age. While synthetic identities often have little history and short durations between first and last seen dates.

Data Types

Knowing more about the type of data may correlate to confirming a real identity.
Pipl returns metadata that describes the type of specific data fields. For example, an email is indicated as personal or work, whether the email is hosted by a free service provider or a one-time, disposable email service. Similarly, phones are marked as mobile, home, or work.

Boolean Indicators

Knowing more about their existence may correlate to confirming a real identity
Pipl returns social media data from networks such as Facebook, Twitter, Linkedin and others. The existence of social media profiles over time may be an indicator of a real identity. For example, a person who has several social media accounts that have been in existence for several years is more likely to be a real person versus one who just created a social media account last month. Additionally, you might find that the mere existence of a job in a profile is a valid signal.

Our digital world runs on trusting who is behind an online identity, but the very concept of identity has fractured into hundreds of data points that fraudsters constantly seek to exploit. **See why Pipl is the first choice when companies need to verify whether identity data actually belongs to the person using it.**