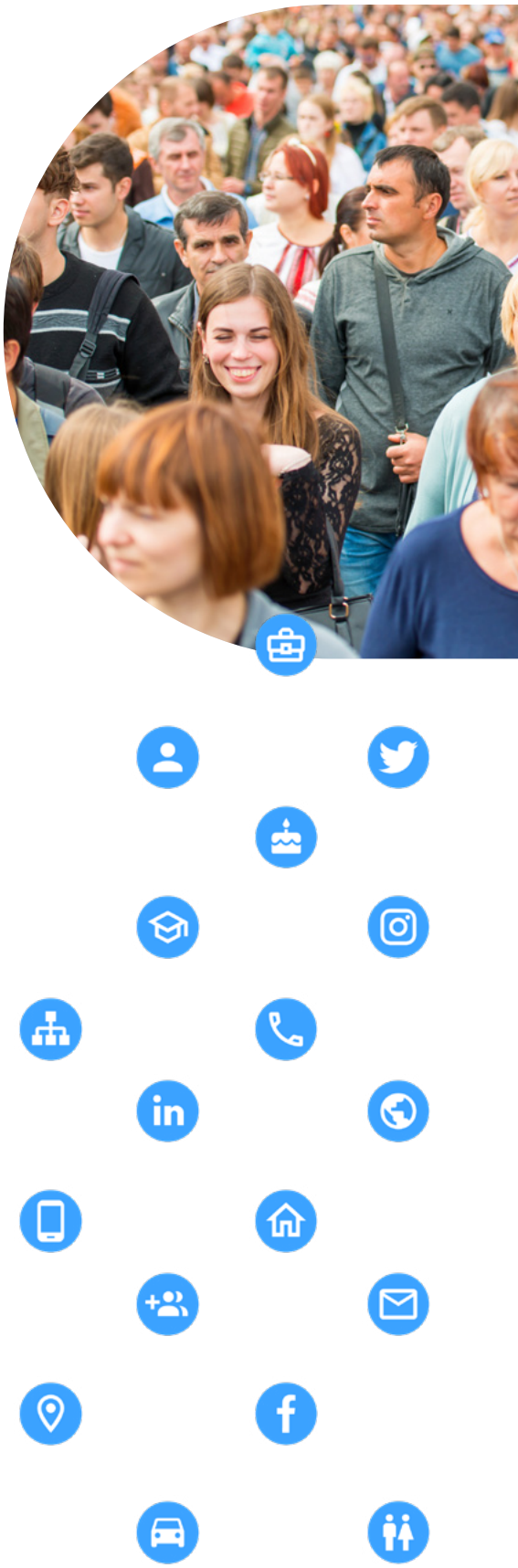
pipl

Using Online Identities in Professional Investigations

Executive Summary

The traditional notion of identity has changed significantly in the past decade. Digitalization and the Internet have created a global online world in which people navigate information, products, and services using digital identities. This paper introduces the concept of online identities and describes how Pipl enables investigators to quickly and easily discover important details and connections related to persons of interest.

The Changing Nature of Identity



The concept of identity has become complicated. Finding people used to mean time-consuming searches, combing through traditional "on-the-record" government and other institution-issued documents. For example, a person's birth certificate substantiates when and where a person was born. A driver's license can add details about physical attributes such as appearance, hair and eye color, and address at the time of issue. Credit records can help uncover details about purchase patterns. These traditional identification vehicles have been important investigative tools for decades. Today however, the Internet, mobile devices, and social media sites are driving a proliferation of digital identity elements.

Identities are Now Fragmented

Each time a person subscribes to a web service, logs onto a new site, or engages in social media activity, they generate multiple identifying characteristics. These often include email addresses, usernames, mobile phone numbers, IP addresses, and cookies.

Websites create user accounts and collect identity elements from the people who use the site. This enables users to easily navigate content and services and be recognized when they return to a site later. An e-commerce site collects identity elements such as account set-up, payment method, "liked" products, and browsing attributes. A social media platform might collect identity elements such as contact preferences, liked topics, and interactions with other users. A person might use one username to comment on a news site and a completely different username on their social media pages.

