A guide to digital identity verification: **the technology and trends**

onfido

Contents

Executive summary

Identity risk management. It's more important than ever. Traditional businesses are taking their operations fully digital, and newer all-digital businesses are growing. Both groups need to onboard new users to their services remotely. But they also need to monitor the risks to their platform as their user base increases.

An all-digital world amplifies the classic threats of service misuse, identity impersonation and money laundering. Businesses want to offer quick, convenient access. But they also want to keep fraudulent users off their platform. And everyone's struggling with the same trade-off.

Users worry too, about keeping their identity information secure. And rightly so—the current state of play is fragmented and out of date. Identity information should not be this easily duplicated or poorly stored. Data breaches should be the exception, not the norm.

So what does this mean? Verifying a person's true identity is the key to managing risk.

In the past we've seen a few different ways to verify an online user's identity. None of them are perfect. Most of them aren't fit for purpose. Geographical reach, user experience and fraud prevention exist in a balance.

The solution—a process that works for users and companies. Let users use the most commonly available forms of legal identity as proof. Let companies get true assurance that users are who they claim to be by combining users' legal identities with biometric analysis.

Doing this at high volume globally is tough. A human-only approach can't deliver accurate results at scale. Thankfully, artificial intelligence (AI) and machine learning (ML) techniques have finally matured—and with them, we can solve online identity: businesses can confidently onboard anyone, anywhere in the world with an identity document (ID) and an internet connection.

This white paper will explain the ML techniques behind Onfido's market-leading technology.

Background

The number of smartphone users is forecast to grow to around 2.9 billion in 2020[1], with smartphone penetration rates increasing too.

Consider the shift in expectation. Users of online services are now connected 24/7 and gravitate towards simple digital user experiences.

What was once known as 'e-commerce' is now simply 'shopping'. And 'online banking' is now just 'banking'. In the iPhone era, mobile has become the new norm. So businesses meet consumers wherever they are. In-person interaction? No longer required.

International transfers, P2P payments, check cashing services and loan origination—all of these transactions now happen in a more convenient and less expensive way, from the consumer's phone. Although it may seem slow to an outsider, traditional brick-and-mortar financial institutions have largely embraced the mobile experience, which has forced them to re-think transactional workflows. A smartphone in a user's palm can be a document scanner, for instance. And with that realization, hundreds of millions of customers can deposit checks without having to visit a branch or ATM. And we're only just starting to unlock the potential of using smartphones as document scanners.

Outside of financial services, entrepreneurs are launching new businesses in hours, not weeks or months. Cheap cloud storage, website design and development, payment processing and even payroll services are a click away. The rise of the sharing economy has created opportunities for anyone with an underutilized asset—primarily properties and vehicles—to benefit economically from renting it to a stranger. The digitization of the economy has also disrupted longstanding for-hire services like taxis and rental cars. And it's made on-demand services possible, with new business models for dog walkers, house-sitters and food delivery.

Something had to give. More online services means more consumer data—and when companies have more data to manage, we see plenty of mismanagement and misuse. Poor data protection has damaged consumer trust and led to unprecedented regulation. We hear about another mass-scale data breach every other day. With so much data being leaked, it's embarrassingly easy for fraudsters to mimic another individual, use a stolen credit card or take over an account. Juniper Research estimates the global cost of data breaches will rise to $2.1 trillion by 2019. The United States alone accounted for 47% of the world's data breaches[2] in 2017.

These data breaches compromised personal information belonging to hundreds of millions of consumers. So the public and private sectors are searching for preventions. These preventions wouldn't rely on a central identifier, or on a single central database managed by a private entity. The last few years have seen increased momentum around this consensus, thanks to the activities of several industry working groups, including The Better Identity Coalition[3].

To regain consumers' trust, industry and government regulators have provided some basic principles to guide data usage and protection. Even sovereign states have had to ensure that legitimate online businesses aren't unknowingly

helping move money for terrorist financing or other nefarious activities. So businesses now have to meet several compliance requirements to protect against terrorism, money laundering and fraud.

In the US and Europe, a number of core regulations for the banking industry have driven financial services companies to put compliance measures in place. These include the Bank Secrecy Act of 1960 and the 2001 Patriot Act, signed after September 11th in response to the discovery of a wide economic network established by terrorists. Banks also have to ensure a meaningful level of customer due diligence (CDD), commonly referred to as KYC or "Know Your Customer" programs, to ensure banking transactions are monitored and not anonymous. KYC directives include establishing the identity of an institution's customers, and understanding the nature of those customers' activities.

The US Financial Crimes Enforcement Network, under the Department of the Treasury, amended existing Bank Secrecy Act regulations to clarify and strengthen CDD requirements for certain financial institutions. The CDD Rule outlines explicit customer due diligence requirements and imposes a new requirement for financial institutions to identify and verify the identity of beneficial owners of legal entity customers—although there are some exclusions and exemptions.

And of course there's the famous GDPR. In Europe, the EU's new General Data Protection Regulation came into effect in May 2018 and set out a new and rigorous set of data privacy requirements. Global companies that collect information from EU citizens have to comply with GDPR regulations or face significant fines. Under GDPR, companies have to implement data protection systems, including policies and procedures for managing and protecting data.

Regulations will continue to introduce controls to protect consumers transacting online. But we can't undo a data breach. With so much data compromised, personally identifiable information (PII) can no longer be the cornerstone of identity when it comes to transacting online. The damage to consumer trust, too, can't be ignored. It's now a significant blocker to widespread adoption of online services. Platform trust is fundamental to the new digital use cases presented above working over the long term. New regulations should help build consumer trust, but at a cost. They'll add friction for users. And many online business models rely on new users signing up quickly and using the service regularly. So these protections need to be considered carefully against user experience.

Protections need to be considered carefully against user experience

This has intensified a challenge that has existed for businesses since the start of the internet age—how do I know a person is who they claim to be? How can a financial services company maintain regulatory compliance around KYC when users are 100% online for every transaction and interaction?

To address this challenge, businesses try to validate identity information provided by users. If the business believes the information's authentic, the user can open accounts and transact online. This process identifies the user and ensures the business is compliant. But it's tough to do it without slowing down the sign up process. It's a trade-off between speed and security.

There are two fundamental ways to implement these protections:

1. A very heavy initial vetting to provide assurance that users are known identities, and that they don't have a previous history or potential of conducting a financial or service crime. This initial vetting is called identity verification and is the core focus of this white paper.

2. A lighter weight vetting that can yield a lower level of service access, but that's coupled with a higher monitoring of the user's behavior once they're in the system. Assuming over time that they don't trigger additional risk thresholds, they can be granted higher levels of service access.

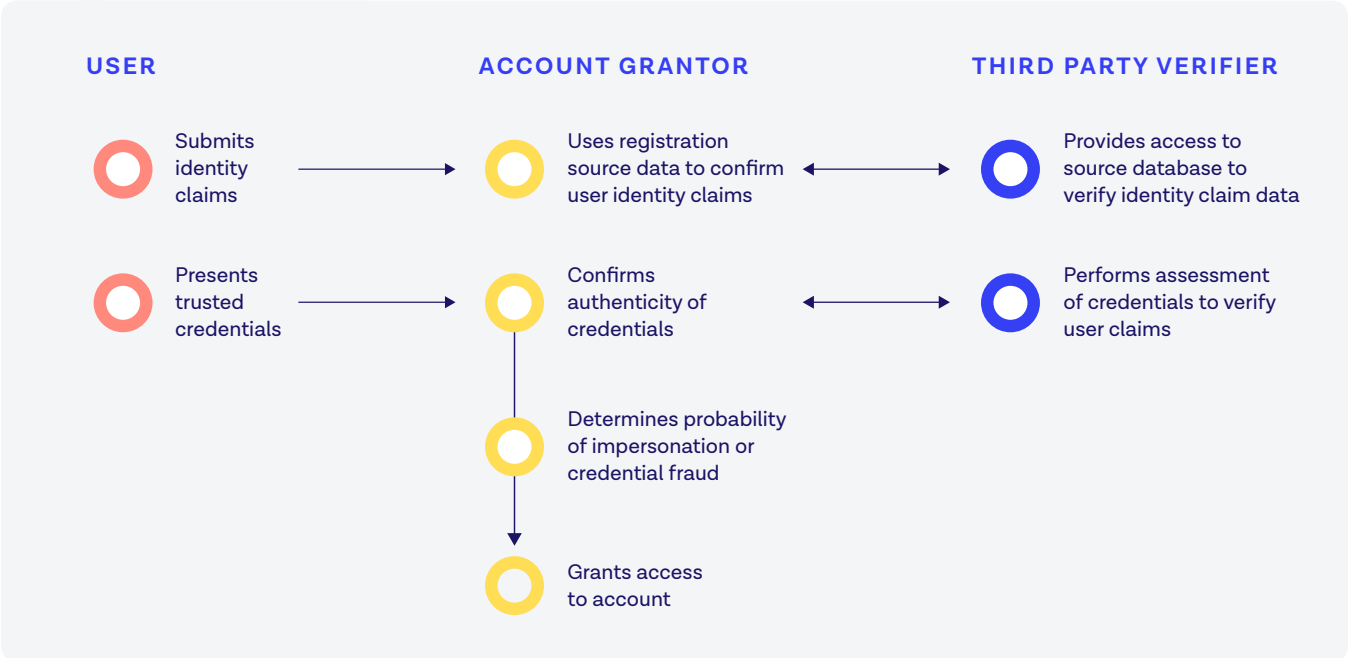
Newer technologies like AI support businesses in gathering important insights about their customers.

