



What You Need to Know About

Digital Identity Verification

What You Need to Know About Digital Identity Verification

Contents

- 01 Overview**
- 02 History**
- 05 Verifying an Identity**
- 08 Methods for Verifying an Identity**
 - Approaches to Verifying an Identity
- 12 Process for Verifying an Identity: Risk Modeling**
- 13 Evaluation of ID Authenticity**
 - ID Forgeries and Identity Fraud
 - Degrees of Fraud Sophistication
- 16 Approaches to Detecting ID Fraud**
 - Data Integrity Analytics
 - Visual Document Authenticity
 - Facial Comparison as Evidence of Ownership
- 20 An Approach to Detecting Document Fraud with Machine Learning**
- 23 The Process of Document Verification**
- 24 FTx Identity Product Platform**
- 27 The Machine Learning Technology of FTx Identity**
 - The Challenges of Verifying an Identity
 - Focus Areas
- 29 The Future of Identity is Here**

Overview

In this day and age, identity risk management is more crucial than ever. As traditional businesses are fully digitalizing their operations, newer all-digital businesses are expanding. Remote user onboarding is essential for both groups' services. However, as their user base increases, they also need to keep an eye on their platform's risks.

The traditional threats of service misuse, impersonation of another person, and money laundering are heightened in a fully digital society. The traditional threats of service misuse, impersonation of another person, and money laundering are heightened in a fully digital society. Businesses seek to provide easy, quick access. However, they also aim to prevent fake users from using their site. The same trade-off is causing difficulty for everyone.

Naturally, users care about protecting their identification information as well, and this is understandable given the approach to matters has been disorganized and outdated. After all, identity information shouldn't be as inadequately stored or susceptible to duplication. Data breaches need to be the occasional exception rather than the rule. What does this mean, then? The secret to managing risk is confirming a person's true identity.

We've seen a few different methods in the past for confirming an online user's identity. All of them lack perfection. Most of them are unsuitable for the job. A careful balance exists between geographical reach, user experience, and fraud prevention. A method that benefits both businesses and users is the answer. Allow users to prove their identity using the most widely used forms of identification. By combining users' legal identities with biometric analysis, companies can be sure that users are who they say they are.

It's hard to do this on a large scale across the world. Scalable accuracy cannot be achieved with a human-only technique. We can finally address the problem of online identification since artificial intelligence (AI) and machine learning (ML) approaches have finally reached maturity. Now, any person in the globe having an identity document (ID) and an internet connection can be confidently onboarded by a business.



History

In this day and age, people are always on the go. They prefer an easy-to use interface that provides them with all their digital needs in one place without any hassle or confusion. What was previously referred to as eCommerce is simply known as shopping. Additionally, what was known as online banking is just referred to as banking. Mobile has taken on new significance in the age of the iPhone. As a result, businesses communicate with their customers wherever they are, so that means that there is no longer the need for face-to-face interaction.

Peer-to-peer (P2P) payments, international transfers, check cashing services, and loan origination are all now performed more easily and affordably through a customer's phone. Imagine all of that convenience in the palm of your hand! Traditional brick-and-mortar financial institutions have mostly adopted the mobile experience, despite what may appear to be a slow adoption rate to outsiders, which has made them reevaluate their transactional workflows. For example, a smartphone in a user's hand can function as a document scanner. Hundreds of millions of customers can now deposit checks without having to go to a branch or ATM.

Entrepreneurs are starting new enterprises outside of the financial sector in a matter of hours rather than days or weeks. With just one click, you can get low-cost cloud storage, website design and development, payment processing, and even payroll services. Due to the growth of the sharing economy, any-

one with idle assets, particularly residences and vehicles, now has the chance to make money by renting them out to others. For instance, a stranger could rent out their townhouse in Nantucket to you for the weekend. Long-standing for-hire services like taxis and car rentals have also been impacted by the economy's digitalization. Additionally, it has paved the way for new business models such as food delivery and dog sitting businesses.