The average social media user has 8.9 social network profiles alone. Add their online banking, shopping, blog, and membership accounts, and it's easy to see how most Internet users are associated with numerous accounts and dozens of identity elements.

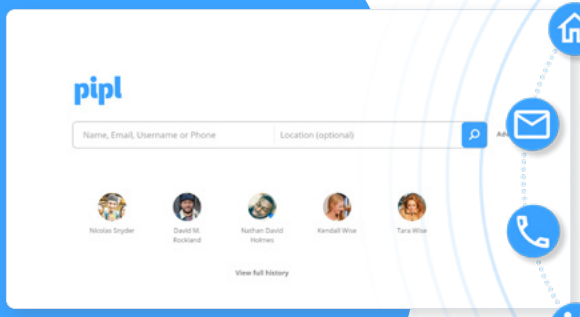
Individuals control some of their identity elements, but other pieces, like browser cookies, are controlled by websites or are completely invisible. Stolen identities are examples of identity elements that are no longer completely controlled by their true owners. Individuals intent on committing fraud often use stolen identity information to create synthetic identities that combine factual and fictitious elements. This complexity can make it difficult to find individuals, corroborate information, and uncover connections between people.

[1] https://financesonline.com/number-of-social-media-users

What Now?

Professional investigators, analysts, and researchers—as well as other professionals whose job it is to find people—now face an intimidating task. Traditional sources of information provide a start. For consumer-level people searches, it's easy to type a name or keyword into a leading search engine and find basic information. However, search engines index **less than 10%** of all web content. Investigators need more data to investigate cases and uncover important connections.

Finding the right information can take hours or days. Collecting and cross-referencing data to substantiate a case can take days or even weeks of painstaking research. Additional intelligence often lies in the deep web, but finding it requires complicated queries and coding. Even then, a lack of effective analysis tools means that search results invariably contain pools of disconnected or irrelevant information. It's clear that traditional search methods simply can't keep up in a world of fragmented identity elements scattered across the global Internet.



Pipl Pulls Together the Pieces

Pipl is the leading provider of online identity information, and our Pipl SEARCH and Pipl API products are reducing customer friction, case resolution times, and risks associated with fraud. We serve investigation professionals in insurance, e-commerce, financial services, legal, government, and law enforcement.

Pipl enables investigators and researchers to build their cases with confidence. They can:

- Find persons of interest
- Verify information about known persons of interest
- Uncover and visualize important details and connections
- Gain relevant, actionable insight

Designed for High-Stakes Applications

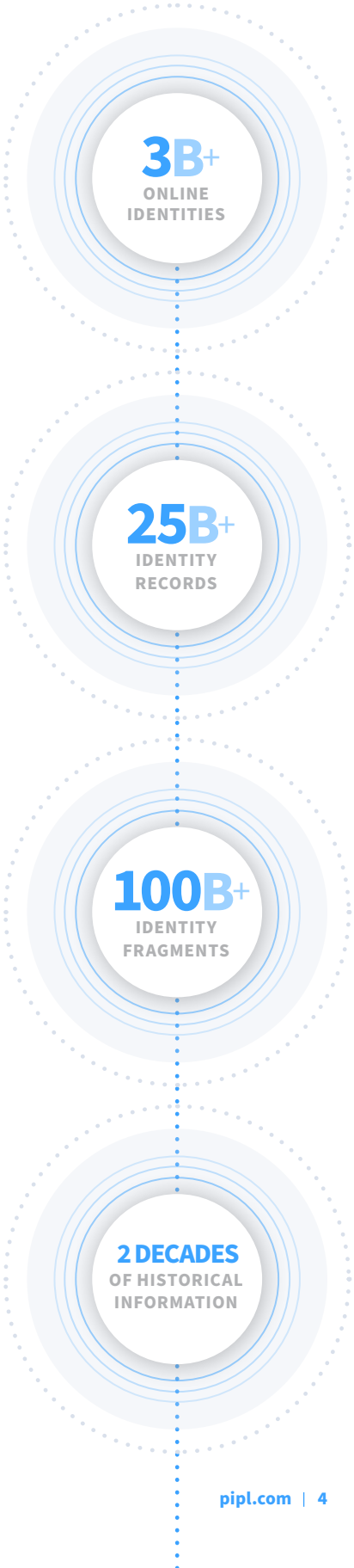
Most identity information repositories are built on large databases containing offline data, which is usually relevant only within a specific geographic region. These sources are only updated when a person's account or official record changes—for example, a change in marital status or an address change.