The next sections will analyze various approaches businesses take to verifying identity information online.

Identity verification

Identity is a set of claims that can be used to describe a unique person. These claims can include permanent traits like date of birth, ethnicity and fingerprints. Or they can include semi-permanent traits like height, weight, eye color and name. These attributes only change slightly over a person's lifetime, or are very difficult to change. Identity claims can also relate to housing and communications information—like home address, phone number and email address—which don't change often, but easily could. Looking at these traits, it's clear that it's not just the current version of the trait that's important. The history, or footprint, of this trait is just as key.

Identity verification is a process that validates these identifying traits and verifies them against a real, physical human. That person can then be issued a digital credential, which they can use to gain access to online accounts, services and activities. When someone claims their identity with a relying party for the first time, their claim needs to be verified to ensure they are legitimate. This use case is typically called 'account onboarding'. It's one of the primary drivers of identity verification. We're familiar with this process as it's core to account activation (e.g. opening a new bank account), for example, or travelling on an airplane. For relying parties, the concept of identity evolves over time as the interactions with the user change in nature. So identity verification after onboarding can also grant access to new services, provide stepped up or tiered services with added privileges, or re-assess the user if questions arise about their identity over time. The general process is:



There are three main things to consider when verifying an identity:

1. **What we know** (e.g. 'out of wallet' questions that can be stored in a database, like your monthly mortgage payment)

2. **What we have** (e.g. a token, credential or trusted ID)

3. **Who we are** (e.g. face, fingerprint or voice—also known as biometrics)

For the a reliable result we'd need to use a multifactor approach. Certain components are pointless without others— a fingerprint doesn't mean anything unless it's associated with a specific name or ID, for instance.

Given the billions of personal records leaked in recent data breaches, claims made using personal information alone ('what we know') have little credibility. They won't prove that a person is who they claim to be. Instead, we need to link the claims to a trusted source so we can determine that the person claiming the identity actually owns it.

A trusted source could be an authoritative registration system—often managed either by the government or by the person directly—to ensure its reliability and consistency. For example, a birth record is an authoritative record issued at time of birth in a hospital. That record then flows into a unique government or citizen ID, and then into the health and education systems as the baby becomes a child. As the user grows up, they get driver's licenses and government-issued passports or national IDs— the three main forms of photo ID. They provide a reliable credential which gives additional assurance of that person's identity.

The ID alone can provide some assurance of identity. But including a human verification can give a more robust result. In the offline world, banks tend to do this by insisting that a user visits a branch to meet a bank attendant face-to-face so the user can prove possession and ownership of the ID. But nobody wants to do that anymore.

Thanks to smartphones, we can replicate the benefits of the branch visit model without the headaches. A user can upload their ID and a picture of their face from anywhere. Then mathematical comparisons between the facial image and the ID picture provide assurance that the ID owner is the person trying to open the digital account.

With video, we can take this level of identity assurance a step further. A user can submit a video alongside the ID, coupling facial comparison with 'liveness'—proof there is a living, breathing ID owner behind the camera. To do this in an online world, the user responds to random challenges in real time: "say the number 429 and turn your head to the left." The video can be analyzed in near real-time with the ID image, to check the liveness of the human and compare the face extracted from the video to the face on the ID.

Once a business verifies an user's identity, they can create an online digital identity. Account access follows, and the platform can issue a stronger credential such as a unique account number, or username/password combination.

Approaches to identity verification

So what's the best approach to verify an identity? It's important to think about:

- **Reliability -** how vulnerable is the approach to previous attacks that may affect the integrity of the source data? For example, if a database is used for testing "out of wallet" questions, it should be evaluated for any historical breaches that mean that data isn't private anymore.

- **Ubiquity -** how widespread is the approach's coverage for the group we're trying to verify? Many businesses are global—but lots of identity methods are local. This won't work.

- **Cost -** how expensive is it to maintain and protect data and credentials? Ensuring they can be trusted will incur processing costs which need to be factored in.

- **User experience -** How easy is it to use? The identity flow should be smooth and fast. After spending lots of money acquiring a customer, too much friction in the sign up process will cause the user to 'drop off'.

Methods of identity verification

Here are some of the more common approaches to identity verification:

Databases

These are systems that house data—data previously collected and verified as part of a registration system. They can be private databases run by for-profit companies, or public databases run by governments. Examples of private databases include credit bureaus and telephone directories. Examples of public databases include government identifiers (Social Security, tax or voter numbers) or the DMV that houses Driver's License data and numbers.

When using databases for identity verification, it's important to consider the cost of access, the fact that historical data breaches will have compromised the data trustworthiness and whether the data can be used commercially under current privacy regulations.

Government-issued IDs

These include:

1. **Driver's license**
2. **Passport**
3. **National identity card**
4. **Residence permit**
5. **Voter identification document**
6. **Tax identification document**

When using government-issued IDs for identity verification, it's important to consider that forgery is common. Detecting misrepresentation of the user and other types of fraud is key. Also, there's no single global schema for a consistent ID format. So to process these documents we need deep expertise, especially to do so globally. Finally, since documents are real world artifacts, manipulating them in a digital format requires some form of scanning and analytics.

Phones and phone numbers

Our phones have become extensions of our lives. To some degree, we see them as a trusted extension of a person's identity. This is especially true when many of us keep our phone number when replacing phones.

The phone itself is able to provide a fairly rich set of fingerprinting data. This is a way to identify the user, as long as the user still owns the phone. It's relatively common in many fraud risk engines within the advertising ecosystem, but it's falling under increased privacy scrutiny by privacy regulators.

As SMS is so widespread, it's an easy way to confirm ownership of a given mobile phone number. While this is a fairly reliable method, it is open to vulnerabilities that mean it could never be foolproof.

Email, social network and instant messaging identities

Much like our mobile identity, our online identity is an important part of our full identity. It can be the source of some fairly reliable components, including:

1. **Email addresses**—while many people have several email accounts (some reflecting differing levels of anonymity), email addresses typically stay around a person in the same way that phone numbers do. So they can be relied upon as unique verifiers of online personas.

2. **Instant messaging profiles**—private handles on WhatsApp, WeChat and other popular systems can give us a link to real people.

3. **Social Networks**—over a billion of us now use Facebook to power our interactions with friends and family, as well as to log into many other services. When social networks are this embedded, they can be relied upon for some inherent trust. So they can help verify certain aspects of a person's real identity.

4. **Professional networks**—LinkedIn offers a similar level of assurance through employment and educational associations, and similar network-level trust assurance, including endorsements.

Ultimately though, social networks are considered lower assurance identity systems. They were designed for casual social interactions used repeatedly over a period of time—not for out-of-the-box, instant use in high risk transactions.

Biometrics

Biometrics are the unique traits of the human body which can be used as personal identifiers. The most common biometrics include:

Fingerprints

Retina scans

Facial traits

Voice patterns

As discussed above, a biometric by itself isn't enough for an initial identity verification. That is, unless the user's biometrics were previously captured and registered to a trusted identity. So they're used as a second factor to authenticate a returning user as the account holder, or to confirm true ownership of other identity credentials.

Some of the other considerations for using biometrics are spoofing (masks, digitizers or other technological cheating), data storage and worries of surveillance.

Summary

The below table gives a side-by-side comparison of each approach to identity verification.

Reliable approaches to verification will use these methods in some combination to properly assess the identity being claimed. Lower assurance models can be used as weaker signals, or in first-pass cascade for a layered authentication.

Identity verification process: risk modeling

When conducting an identity verification, companies will balance their need to detect identity fraud—or ensure they know their user for compliance—against user experience. The balance between accuracy and speed informs a company's choice of identity verification method, and the inputs can be used to calculate an overall identity risk score.

This is called a *risk engine.* It takes in a number of signals (like the ones from the previous section) and enables a decision to be made about whether a fraudulent user is attacking the system within a certain risk threshold.