More online services equate to more customer data, and when businesses have more data to manage, there is a lot of data misuse and mismanagement. Customer trust has been harmed by inadequate data protection, which has resulted in unprecedented regulation. Another significant data breach is reported every other day. With so much information being released, it's shockingly simple for scammers to pass as another person, use a stolen credit card, or tap into an account.

History Cont'd

Both the public and private sectors are looking for preventative measures to protect their personal data. These safeguards wouldn't be dependent on a single central database run by a private entity or a central database. Thanks to the efforts of various industry working groups, this understanding has gained pace over the last few years.



Industry and governmental officials have offered some fundamental guidelines for data usage and protection in an effort to win back customers' trust. Even sovereign states have had to take precautions to make sure that respectable online companies aren't unintentionally aiding in the transfer of funds for terrorist funding or other illegal operations. As a result, organizations now have to adhere to a number of compliance standards in order to safeguard against fraud, money laundering, and terrorism.

A number of fundamental banking regulations in the United States and Europe have compelled financial services organizations to implement compliance procedures. These include the Bank Secrecy Act of 1960 and the 2001 Patriot Act, which was passed in response to the September 11 attacks after it was discovered that terrorists had created a sizable economic network. In order to ensure that banking transactions are monitored and not anonymous, banks must also ensure that a relevant level of customer due diligence (CDD), often known as KYC or know your customer programs, is in place. According to KYC guidelines, customers of an institution must have their identities verified and their activities must be understood.

The Department of the Treasury's US Financial Crimes Enforcement Network revised the Bank Secrecy Act's current rules for several financial firms. Although there are several exceptions and exclusions, the CDD rule specifies clear customer due diligence obligations and imposes a new requirement on financial institutions to identify and verify the identities of beneficial owners of legal entity customers. Of course, there is also the well-known GDPR. A new and stringent set of data privacy regulations were established in Europe in May 2018 with the implementation of the Europe's new General Data Protection Regulation. Global businesses must abide by GDPR requirements if they collect information from European citizens, or they risk paying hefty fines. The GDPR mandates that businesses have data protection

History Cont'd

systems, including policies and procedures for handling and safeguarding data.

To protect online shoppers, regulations will keep adding restrictions. However, a data breach cannot be reversed. Personally identifiable information (PII) can no longer serve as the foundation of identity when it comes to online transactions because so much data has been stolen. It is impossible to disregard the harm done to customer trust. It currently poses a considerable barrier to the general use of online services. The long-term viability of the new digital use cases discussed above requires platform confidence. However, there will be a price for the new regulations. They'll make things more challenging for users. Additionally, a lot of online business strategies depend on brand-new users joining right away and using the service frequently. Due to this reason, it is important to carefully weigh these precautions against user experience.

How can I tell whether someone is who they say they are?—Businesses have struggled with this issue ever since the dawn of the internet era, and it is now becoming more serious. How can a financial services company maintain regulatory compliance around KYC when all communications and transactions with users happen entirely online? To get around this problem, businesses try to confirm the identity information that users provide. If the company thinks the information provided by the user is accurate, the user

may create accounts and make online transactions. The user is identified through this process, which also ensures compliance for the business. It is challenging since the sign-up procedure must not be slowed down. Speed and security must be compromised.

There are essentially two approaches to put these safeguards in place:

1. A very thorough first screening to ensure that users are known individuals and have no criminal record or potential to commit financial or service crimes. Identity verification is the term for this initial screening, which is the main topic of this white paper.
2. A less thorough screening process that results in a lower level of service access but is combined with a higher level of user behavior monitoring once they are within the system. They may be given higher degrees of service access in the future, provided that no new risk criteria are set.

Recent innovations in technology, such as AI, let companies learn crucial information about their customers.

The following sections will examine several methods used by businesses to validate customer identity information online.