Unlike traditional databases, Pipl delivers immediate, relevant insight to help investigators find people and resolve cases faster. We continuously scour the web, social media, and exclusive sources to collect identity elements, as well as other physical and digital data. International data sources provide unmatched global coverage.

Our powerful algorithm performs sophisticated data analysis to deliver more than 3 billion online identities, based on more than 25 billion identity records built from more than 100 billion identity elements amassed over two decades.

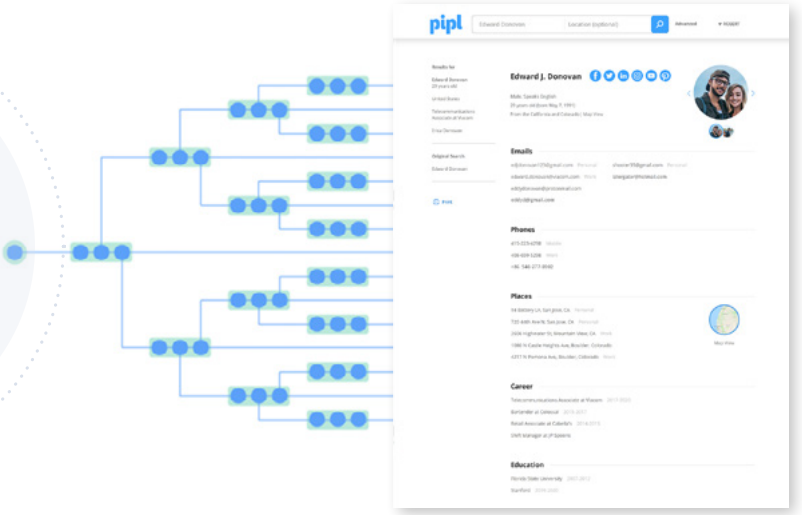
The result is an index of high-confidence, online identities including:

- 2.6 billion phone numbers
- 1.5 billion email addresses
- 97% coverage of U.S. adults; 90% with phone numbers and 63% with social profiles



How Pipl Works

When an investigator begins a search, the Pipl identity resolution engine executes multiple recursive searches to deliver rich, reliable profile data in seconds. Our advanced algorithm finds records that match the initial search term. It then identifies other identity elements within these records and executes searches against those elements. The algorithm continues to search and find other matches until there is no more data to be found. As it searches, it cross-references each identity element between records to statistically verify data accuracy and uncover connections. The result is a trusted online identity.



The Pipl identity resolution engine continuously analyzes, corroborates identifiers, and makes important connections.

Applications

Investigators use identity resolution to enable graph analytics or render a “single view” of subjects. Pipl supports many different use cases, including:

- Private investigation
- Journalistic investigation
- Criminal investigation
- Cybercrime investigation
- Background investigation
- Anti-money laundering (AML)
- Bail enforcement
- Anti-terrorist intelligence
- Know Your Customer (KYC)
- Asset recovery
- Benefits fraud
- Tax fraud
- Identity theft
- Fraud, waste & abuse substantiation
- Property & casualty insurance claims fraud

Pipl in Practice

We partner with organizations that have risk management, investigation, and due diligence needs. Our customers span organizations from government agencies to global brands—helping them find individuals, corroborate information, and uncover connections between people. Although each customer has specific uses for Pipl, the following example provides a deeper understanding of how the Pipl identity resolution platform works when searching for a particular subject.

Stage 1

An investigator enters a subject's name and last known location to set search parameters. The name and location are matched to a LinkedIn account in the Pipl index. Pipl finds an email address in the LinkedIn profile and begins building a collection of related identity elements for this person.