The goal of any such risk engine is to build a risk model that can identify possible frauds efficiently without affecting the large majority of legitimate users. This becomes a classic analytics challenge. Since it also currently involves so much human judgment, trained experts are also needed to design, inform, oversee and at times directly participate in the risk system.

The signals described above provide varying results in reliability, accuracy and user experience. Using the right combination of signals from a variety of approaches is the most comprehensive approach, allowing trusted users to pass through and detecting fraudsters.

To design an identity verification process, consider all of the methods in the comparison matrix and then select the set that fits the business' risk profile, speed thresholds and user experience goals. The low reliability of database, device and social network approaches means they should always be used in combination with IDs or biometrics to fully assess the identity and risk of a person online.

The next section will start to explore verification of IDs in depth, and the advantages of including it in overall identity verification and risk analysis.

ID authenticity assessment

When a user provides a legal ID it's typically one of the following types:

1. Driver's license
2. Passport
3. National identity card
4. Residence permit
5. Voter identification document
6. Tax identification document

There are almost 200 countries that issue IDs. When various valid historical versions are taken into account, the total population of IDs in circulation at a given time usually exceeds over 6,000 unique documents.

Each ID has a unique format and presentation of data. They also include security features like watermarks, checksums, and a machine-readable zone (MRZ), where the identity data is also encoded in optical character recognition (OCR) format to help prevent certain types of tampering.

Each of these factors are important when accepting an ID as proof of a legal identity. So the the Rosetta Stone of source IDs is an extremely large and critical data set.

When using government-issued IDs as a source of identity in a typical digital account onboarding process, consider the below:

1. The ID needs to be converted to a digital format and all information digitally extracted.
2. The ID needs to be remotely confirmed as authentic, in its digital form.
3. The ID holder needs to be confirmed as the ID owner through an additional process.

Like any security process, the above steps should build into a process that doesn't inconvenience the majority of trusted users, while still pinpointing a high percentage of the fraudsters trying to cheat the system.

Additionally, an offline verification process allows physical inspection of the ID. It's easy to gather contextual information for both the user and the ID. In an online process, some of these signals are absent. So it relies solely upon the digital images of the ID and document holder. These processes need to be even more effective than pure human inspection to make up for the lack of contextual signals. Other factors that can help signal a fake or impostor come into play.

ID forgeries and identity fraud

Fraudulent IDs come in many varieties. Below are some of the most common.

- **Forged IDs. Fraudsters will illegally change information on the ID so that they can modify part or all of the identity:**

  - **Changing the variable information**
  - **Inserting real pages from another ID**
  - **Removing pages or specific information**
  - **Applying false stamps or watermarks**
  - **Digitally altering or adding information on an image of an original ID**

1. **Counterfeit IDs. This is a total reproduction of the original ID. Typically a fraudster will obtain a template and insert their own information and photo. These are a popular option and can also be purchased illegally.**

2. **Blank stolen IDs. Unpersonalized original IDs leak from the manufacturing supply chain. Fraudsters then insert false information.**

3. **Fraudulently obtained genuine IDs. Fraudsters lie on their applications.They might use a photo of someone else, apply with a fake ID or use different personal details. Authorities then issue them genuine original IDs containing this false information.**

4. **Fantasy or camouflage IDs. Fraudsters create issuing authorities that do not exist or are not allowed to issue IDs. (e.g. Utopia, Rhodesia, British Honduras, World Service Authority, Republic of Texas).**

5. **Impostor IDs. The ID itself is original, but is used by someone other than the legal holder of the ID.**

6. **Compromised or sample IDs. A government-issued sample or image of an ID that's publicly available. Examples include IDs shared on the web, IDs in TV shows or presentations, or IDs reported as stolen or compromised to the police.**

> Each of these categories has different fraudulent characteristics, requiring different detection methods. To determine these fraud techniques systematically, you need a set of ID analytics to inspect each ID. These ID analytics ensure that the ID hasn't been modified, that it's a legitimate ID and that it belongs to the person presenting it.

Here is an example set of fraudulent IDs:

Original

Invalid ID number

Security feature

Image editing

Pixel tampering

Font verification

Levels of fraud sophistication

These different ID fraud approaches fall into different tiers of sophistication:

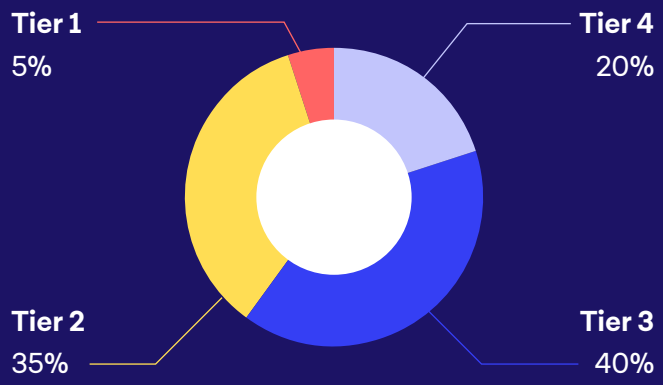
Tier 4.
These are unsophisticated amateur attempts at committing ID fraud. These include fantasy or camouflage IDs (such as the Global Citizen passport, or British national identity card) that don't exist. An estimated 20% of all fraudulent IDs are Tier 4s.

Tier 3.
These are poorly manufactured or altered ID templates. You can usually detect these cases by looking at data validations, such as ID number formats or lack of data consistency within the ID. An estimated 40% of all fraudulent IDs are Tier 3s.

Tier 2.
These are sophisticated ID forgeries and counterfeits. All of the data present on the ID is correct and makes logical sense. But trained experts can spot small variations in fonts, layout and security features—and so can a highly optimized technology detection process. An estimated 35% of all fraudulent IDs are Tier 2s.

Tier 1.
These are the most sophisticated attempts at ID fraud. Attacks on the supply chain of ID issuance and manufacture result in stolen blank IDs. Some fraudsters also illicitly obtain genuines through deceit and social engineering. There are also the highly sophisticated counterfeits available on the black market. These usually sell for large sums of money—and they're often created by criminal organizations collaborating with governments. These IDs can fool even the best-trained document experts, as well as any machine-based approach. The only way to catch these cases is through extensive analysis of all the IDs available to the individual, as well as database and biometric cross-checks. An estimated 5% of all fraudulent IDs are of this variety. This is increasing as ID verification techniques continue to evolve their own robustness.

Estimated coverage



Evaluating the various aspects of the ID image simultaneously requires a toolkit of techniques. The more sophisticated the fraud technique, the more techniques needed to detect the fake.

The next section will explore these detection techniques further.

Summary

Tier 1

- The most sophisticated attempts at ID fraud.
- Can fool even the best-trained ID experts.
- Estimated 5% of all frauds.

Tier 2

- All of the data present on the ID is correct.
- Small variations in fonts, layout and security features.
- Estimated 35% of all frauds.

Tier 3

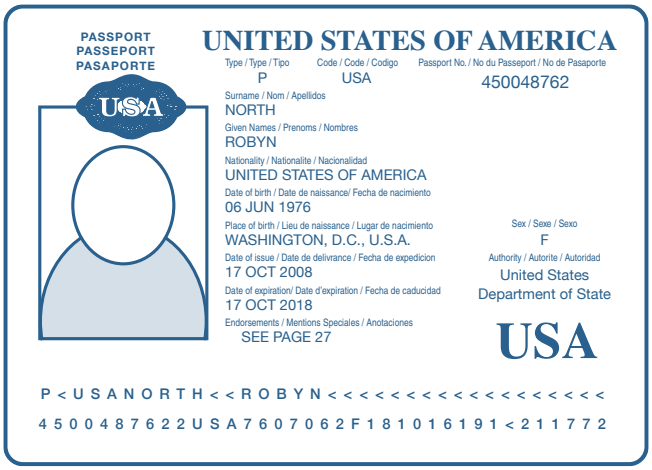
- Poorly manufactured or altered ID templates.
- Lack of data consistency within the document.
- Estimated 40% of all frauds.

Tier 4

- Unsophisticated amateur attempts.
- Fantasy or camouflage IDs.
- Estimated 20% of all frauds.

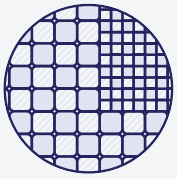
ID fraud detection techniques

For every fraud technique, there's a corresponding detection technique. Many IDs have specific design and layout templates, and built-in security features. These make it harder for fraudsters to illegally modify an ID, and easier for investigators to detect tampering.



There are three major categories of fraud analytics:

Data integrity analytics evaluates the integrity of the text and numerical fields included in the ID, both in free text and encoded in MRZ fields.



Visual document authenticity analytics evaluates the integrity of the scanned ID image, identifying any digital modifications (usually in the form of anomalies).