Verifying an Identity

Identity is a collection of claims that can be used to characterize an individual. Permanent characteristics like the date of birth, ethnicity, and fingerprints can be mentioned in these claims. Or, they could consist of semi-permanent characteristics like name, height, weight, and eye color. Over a person's lifetime, these qualities either don't change much or are highly challenging to alter. Identity claims may also relate to contact and housing details, such as home addresses, phone numbers, and email addresses, which don't typically change but could. It is evident from these features that the trait itself is significant and not just its most recent version. Equally important is the trait's history.

So, what is identity verification? It is the process of validating these distinguishing characteristics and comparing them to those of an actual human. A digital credential that allows access to internet accounts, services, and activities can then be given to that person. When someone claims their identity with a relying party for the first time, their claim needs to be verified to ensure they are legitimate.

Generally, this use case is referred to as account onboarding. It acts as one of the primary driving forces behind verifying an identity. This process is well-known to us because it is essential for a variety of activities, such as opening a new bank account or flying. The concept of identity evolves over time for relying parties as the nature of interactions with the user changes. Having said that, verifying an identity can also be used to allow access to new services, offer stepped-up or tier-based services with increased rights, or reevaluate the user if issues with their identity come into place later on.



Verifying an Identity Cont'd



When confirming someone's identity, you should take three things into account:

- **What is Known** (e.g., out-of-wallet questions that can be stored in a database, like your monthly mortgage payment)
- **What We Possess** (e.g., a token, credential, or trusted ID)
- **Our Identity** (examples of biometrics include voice, fingerprints, and faces)

We'd need to employ a multifactor strategy to provide a trustworthy result. Without other elements, some parts are useless; for example, a fingerprint is useless unless it's connected to a specific name or ID.

Claims made using only personal information ("what we know") lack credibility given the billions of personal records exposed in recent data breaches. They won't provide evidence to support a person's identity. Instead, in order to confirm that the individual claiming the identity genuinely owns it, we need to connect the claims to a reliable source.

An authoritative registration system, which is frequently controlled by the government or the individual using it, is an example of a reliable source. A birth record, for instance, is a legally binding document that is produced at the time of a hospital birth. As the newborn grows into a child, this information is transferred into a special government or citizen ID and then into the health and educational systems. The three primary types of photo ID—driver's licenses, passports, and national identification cards—are obtained when the person gets older. They offer a trustworthy document that adds to the certainty surrounding that person's identity.

Verifying an Identity Cont'd

A valid ID by itself can help confirm identification. However, adding human verification can produce a more reliable outcome. In the real world, banks typically do this by requiring a user to physically visit a branch and speak with a bank representative in order to demonstrate ownership and possession of the ID. However, no one wants to do that any longer given the convenience of the digital age.

By logging into an app on a smartphone, we can recreate the advantages of a branch visit model without the hassles. Anywhere users can upload their ID and a photo of their face. The owner of the ID is then confirmed to be the individual attempting to open the digital account by mathematical comparisons between the facial image and the ID picture.

We can increase the level of identification assurance with video. A user can upload a video along with the ID to combine face recognition with “liveness,” which serves as evidence that the person wearing the ID is a real, live person. In an online setting, the user replies in real-time to random tasks like “turn your head to the right” or “nod twice.” To verify that the person in the video is actually real, it is possible to analyze it in close to real-time alongside the ID image and compare the extracted face to the face on the ID. A company can generate an online digital identity once it has confirmed a user’s identification. After that, the platform can offer a stronger credential, such as a distinct account number or username/password combination, allowing access to your account.