The subject's name and location also match a Craigslist advertisement (corroboration) with an attached phone number, which is added to the collection. The ad contains an email address that differs from the LinkedIn email address. Because the name, phone, and location match one another, the alternate email is added to the collection.

The subject's LinkedIn profile and Craigslist ad are now identifiers, because they contain multiple identity elements. These include a phone number, two email addresses, a work history, and a location, which corroborate the data.

Stage 2

The Pipl algorithm now conducts separate, individual searches of the phone number, two email addresses, and location. The Craigslist email address matches both Facebook and Twitter social networking sites. With these matches, Pipl adds the new social network identifiers to the collection. When identity elements match multiple times, they pass the test for statistical probability and increase the Pipl confidence score—making it more likely to produce a trusted online identity. The Pipl algorithm repeats these incremental searches on each identifier until it finds no new identity elements for the subject. Pipl then displays results according to a list of high-confidence matches and a separate list of possible matches.

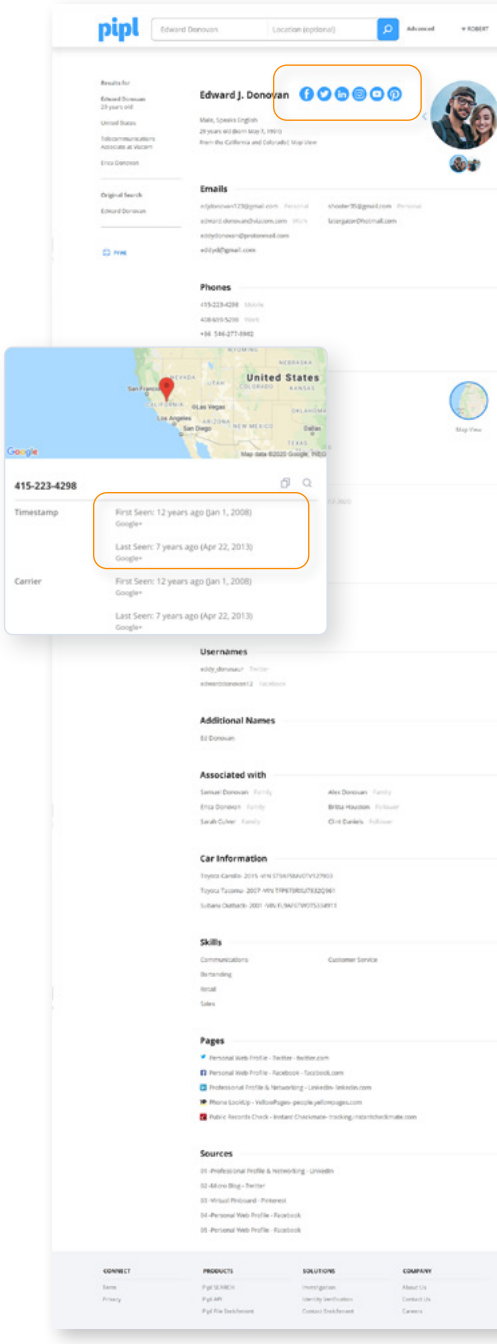
Stage 3

Pipl combines the identity elements and identifiers into a straightforward, organized online identity, as shown in Figure 3. Data is arranged in order of corroborating matches and the most recent sources for at-a-glance ease of use.