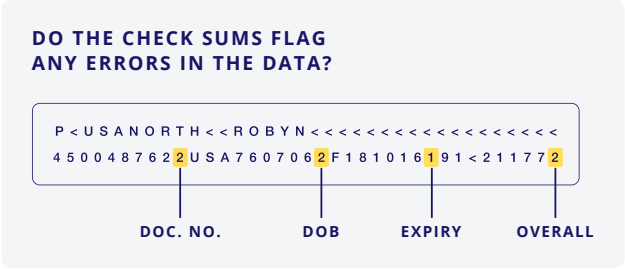
Proof of ownership evaluates a digital image of the user—the person claiming to be the ID holder—and compares it to the image on the identity card.

Data integrity analytics

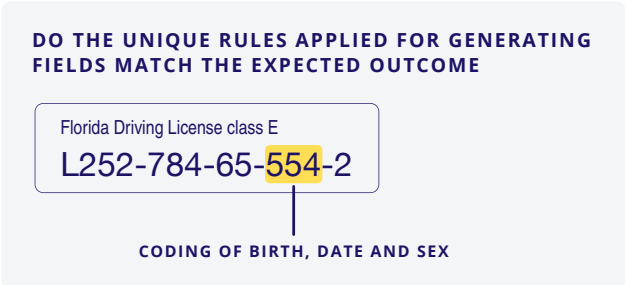
Many IDs have built in security features that can be systematically checked to ensure the ID is authentic. These include:

Data validations. An ID must contain valid data in specific formats on all areas The formats are defined by the source template. This includes encoded MRZ areas which uses check digits to protect data integrity.

Example data includes gender, document registration numbers, expiry date, date of birth and MRZ.



Algorithm validations. Some IDs include very specific embedded rules, in the form of mathematical algorithms, to identify any flaws in ID integrity.

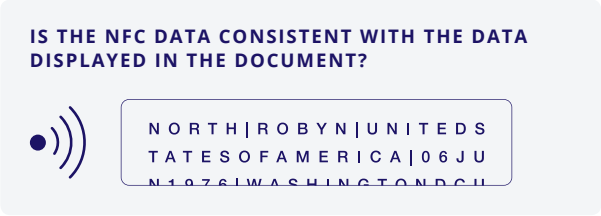


Data consistency. Certain data fields that are present in multiple locations on the ID must match. Certain data is encoded in machine readable formats, and when decoded will map directly to fields presented in clear text.

Fraudsters can miss these and put different values in each location, including the back of the ID.

Examples include date of expiry, document type, nationality, issuing country, document numbers, gender, date of birth, first or last name.

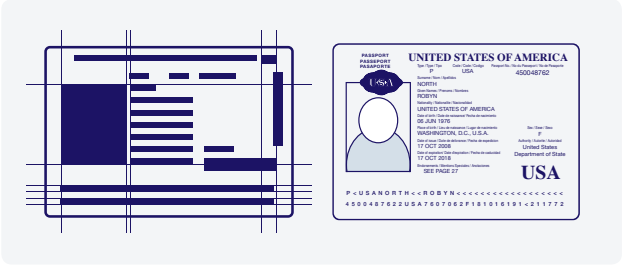
NFC chips. On many IDs, the NFC chip contains additional information. Appropriate readers can access this information and compare it to information displayed on the ID.



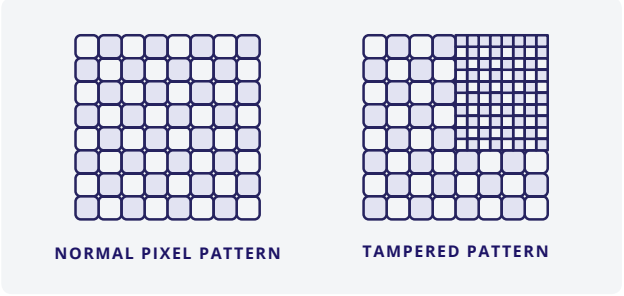
Visual document authenticity

This white paper focuses on online processes. It therefore assumes that the ID being evaluated is a digital image, and the user will present themselves via a digital image capture—either a photo or a video. Digital analysis is the best way to evaluate the authenticity of these images, because it can spot the forgeries that might be more prevalent when a human inspector is not present. There are a number of ways to evaluate the ID and user, which can help identify possible tampering and impersonation from multiple perspectives:

Document template comparison. Comparing the submitted ID against the known document template can identify errors or inconsistencies.



Digital tampering. Fraudsters can use software programs to modify a digital image of an ID. This might mean replacing or modifying the photo of the user. They then submit the tampered copy, with the falsified picture. Onfido's technology analyzes residual pixel information and image metadata to detect when pixels have been copied or modified—a key signifier of tampering.



Font anomalies. Fraudsters will often try to change fields of data but will leave behind font inconsistencies while doing so.

Document copies. Fraudsters will also try to upload a copy of the ID instead of submitting the original ID—often a modified version.

Security features. IDs all have some form of built-in security features. Evaluating these can ensure authenticity.

Examples include:

- Digital watermarks
- Barcodes
- Embossed text

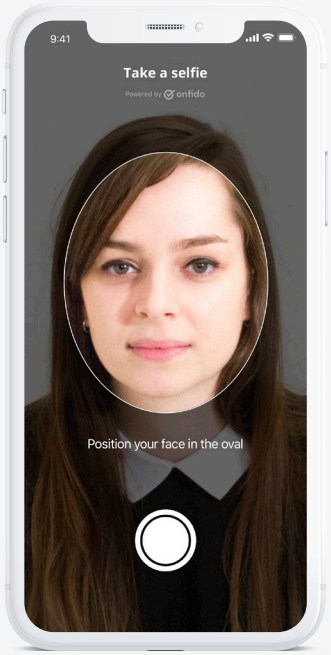
These vary by ID type and version.

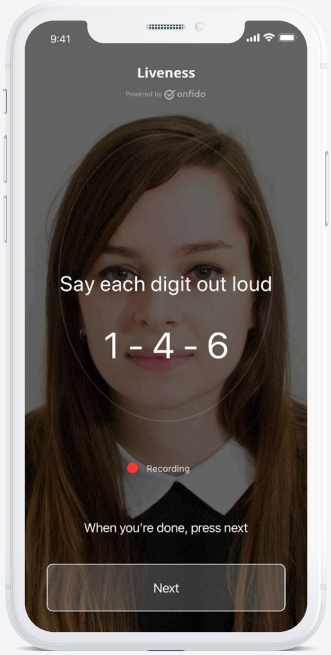
Proof of ownership: facial comparison

As well as confirming that the ID is authentic, we need to confirm that the user presenting the ID is its authorized owner. To prove ownership, we can compare a live capture of the user with the ID photo to confirm facial similarity.

There are two different capture experiences which each provide facial images in near real-time—selfie photos or video liveness.

Fraudulent users will attempt to spoof the system with a printout of the target's face, or photos from their Linkedin profile, to sophisticated 3D masks. Online verification flows, along with mathematical comparison techniques, can protect against these attack vectors. We can minimize the risk of photos from other sources being used by ensuring that the image or video submitted was created on the user's device, and detecting the time the photo or video was created.





Selfie photos protect against common spoofing attempts, like providing a photo of a photo, by using passive anti-spoofing algorithms.

Video liveness provides the highest level of assurance against sophisticated spoofing attempts, including deep fakes. Instead of taking a selfie, the user films themselves reading out numbers and performing randomized movements. This proves liveness in a "challenge and response" interaction.

A machine learning approach to document fraud detection

But how do we detect fraud at scale in a digital context? This could be a fraudulent ID, an impostor, or both.

We have to confirm that IDs are authentic and belong to the user. This means inspection by either (1) trained human experts or (2) machines built to programmatically assess an ID and determine its risk of fraud. Manual inspection alone—while relatively accurate if performed by an expert—doesn't scale well for high volumes for online account onboarding and transactions.

Enter the machines. When it comes to machines, there are two different approaches. The first is a classic heuristic model. This uses pre-defined techniques and applies them directly via static software. The second is a machine learning approach that uses data to train the model. It continually learns as it processes more data, improving its performance.

Heuristic models are limited in the degree of fraud sophistication they will catch, as illustrated below:

***Machine* from this point forward will refer to machine learning systems.**

Building an optimal solution, whether human or machine-based, involves trade-offs:

	PROS	CONS
 **Human**	• Starts processing IDs quickly with basic training • Makes more complex judgment and experience decisions • Generates data to train machine learning models • Double checks machines to reduce false positives • Can interpret contextual signals	• Built-in limitation on how fast they can complete processing • Fatigue degrades performance over time • Cost will always be more than a machine, especially at high scale • Can't effectively tell if two unfamiliar faces are a match
 **Machine**	• Never tires • Will always get faster • Manages huge sets of IDs and rules needed for global scale • Can spot certain fakes that humans can't • More precise and quantitative	• Needs to be trained • Needs reaction period for new docs and fraud techniques • Not always clear when to escalate to human judgement

Regarding facial comparisons, many studies show that it's difficult for humans to identify whether two unfamiliar faces are a match[4], and that face matching algorithms often outperform humans[5].