Methods for Verifying an Identity



What is the most effective way to verify someone's identity? It's crucial to consider:

- **Trustworthiness:** How susceptible is the strategy to past attacks that might compromise the accuracy of the source data? For instance, a database used to test out of-wallet questions should be examined for any prior data breaches that would indicate that the data is no longer private.
- **Ubiquity:** How broadly is the approach used for the population whose existence we're trying to confirm? While many businesses are global, many identity management techniques are local. This will fail.
- **Cost:** How much does it cost to protect and preserve data and credentials? Processing costs associated with making sure they can be trusted will be taken into account.
- **User Experience:** How user-friendly is it? The identity flow should be quick and seamless. If there is too much difficulty during the sign-up process after paying a lot of money to acquire a customer, the user will "drop off."

Methods for Verifying an Identity Cont'd

Approaches to Verifying an Identity

Here are a few of the more approaches to confirming an identity:

Databases

These are data storage systems for data that has already been gathered and verified as part of a registration system. They may be public databases managed by governments or private databases controlled by for-profit businesses. Telephone directories and credit bureaus are two examples of private databases. Government identifiers (such as social security, tax, or voter numbers) or the DMV, which holds information and numbers related to driver's licenses, are examples of public databases. When using databases for verification, it's crucial to take into account the cost of access, the likelihood that past data breaches have weakened the data's validity, and if the data can be used for commercial purposes in accordance with the law.

Government-Issued IDs

These consist of:

1. **Driver's License**
2. **National Identification Card**
3. **Passport**
4. **Residence Permit**
5. **Tax Identification Document**
6. **Voter Identification Document**



It's crucial to keep in mind that forgery is widespread when utilizing government-issued IDs for identification verification. The key is identifying user misrepresentation fraud and other forms of fraud. Additionally, there isn't a single worldwide schema for a standard ID format. Therefore, in order to process these documents, especially on a global scale, we need extensive experience. Last but not least, since papers are physical artifacts, changing them digitally necessitates some kind of scanning and analytics.

Methods for Verifying an Identity Cont'd

Phones and Phone Numbers

Our lives now extend to our phones in many ways. They serve as an extension of a person's identity that we can trust to some extent. This is especially true considering how many of us maintain our phone numbers while switching out phones. An adequate amount of fingerprint data can be provided by the phone itself. When the user still has possession of the phone, this might be used to identify them.

Although it's rather widespread in many fraud risk engines inside the advertising sector, privacy regulators are becoming more concerned about it. Given the popularity of SMS (short message service), it is simple to verify who owns a certain mobile phone number. Although this is a very reliable strategy, it has flaws that prevent it from ever being completely reliable.

Social Network, Email, and Instant Messaging Identities

Our internet identity is a crucial component of our overall identity, just like our mobile identity is. It might be the source of a few rather dependable parts, such as:

1. **Email Addresses:** Despite the fact that many people have many email addresses, some of which reflect varying degrees of anonymity, email addresses often stick with a person in the same way that phone numbers do.
2. **Instant Messaging Profiles:** Private handles on popular platforms like WhatsApp, WeChat, and others can connect us to actual people.
3. **Social Networks:** Facebook is used by over a billion people to log into many different services as well as to engage with friends and family. Social networks have some built-in trust when they are this deeply ingrained. As a result, they can help in confirming some aspects of a person's true identity.
4. **Professional Networks:** Through job and educational affiliations, as well as similar network-level trust assurance features like endorsements, LinkedIn delivers a comparable level of trust. However, social networks are ultimately regarded as identity systems with lower assurance. They weren't made to be utilized right away in high-risk transactions; rather, they were intended to be used regularly over time in casual social interactions.