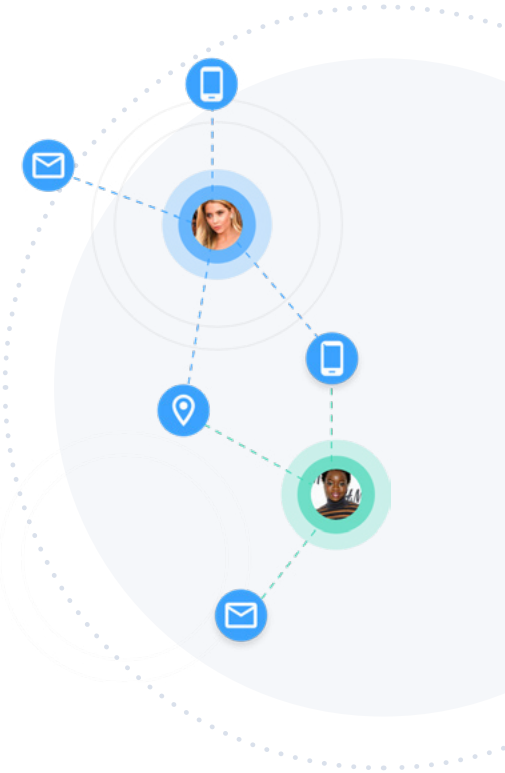
Stage 4

Pipl is interactive, enabling investigators to explore new avenues. Every click initiates a new search on a specific identity element. For example, investigators can click on a social media icon or username and go to the user profile. They can execute advanced searches to look for people in different locations, to explore associations between people, and to refine their results. Hovering over each identity element displayed shows underlying details including metadata, source data, and timestamps.

An organized, interactive online identity



Collect, Corroborate, and Connect



Professionals looking for online identity information rely on Pipl to save significant amounts of time. Results are delivered within seconds. Users receive fresh data, because as a real-time search engine, Pipl continuously updates data to track a subject's evolving identity elements and identifiers. Up-to-date information makes online identity data actionable—investigators can quickly know where a person might be living and how to contact them.

Social media profile information can deliver even more recent and detailed insights about a person. Pipl enables investigators to connect individuals with social media names, even when the names don't match. These connections can shed additional light on a person's activities, associations, and movements. Pipl helps minimize investigative "dead ends" and enables professionals to follow more paths and gain better context for decision making.

Visualizing Connections

With Pipl, online identity data resides in one place. Investigators no longer need dozens of different tools and web searches to find, retrieve, organize, and analyze information from multiple disconnected sources. Pipl online identities make it simple to see critical information at a glance and then drill down for comprehensive underlying detail.

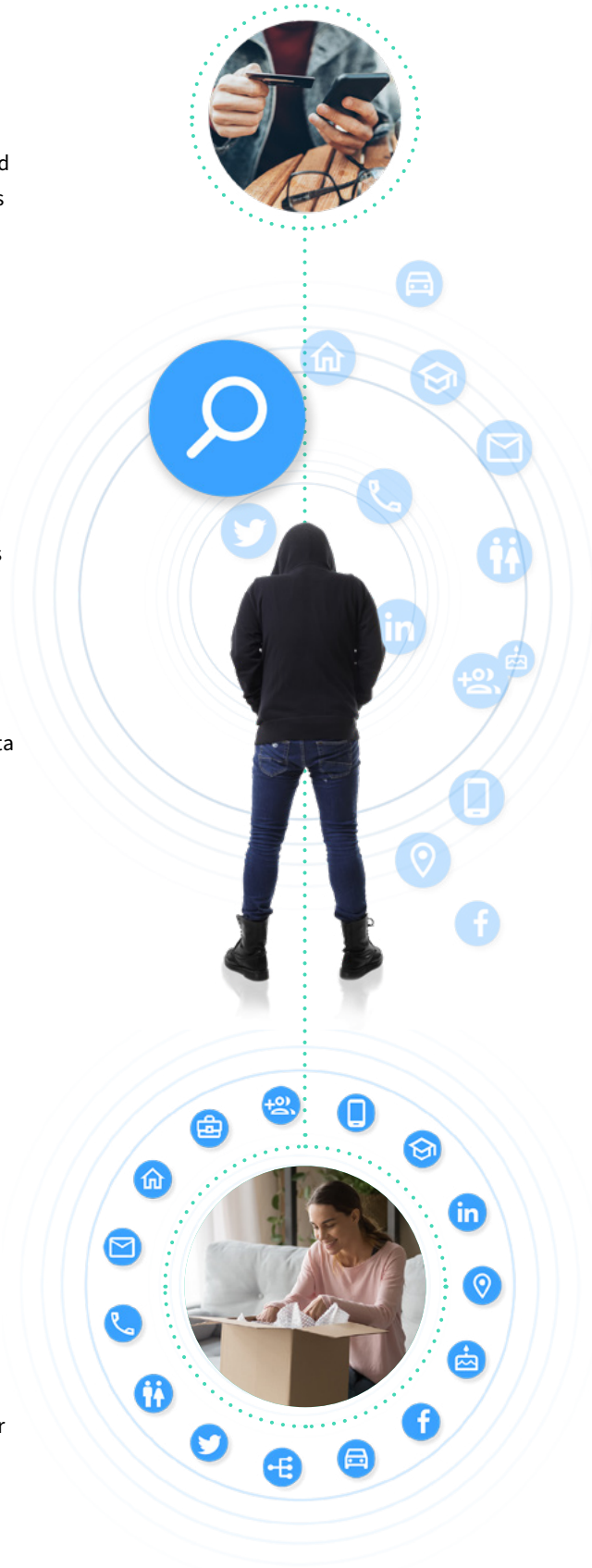
Storyboard capabilities immediately uncover and illustrate connections between people. This allows investigators to compare online identity information from several people and overlay the information to instantly discover important connections. Within seconds, an investigator can determine the relevance and importance of people's associates to their specific case.