We can draw a few conclusions from this:

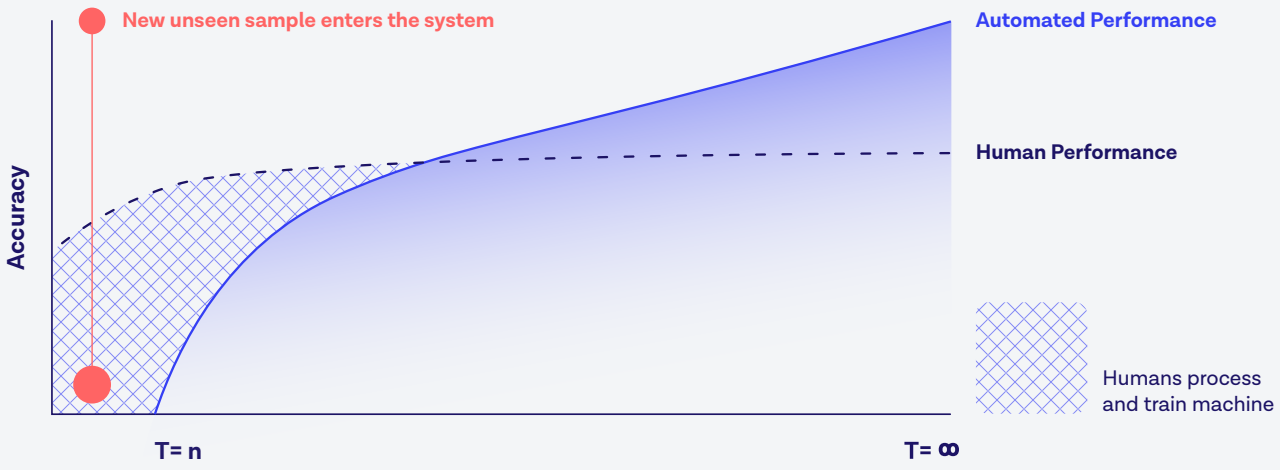
1. A solution manned by humans alone will:

- Not catch all fakes, [6] and will miss a significant number at higher scale and sophistication
- Have limitations in scale, both in terms of cost and processing time
- Make mistakes both positively and negatively detecting fraud

2. A solution manned only by machines will:

- Catch a higher degree of possible fakes, specifically around new sophisticated attacks
- Require human intervention to identify new fraud techniques and train machine learning models with significant data sets
- Require significant software development expertise and machine learning infrastructure to build and maintain models

In reviewing (1) and (2), it is clear that a hybrid approach, using both humans and machines, is the ideal way to solve the problem. This concept is illustrated below:



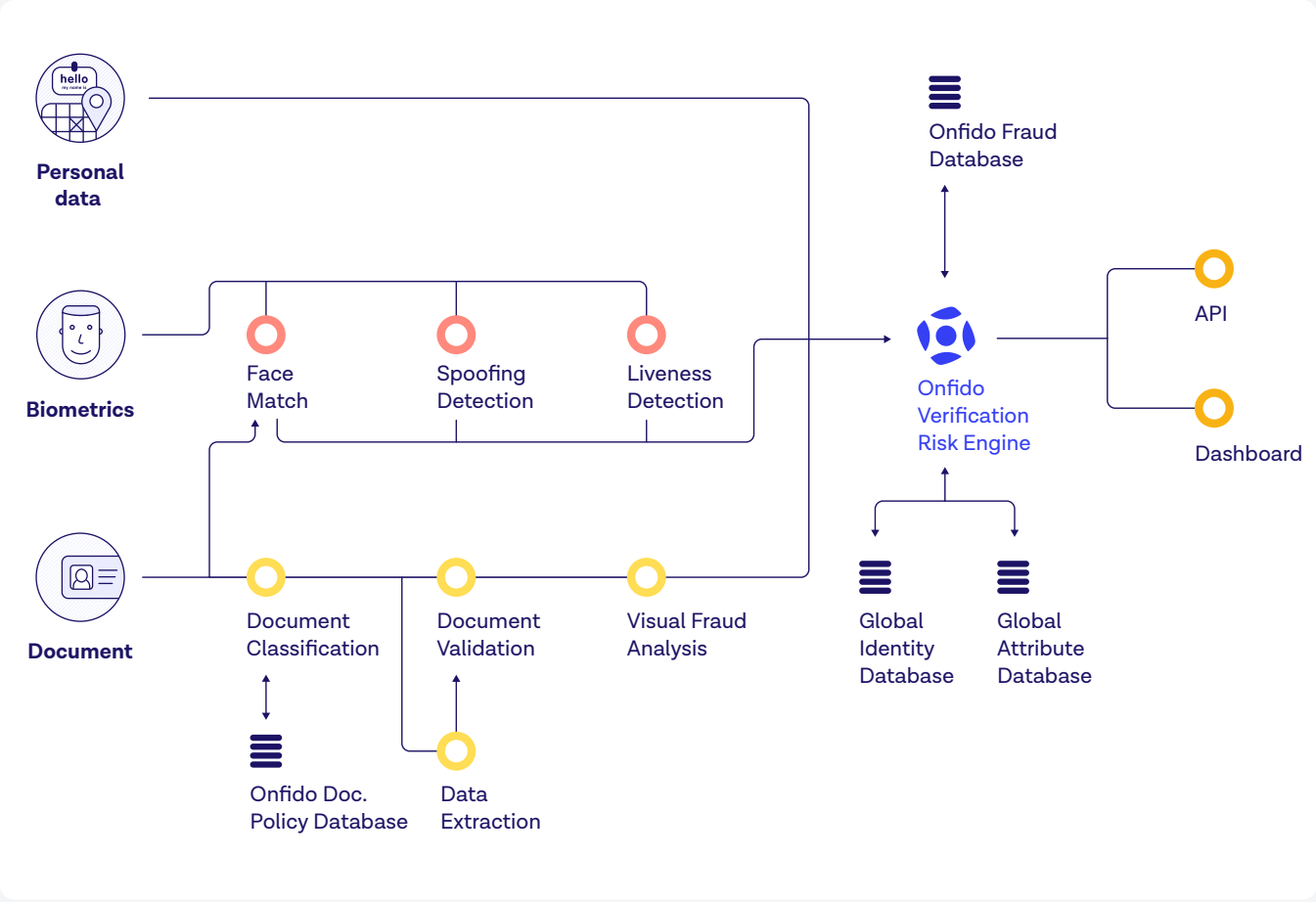
Document verification process

To verify an identity using a government-issued ID, the process looks like this:

Document and biometric capture	• Digitally scan the ID, ensuring proper image quality for analysis • Collect an authentic photo or video of the applicant • Collect any device or environment fraud signals and fingerprints
Document classification	• Document type • Document issuing authority • Document version
Extraction and validation	• Read and capture the relevant PII and document characteristics from the text and MRZ areas • Validate all document template format, data and security rules
Document authenticity analytics	• Analyze the digital document characteristics to ensure it has not been tampered with or altered • Confirm it is not a previously known or publicly available fake • Verify the document issuer status with the issuing authority (where possible)
Proof of ownership: facial comparison	• Compare the captured image of the user to the photo in the ID to confirm facial similarity and compute risk confidence • Add further human liveness testing to defeat spoofing
Supplementary signals	• Verify the information on the ID with a trusted identity database to prove the identity exists • Assess device network information to ensure it doesn't possess a known fraud profile