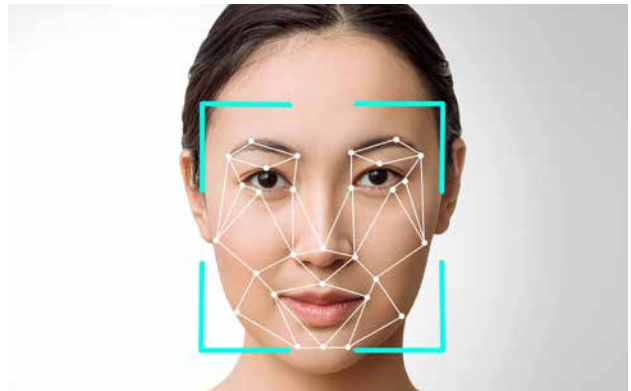
Methods for Verifying an Identity Cont'd

Biometrics

The distinctive characteristics of the human body that can be utilized as personal identifiers are known as biometrics. The most common biometrics are:

- **Fingerprints**
- **Retinal Scanning**
- **Facial Features**
- **Voice Patterns**

A biometric is insufficient for an initial identification verification, as was previously discussed. Unless the user's biometrics have already been recorded and registered to a reliable identity. They are therefore employed as a second factor to confirm the actual ownership of additional identification credentials or to authenticate a returning user as the account holder. Spoofing (masks, digitizers, or other technological cheating) concerns, data storage issues, and privacy concerns are a few more factors to take into account while employing biometrics.



Process for Verifying an Identity: Risk Modeling



Companies will weigh the user experience against their need to stop identity fraud or make sure they are familiar with their users for compliance when undertaking verifying an identity. A business's choice as far as an identity verification process is influenced by the trade-off between accuracy and speed, and the inputs can be utilized to determine an overall identity risk score. It is referred to as a risk engine. It gathers several signals and enables a determination of whether a fraudulent user is trying to attack the system within a specific risk threshold.

Any such risk engine aims to create a risk model that can quickly spot potential fraud without hurting the vast majority of authentic users. This turns into a standard analytics problem. Trained experts are also needed to design, inform, supervise, and occasionally directly participate in the risk system due to how much human judgment is now required.

There are differences in the reliability, accuracy, and user experience of the signals described. The most thorough method for enabling trusted individuals to get through and identifying fraudsters is to combine the right signals from a number of different ways. Consider all of the identity verification techniques in the comparison matrix before choosing the set that best suits the company's risk profile, speed requirements, and user experience objectives. In order to thoroughly assess a person's identification and risk online, IDs or biometrics should always be used in conjunction with databases, devices, and social network approaches due to their limited reliability.

Evaluation of ID Authenticity

One of the following types is usually present when a user presents a legal ID:

1. **Driver's License**
2. **National Identification Card**
3. **Passport**
4. **Residence Permit**
5. **Tax Identification Document**
6. **Voter Identification Document**



IDs are issued by around 200 countries. When all valid historical versions are considered, there are typically more than 6,000 different types of IDs in use at any given time. Each ID has a distinct data presentation and format. In order to help prevent some sorts of tampering, they also incorporate security measures like watermarks, checksums, and a machine-readable zone (MRZ), where the identity data is also recorded in optical character recognition (OCR) format. Each of these elements is crucial when considering an ID as legal identity documentation.

It's important to keep the following in mind while using government-issued IDs as a source of identification throughout a typical digital account onboarding process:

- All information must be digitally extracted from the ID and converted to a digital format
- The ID's digital form has to be remotely verified for authenticity
- The owner of the ID must be verified through a separate procedure

Like any security procedure, the procedures just mentioned should be polished into a procedure that doesn't inconvenience the vast majority of reliable users while still identifying a sizable portion of system cheaters. To add, an offline verification process enables a physical inspection of the ID. Contextual data can be easily gathered for both the ID and the user. Some of these indications are not present in an online procedure.

As a result, it only uses the ID and document holder's digital photos. To compensate for the absence of contextual signals, these methods must be even more efficient than simple human examination. There are additional elements that may be used to detect a fraud or imposter.