Fusing On- and Off-the-Record Data

Finding reliable data for thin-file subjects can stymie an investigation. Often, there is a lack of data about millennials, multinationals, and people in jurisdictions not covered by traditional identity data sources. Almost 1 billion people in the world have no on-the-record identity at all. Even with access to international databases, on-the-record information does not provide a comprehensive understanding of a person and their identity.

However, digital capabilities and mobile communications make it possible to improve investigative results with online identities. Currently, there are **4.8 billion mobile phones** in the world—of which 3.5 billion are smartphones. This means even individuals in developing countries or remote areas without on-the-record data are building online identities that can be discovered by Pipl.

Pipl fuses identifiers from both on-the-record and off-the-record online data. This makes it possible for investigators to search for subjects who might cross jurisdictions or operate sporadically. Pipl's ability to integrate and analyze different data types also adds dimensions that investigators need to dynamically "connect the dots" in scenarios such as these:



Locating a Person of Interest

- Discover if a person's name and last seen residential address (offline) match online data such as usernames, social media handles, and email addresses (online)
- Learn if the person is using aliases
- Look for a person's associates living at the same address
- Search for a person using a different name or different spelling of the name
- Link online social handles and email addresses to the same physical address
- When a person of interest can't be found through an offline data source, use online social media data to gain clues about their current location and activities
- Uncover relevant associations between parties (both on and offline) to help locate a person of interest

Uncovering Associations Between Persons of Interest

- Identify associated people, such as family members, living at the same address
- Discover relationships between people through social (online) data that uncovers similar educational experiences, common causes, or special interest groups (both off and online)
- Learn if people ever shared the same address
- Use data from professional networking sites (online) to find out if individuals have ever worked together

[2] Smartphone and Mobile Phone Users Worldwide, BankyMyCell.com, August 13, 2020

Proving or Disproving Claims or Testimony

- Corroborate when online sources indicate activity or behavior that conflicts with what's been represented to an investigation
- Support subpoena requests with data from a person's social media and other online accounts that substantiates that the content being requested exists
- Enrich offline information with online identifiers to determine when a false identity is being used
- Uncover synthetic identities faster with online data that indicates whether the subject is a real person or not

Secure Information Management

As a SOC 2-certified vendor, we manage our platform's information security, availability, processing integrity, confidentiality, and privacy. This provides customers with assurance that they can maintain compliance with privacy and security regulations. It also reduces the risk of compromising the integrity of their information systems. Pipl customers can implement our software-as-a-service platform or API and confidently integrate Pipl with existing data and analysis platforms.

ABOUT PIPL

Pipl is the world's leading provider of online identity information. Our Pipl SEARCH and Pipl API products are reducing customer friction, case resolution times, and the risks associated with fraud. We serve fraud and investigation professionals in insurance, e-commerce, financial services, legal, government, and law enforcement. Pipl's unmatched global coverage includes more than 3 billion identities cross-referenced from more than 25 billion individual records to create the leading online identity index.