Onfido product platform

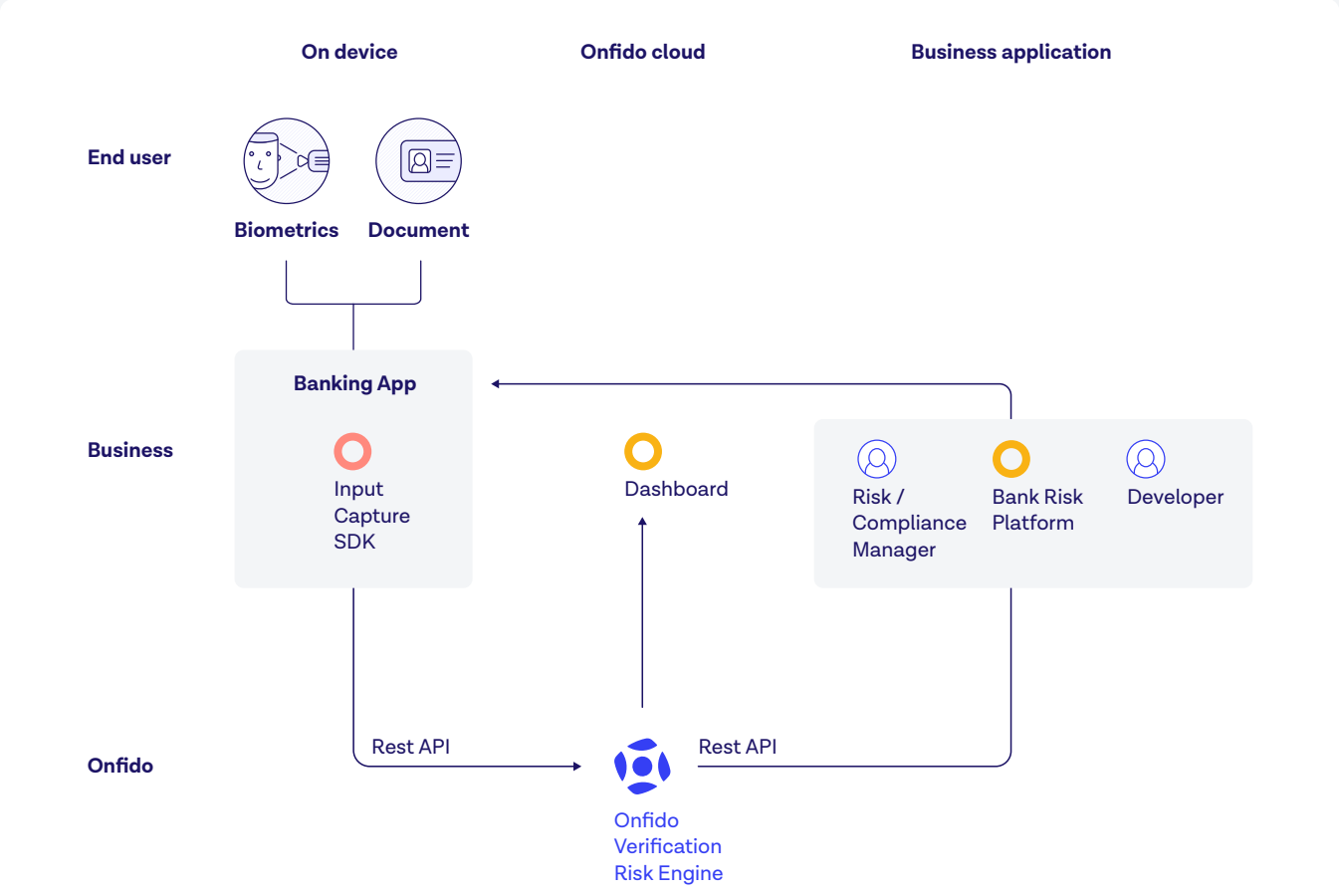
Onfido's identity verification (IDV) platform is a risk engine that combines various signals to assess whether a user is in fact the person they claim to be. The platform uses global identity documents as a primary source of identity, and compares them to a digital capture of the user in either a photo or a video. Supplementary databases can provide additional verification or augmentation of identity attributes.



The Onfido IDV platform is cloud-based, and is easy to integrate into common business applications and back-office risk systems via a flexible set of SDKs and REST APIs.

The cloud platform brings the following benefits:

1. It scales globally to verify anyone, anywhere in the world
2. It's low cost to operate, which brings better value to its users than expensive and hard-to maintain on-premise software..
3. The system is continually evolving with new features, upgraded technology and additional document coverage with little to no client-side integration needs. This keeps system users fully up-to-date and able to address ever-evolving fraudsters.



Onfido's platform is also built to meet global regulations. It delivers both high performance and compliance with complex global data privacy and sovereignty restrictions. And when users request it, Onfido can delete their user records within seconds, in accordance with appropriate data retention policies.

Supervised machine learning

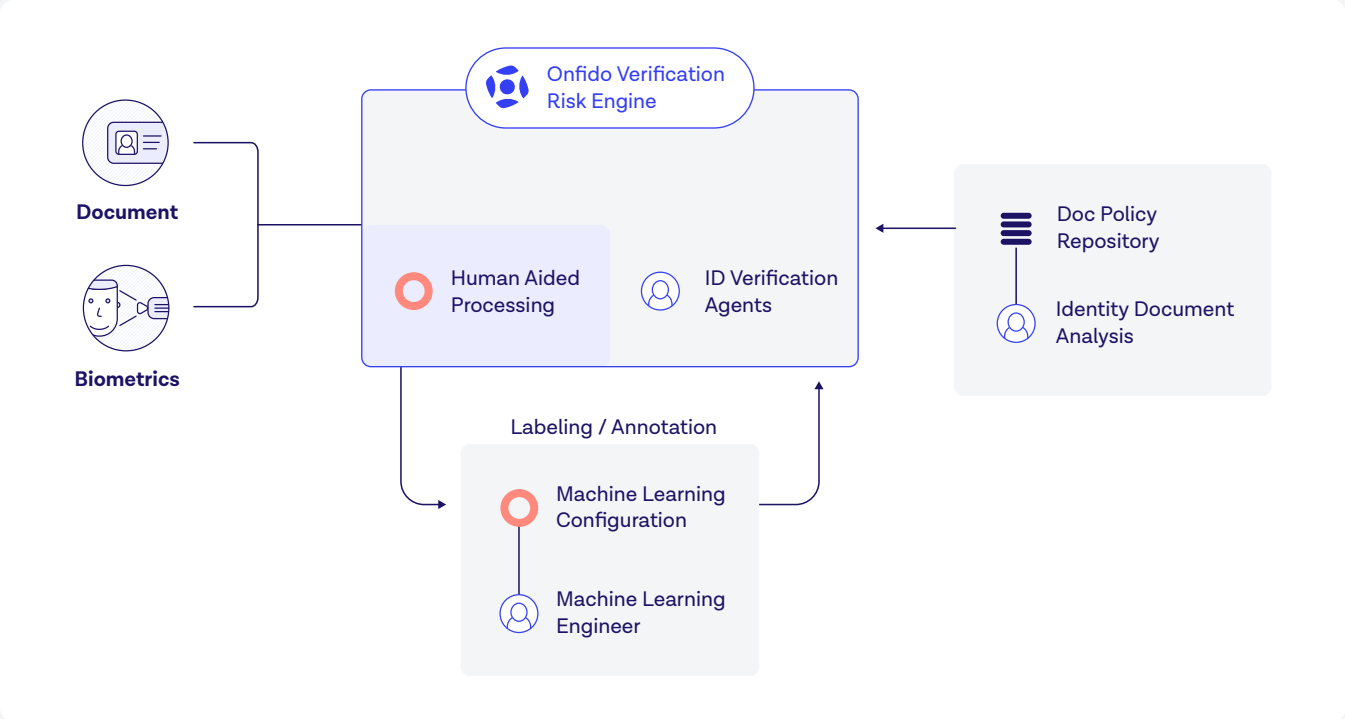
It's universally recognized in today's enterprise software market that to meets the needs of large global businesses at scale, machine learning is the way. Global companies pose a series of complex challenges for identity verification in particular.

The difficulty of processing a large, dynamic set of global government-issued IDs (and their associated versions), for one. To classify those IDs, extracting their data and evaluate their authenticity isn't easy. The system needs to perform effectively, consistently and quickly, according to the business process it serves. Onfido believes that a machine learning-based approach is the only viable way to achieve these results.

The Onfido IDV platform was designed and built using deep learning, a machine learning technique designed to mimic the way the human brain works (layers of neurons in the neocortex) for the broadest, most effective performance.