Evaluation of ID Authenticity Cont'd

ID Forgeries and Identity Fraud

There are numerous types of fake IDs. Some of the most common are listed below:

- **Forged IDs:** Fraudsters will illegally alter the information on the document in order to change all or a portion of the identity:
 - Changing the information that is variable
 - Adding real pages from a different ID
 - Eliminating certain pages or information
 - Applying fake stamps or watermarks
 - Modifying or adding information digitally to an image of an original ID
- 1. **Counterfeit IDs:** This is an exact replica of the original ID. A fraudster will usually purchase a template and add their own details and photo. These are a common alternative that may also be obtained illegally.
- 2. **Stolen Blank IDs:** Unpersonalized original IDs have been leaked from the manufacturing supply chain, and after this occurs, fraudsters add false information to them.
- 3. **Genuine IDs Obtained Fraudulently:** Fraudsters fabricate information on their applications. They might apply using a phony ID, a photo of another person, or a different collection of personal information. Authorities then provide them with genuine original IDs that are printed containing this false information.
- 4. **Fantasy or Camouflage IDs:** Fraudsters fabricate issuing authorities that do not exist or are not authorized to issue IDs.
- 5. **Imposter IDs:** The ID is original; however, it is being used by someone other than the rightful owner.
- 6. **Compromised or Example IDs:** A publicly accessible sample or image of an ID that has been issued by the government. Examples include IDs shared online, IDs from TV shows or presentations, or IDs that were reported as stolen or compromised to the authorities.

Evaluation of ID Authenticity Cont'd

Each of these types has a unique fraudulent trait, necessitating a unique means of detection. You require a set of ID analytics to examine each ID in order to identify these fraudulent practices in a systematic manner. These ID analytics make sure that the ID is authentic, hasn't been tampered with, and belonged to the individual presenting it.

Degrees of Fraud Sophistication

These many ID fraud techniques can be divided into different tiers of sophistication:

Tier 4: These attempts at ID fraud are unsophisticated amateur efforts. These include fictional or camouflage IDs (such as the Global Citizen passport, or British national identity card) that don't exist.

Tier 3: These ID templates were either improperly made or altered. Data validations, such as ID number formats or a lack of data consistency within the ID, can frequently detect these scenarios.

Tier 2: These ID counterfeits and forgeries are sophisticated. Each piece of information on the ID is accurate and makes sense. However, highly optimized technology detection processes can also pick up on small variations in fonts, layout, and security features, just like skilled experts can.

Tier 1: These ID fraud efforts are the most sophisticated ones. Blank IDs are stolen as a result of attacks on the supply chain of ID issuance and manufacture. Through deceit and social engineering, some fraudsters also gain genuine IDs illegally. In the black market, there are even highly sophisticated imitations. These frequently result from criminal organizations working with governments and usually sell for a lot of money. These IDs have the ability to deceive any machine-based method as well as the best-trained document experts. The only method to identify these cases is by thoroughly examining all of the IDs that the person has access to, as well as performing database and biometric cross-checks.

A toolkit of techniques is needed to evaluate the ID image's many components at once. More methods are required to spot the fake the more sophisticated the fraud strategy.

These detecting techniques will be covered in more detail in the next section.

Approaches to Detecting ID Fraud

There is a corresponding detection strategy for every fraud technique. Many IDs come with pre-built security measures as well as particular style and layout templates. These make it more difficult for fraudsters to illegally alter an ID and make it simpler for investigators to spot tampering.

The three main types of fraud analysis are as follows:

Analytics for Data Integrity

Examines the accuracy of the ID's text and numerical fields, both in free text and encoded in MRZ fields.

Analytics for Visual Document Authenticity

Checks the ID image's authenticity and looks for any digital alterations (typically in the form of anomalies).

Ownership Proof

Assessing a user's digital image (the individual claiming to be the ID holder) and comparing it with the photo on the ID.

Data Integrity Analytics

Numerous IDs come equipped with security features that can be systematically examined to verify the ID's authenticity. These consist of:

Data Validations

An ID must have accurate information in predetermined formats in all categories. The source template establishes the formats. This includes the encoded MRZ areas that use check digits to safeguard the integrity of the data. Data examples include gender, document registration numbers, expiration date, date of birth, and MRZ.