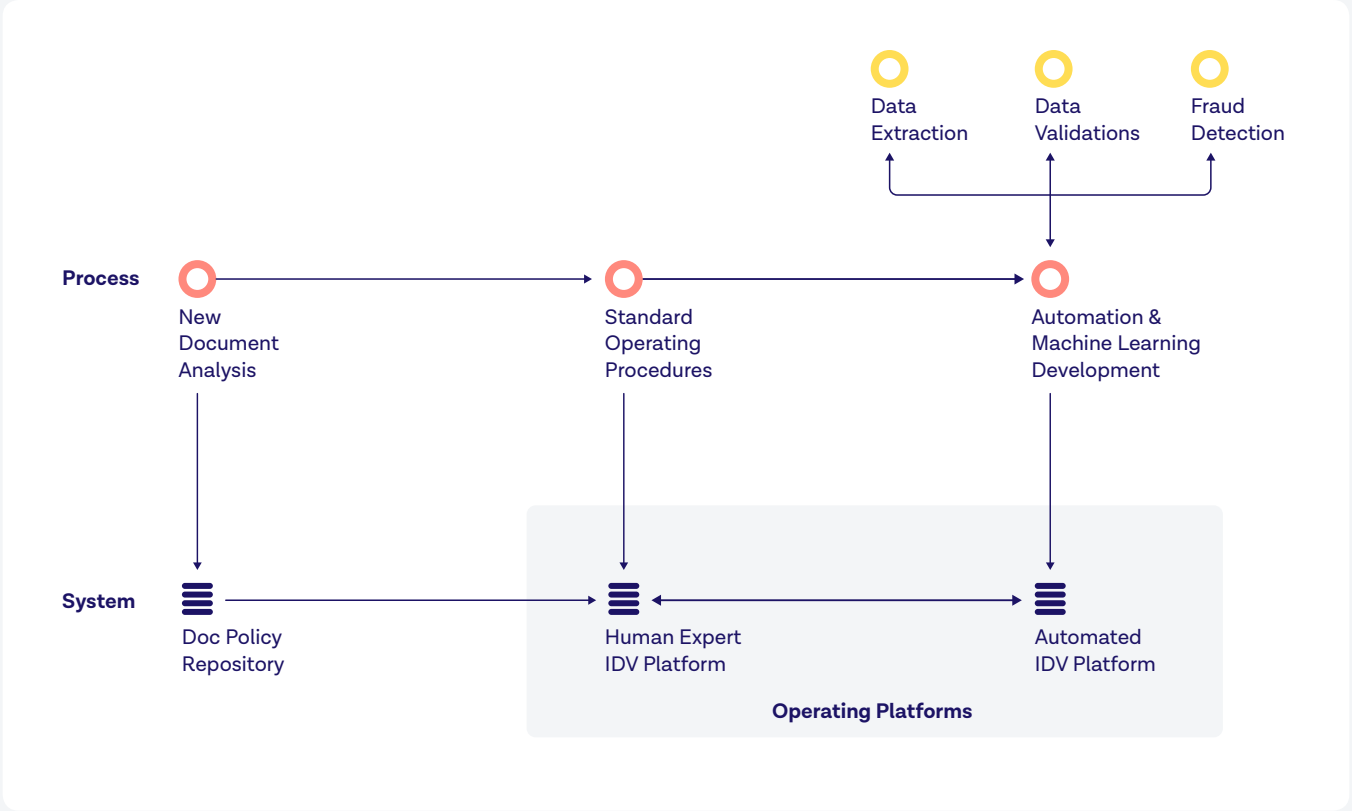
Onfido uses trained human experts to process new IDs not seen previously by the system. These are then classified and labeled to help train the deep learning models. Once the models are tuned to acceptable business performance, the human experts look at other, newer IDs, or manage quality control and advanced forensics (including the detection of more sophisticated frauds and continual system performance).

Similarly, Onfido's facial verification uses a combination of machine learning algorithms and human agents. This gives a better result than a human or a machine agent alone could achieve. [7]



Document repository

Onfido's Document Specialist team are the first to analyze an ID. This team are world-class experts from both the public and private sectors with decades of experience in document fraud. They encode the templates into Onfido's document policy repository, known as DocuPedia. This system is then used to drive the definition of the initial manual processes, and subsequent automated processes. This is illustrated below:



Biometric technology for facial verification and anti-spoofing

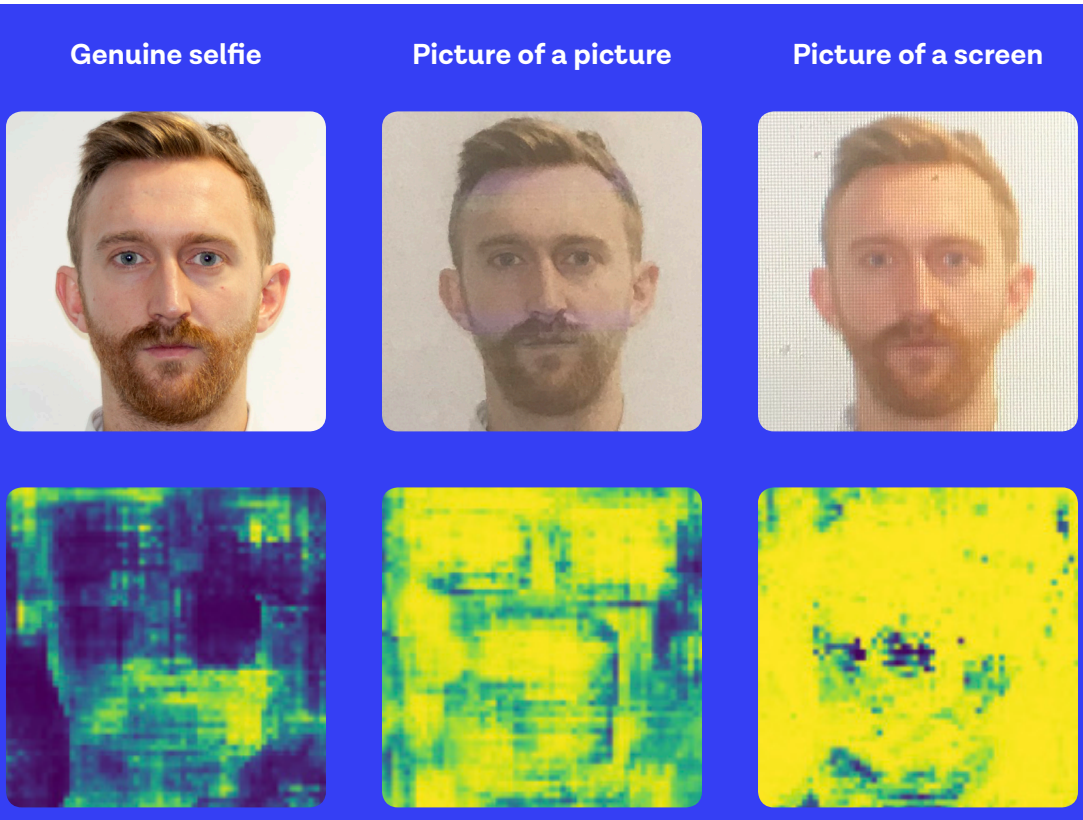
We need a complex set of processes to analyze a wide range of users. We need to determine that those users are real human beings. We need to know that they're not being impersonated. And we have to confirm that they match the image on their ID. Below is a summary of some of the processes that Onfido uses:

Several of these processes require machine learning models. For spoofing detection on selfie photos, the algorithm goes through the image and extracts patches. It analyzes each patch for evidence of it being a printout of a photo or a digital screen.

Here the visualization shows yellow for each patch that has a spoof-like texture. The more yellow there is in the image, the more likely it is to be a spoof.

Some of the processes that Onfido uses

- **Voice (audio processing)**
  Voice to text processing to verify digits.
- **Headpose (face tracking)**
  Tracking face movement to verify movements.
- **Lip (mouth tracking)**
  Tracking lip movement to make sure it matches the voice.
- **Sherlock (texture analysis)**
  Texture based analysis finds photos of photos and photos of screens.
- **Agents (human review)**
  Humans uplift when the machine can't reach a conclusion.



This can also be applied to video images, as illustrated below. Spoofs are represented via the dense red areas, which indicate if it is a video of a printed out screen, for instance.

Finally, evaluating human 'liveness'. Assessing whether the face shown is 3D—by analyzing it as it turns to one side—can flag suspicious movements, which may indicate masks or other spoofing techniques.

The next section will look deeper into Onfido's technology, and how Onfido uses machine learning.

Video analysis

Genuine video

Spoofed video

Head turn 'liveness'

Accept

Reject

Onfido machine learning technology

Identity verification challenges

All of the identity verification processes outlined above require judgment calls. For example, whether two faces are the same identity, or whether the font used in the MRZ is correct. Machine learning powers those judgement calls by taking a digital image as an input, and outputting a decision to a degree of confidence.

Images add a layer of complexity to this process. For that reason, Onfido's technology also relies on elements of computer vision. Images are challenging because they're high dimensional representations, with a low signal-to-noise ratio. This means that of all the pixels in an image, only a small proportion will directly influence the decision, as often the object of interest will only make up a small proportion of the image. This effect is amplified when it comes to building a fraud detection system, as we're often looking for small inconsistencies. And since users capture images in an unconstrained environment, the system has to be trained to handle photos of varying quality and distortions—just like the human brain.

Onfido's systems capitalize on the latest in deep learning techniques. Such methods have the advantage over more traditional pre-2011 computer vision pipelines. One key difference is that in a deep learning system we're able to learn the image features rather than making use of hand-engineered ones. Doing so gives the model more expressive power and allows us to model more complex functions. These kinds of models make it possible to leverage massive amounts of data, both external and internal. And more data means better results.

Onfido's statistical models help make fraud prevention decisions quickly and at scale. Our technology detects fraud by looking for anomalies in the data it receives. In an anomaly detection problem, the amount of 'normal' data will outweigh the amount of anomalous data. Conversely, most datasets in academia contain balanced datasets, meaning that all classes are well represented.

A typical dataset from an online business that needs identity verification will only contain a few examples of fraud—typically less than 1%. This influences how the decision process is built. If we can only model the normal class, decision-making must be based on how far away a new fraudulent submission is from the norm represented by the model.

Humans and machines

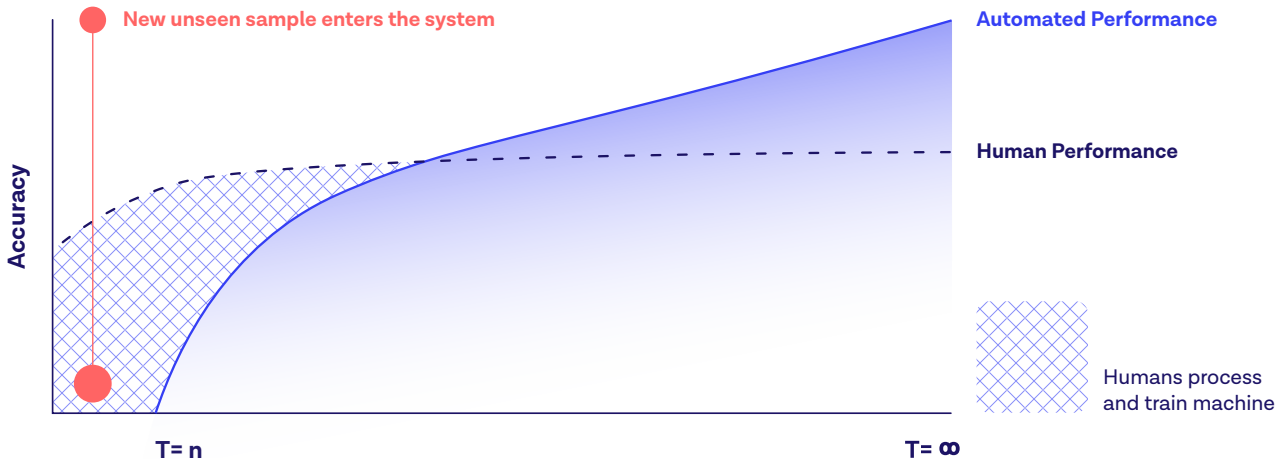
Many machine learning systems are ‘human-in-the-loop’ models, such as supervised learning.

In supervised learning, a label y is given for each input x. This information then trains the model by making use of the predicted value ŷ and the expected value y. This label often comes from human annotators—either a single annotator, a well trained specialist, or a group of specialists. In any case, human knowledge is the gold standard, the source of truth.

Onfido has built on this idea to develop a ‘hybrid’ machine learning system. In this system, human knowledge—the ability to generalize and the capacity to learn new tasks quickly—is combined with automatic decision making. The goal is to provide the highest level of accuracy, while maximising completion rates.

Onfido’s facial comparison and liveness systems use proprietary facial matching technology alongside expert human review. This enables the automation of a large percentage of verifications, while examples of high uncertainty are escalated to human review. This combination delivers a high completion rate.

Onfido is invested in creating processes that allow human decision making to work symbiotically with machine learning. Tightening the loop from prediction to training powers systems that constantly evolve and improve.



Areas of focus

Onfido's focus to date has been on two problem domains—facial biometrics and identity documents. Creating proprietary models trained on owned data has improved system performance considerably on both.

Document data extraction is one example. Optical character recognition (OCR) is already very commoditized, with excellent solutions from the likes of Abbyy, Google Vision and Microsoft Cognitive Services. By developing a proprietary OCR system specifically optimized for matching images on IDs, we've been able to boost performance by 30-70%.

Similarly, Onfido's facial recognition system has been trained to boost performance by over 30% when compared to state-of-the-art general purpose solutions.

Recasting as an anomaly detection problem means that we can model the true class, and use the fraudulent examples to tune parameters. This means that far more data can be leveraged from the true class. As a specific example, when we applied this approach to MRZ font verification, it halved the number of false positives generated by the model.

	SECURITY	SCALE	SPEED
Machine learning	ML can detect fakes that humans can't	ML learns to recognise global IDs	Sub 15 second results
Human	Humans QC to eliminate false positives	Human experts provide training data for ML models	Human labelling for long-tail IDs with 2min average results

The future of identity

Advances in smartphone ownership, machine learning, privacy awareness and cloud architecture have made identity verification technology available to everyone, not just higher risk financial institutions. And the move away from vulnerable, centralized PII databases calls for a new model—both in terms of policy and technology.

Identity verification protects consumers by guarding their identity against fraudsters. We all hate it when our accounts are hacked, so consumers want robust identity proofing as much as businesses do.

But for a system to be ubiquitous, it needs not only to be trusted, but convenient—so that it doesn't get shunned by users. The ability to port, or transfer, identity between providers safely and securely will be key to this. It removes friction for consumers, and reduces cost for businesses. At Onfido, we're investing in this future. In Europe we lead the 'Fintech Delivery Panel'[8] to pilot a scheme for users to securely re-use their identity across different financial services. In the US, we work with the Better Identity Coalition as they move towards a similar goal, as well as tackling the problem of using the Social Security Number as a central identifier. Our vision is to build an open world where identity is the key to access.

By bridging offline identities (government issued IDs) with a new, digital credential created upon verification of an identity claim, some exciting new possibilities can open up for the next chapter of identity services.

Decentralized identity. Today, no one wants to enable any single company to house and manage identity information for all users globally and digitally. These systems are impossible to keep safe, and once compromised, bring down the trust of the whole system. A better system would be owned by the end user, and housed and stored in a distributed manner. There would be no single owner, unless the user themselves decided there should be one. With blockchain and distributed ledgers, the technological foundation is in place already. Once the legal structures for relying parties can be put in place, this system could become a reality.

Standardized language for verified claims. There are many types of verifiers from businesses, services providers and government agencies. Without a common language, these various systems cannot interoperate and gain efficiencies and scale. The W3C[9] is driving crucial work around creating a common language for verified claims. This would provide that important set of relationships, and enable a new way to share and consume various identity claims about a user.

Thanks to these and other innovations, digital identity is a rapidly developing area. And Onfido—with ever-evolving technology, human expertise and a partner system that includes Microsoft and Salesforce—is helping to drive this change.

onfido

Endnotes

1 https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

2 https://www.itgovernanceusa.com/blog/us-dominates-the-world-in-data-breaches/

3 https://www.betteridentity.org/news/2018/2/23/press-release

4 Megreya, A. M. & Burton, A. M. Matching faces to photographs: Poor performance in eyewitness memory (without the memory). Journal of Experimental Psychology: Applied 14, 364–372 (2008)

Kemp, R., Towell, N. & Pike, G. When Seeing should not be Believing: Photographs, Credit Cards and Fraud. Applied Cognitive Psychology 11, 211–222 (1997)

Burton, A. M., Wilson, S., Cowan, M. & Bruce, V. Face Recognition in Poor-Quality Video: Evidence From Security Surveillance. Psychological Science 10, 243–248 (1999).

White, D., Kemp, R. I., Jenkins, R., Matheson, M. & Burton, A. M. Passport Officers' Errors in Face Matching. PLoS ONE 9, (2014).

5 O'Toole, A. J. et al. Face Recognition Algorithms Surpass Humans Matching Faces Over Changes in Illumination. IEEE Transactions on Pattern Analysis and Machine Intelligence 29, 1642–1646 (2007).

Phillips, P. J. & O'Toole, A. J. Comparison of human and computer performance across face recognition experiments. Image and Vision Computing 32, 74–85 (2014).

Lu, C. C. & Tang, X. Surpassing Human-Level Face Verification Performance on LFW with GaussianFace. Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI-15), Oral Presentation (2014).

6 http://www.dailymail.co.uk/sciencetech/article-2728172/Passport-officers-no-better-public-spotting-fakes-Officials-fail-recognise-false-documents-15-cent-time.html

7 Scheirer, W. J. et al. Report on the BTAS 2016 Video Person Recognition Evaluation. 2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS) (2016). doi:10.1109/btas.2016.7791198

8 https://technation.io/about-us/fintech-delivery-panel/

9 https://www.w3.org/2017/vc/WG/

onfido

Get in touch:

www.onfido.com
info@onfido.com

London | **San Francisco** | New York | **Lisbon** | New Delhi | **Singapore**