Algorithm Validations

To detect any ID integrity problems, certain IDs contain highly specific embedded rules in the form of mathematical algorithms.



Approaches to Detecting ID Fraud Cont'd

Data Consistency

Examples include the expiration date, kind, issuing country, nationality, document type, gender, and first or last name.

NFC Chips

The NFC (near-field communication) chip on many IDs contains additional data. Readers who are authorized can access this information and compare it with the data on the ID.

Visual Document Authenticity

The online processes are focused on in this white paper. Therefore, it is assumed that the ID being assessed is a digital image and that the user will identify themselves by the capture of a digital image, such as a photo or a video. The best method for determining whether these photographs are legitimate is through digital analysis because it can identify forgeries that might be more common when a human examiner is not present. There are various methods to assess the ID and the user, which can assist in identifying potential tampering and impersonation from various angles:

Comparing Document Templates

Errors or inconsistencies can be found by comparing the provided ID to the known document template.

Digital Tampering

Fraudsters have the ability to alter an ID's digital image using software. This could entail changing or removing the user's image. They then turn in the altered copy along with the fabricated image.

Font Anomalies

Fraudsters frequently attempt to modify data fields but leave behind font inconsistencies in the process.

Copies of a Document

Instead of submitting the original ID, fraudsters will also attempt to upload a replica of it, usually in a modified form.

Approaches to Detecting ID Fraud Cont'd

Security Features

All forms of IDs include some sort of built-in security features. These can be evaluated to guarantee authenticity.

Here are some examples:

- **Digital watermarks**
- **Barcodes**
- **Embossed text**

The versions and ID types of these differ.



Facial Comparison as Evidence of Ownership

We need to ensure that the user providing the ID is its authorized owner in addition to ensuring that the ID is authentic. We can verify facial likeness between a live user capture and the ID photo to demonstrate ownership. Selfie photos or video liveness are two distinct capture experiences that both offer facial images in near real-time.

Using sophisticated 3D masks or a printout of the target's face, for example, or images from their social network account, fraudulent individuals will try to fool the system. Mathematical comparison methods and online verification routines help defend against these attack vectors. By confirming that the image was made on the user's device, we can reduce the possibility of images from other sources being used.

Selfie Images

They employ passive anti-spoofing algorithms to defend against common spoofing attempts, such as supplying a photo of a photo.

Video Liveness

The maximum level of protection against complex spoofing efforts is provided, including deep fakes. The user partakes in tasks like doing random gestures in place of snapping a photo to verify their identity. This demonstrates the “challenge and response” interaction's liveness.

An Approach to Detecting Document Fraud with Machine Learning

But in a digital environment, how do we spot fraud on a large scale? This can be a fake ID, an impersonator, or both. It is necessary for us to establish that IDs are real and belong to the users. This entails examination by either relying on skilled human specialists or a machine designed to programmatically evaluate an ID and assess its fraud risk. Even when carried out by a professional, manual inspection doesn't scale effectively for big numbers of transactions and account onboarding for online accounts.

The Machines

There are two distinct strategies when it comes to machines. The first model is a traditional heuristic one. This employs static software to directly apply pre-defined techniques. The second strategy employs machine learning and trains the model using data. As it analyzes additional data, it continuously learns, which boosts its efficiency. Heuristic models can only detect fraud of a certain level of skill.

Trade-offs are necessary to create an optimal solution, whether it is a machine or a human.

Human Pros:

- Rapidly begins processing ID with basic training
- Makes more intricate judgment and experience decisions
- Creates data for machine learning models to be trained
- Machines are double-checked to reduce false positives
- Can recognize contextual signals

Human Cons:

- Built-in restriction on how quickly they can finish processing
- Over time, fatigue reduces performance
- Cost will always exceed that of a machine, particularly on a large scale
- Unable to accurately determine whether two unknown faces match

An Approach to Detecting Document Fraud with Machine Learning Cont'd

Machine Pros:

- Never grows weary
- Will always accelerate
- Organize enormous sets of IDs and regulations required on a global scale
- Can prevent some fakes that human couldn't
- Greater accuracy and precision

Machine Cons:

- Must receive training
- Reaction period is necessary for new documents and fraud tactics
- It's not always obvious when to resort to using human judgment

Numerous studies on facial comparisons reveal that people frequently struggle to distinguish between two unfamiliar faces that are a match and that face matching algorithms frequently surpass them.



An Approach to Detecting Document Fraud with Machine Learning Cont'd



From this, we can conclude a few things:

1. A human-only operated solution will:

- Not be able to catch all fakes, and many that are more elaborate and sophisticated will be missed
- Have scale restrictions, both in terms of price and processing time
- Make mistakes positively and negatively when detecting fraud

2. A solution that relies solely on machines will:

- Catch more potential fakes, especially when it comes to new, sophisticated attacks
- Require human intervention as a means to identify new fraud techniques and the training of machine learning models with significant data sets
- Require extensive machine learning infrastructure and software development knowledge to create and maintain models

The Process of Document Verification

A government-issued ID can be used to verify someone's identity by taking on the following process:

Document and Biometric Capture

- Perform a digital scan of the ID to get the best image quality for analysis
- Gather a real photo or video of the candidate
- Obtain fraud indicators and fingerprints from any device or environment

Classification of Documents

- Type of document
- Document that is issuing authority
- The version of the document

Extraction and Validation

- Read and note any important PII and document characteristics from the text and MRZ areas
- Check all data, forms, and security requirements for document templates

Analytics for Document Authenticity

- Examine the characteristics of the digital document to make sure it hasn't been modified or tampered with
- Verify that it is not a fake that has been exposed or is widely available
- Check with the issuing authorities to see if the document has been issued

Identification Through Facial Comparison

- To determine whether the user's captured image and the ID's photo have similar faces, calculate risk likelihood
- Increase the number of human liveness tests to prevent spoofing

Additional Signals

- To demonstrate the identity, check the information on the ID with a reliable identity database
- Check the information of the device network to make sure it doesn't have a known fraud profile

FTx Identity Product Platform

The identity verification (IDV) platform designed by FTx Identity is a risk engine that integrates several signals to determine whether a user is really who they say they are. The platform compares them to a digital image of the person in a photo and uses international identity documents as its primary source of identity.

Machine Learning

In today's enterprise software market, it is widely acknowledged that machine learning is the best method for meeting the needs of significant global businesses on a huge scale. Global businesses present a variety of intricate issues, particularly for verifying an identity. For starters, it is challenging to analyze a sizable, dynamic set of international government-issued IDs (and the variants that go along with them). It is difficult to categorize those IDs, retrieve their data, and determine their veracity. The system must operate efficiently, reliably, and promptly. A machine learning-based strategy is deemed by FTx Identity to be a practical means of accomplishing these outcomes.

Utilizing Biometrics for Anti-Spoofing and Facial Verification

To examine a broad range of users, we require a sophisticated set of procedures. We must confirm that those users are actual people. We need confirmation that they are not being impersonated. And we need to make sure their picture matches the one on their ID. A list of some of the procedures used by FTx Identity is provided below.

Many of these procedures call for machine learning models. The system scans the image and extracts patches to identify spoofing in selfie photos. Each patch is examined for signs that it is a reprint of a photo or a digital screen.

Several of the procedures used by FTx Identity include:

- **Headpose (Face Tracking)**
 - Tracking facial movement to confirm movements
- **Sherlock (Texture Analysis)**
 - Finding images of images and images of screens using texture-based analysis

The technology used by FTx Identity and our usage of machine learning will be thoroughly examined in the next section.

The Machine Learning Technology of FTx Identity

The Challenges of Verifying an Identity

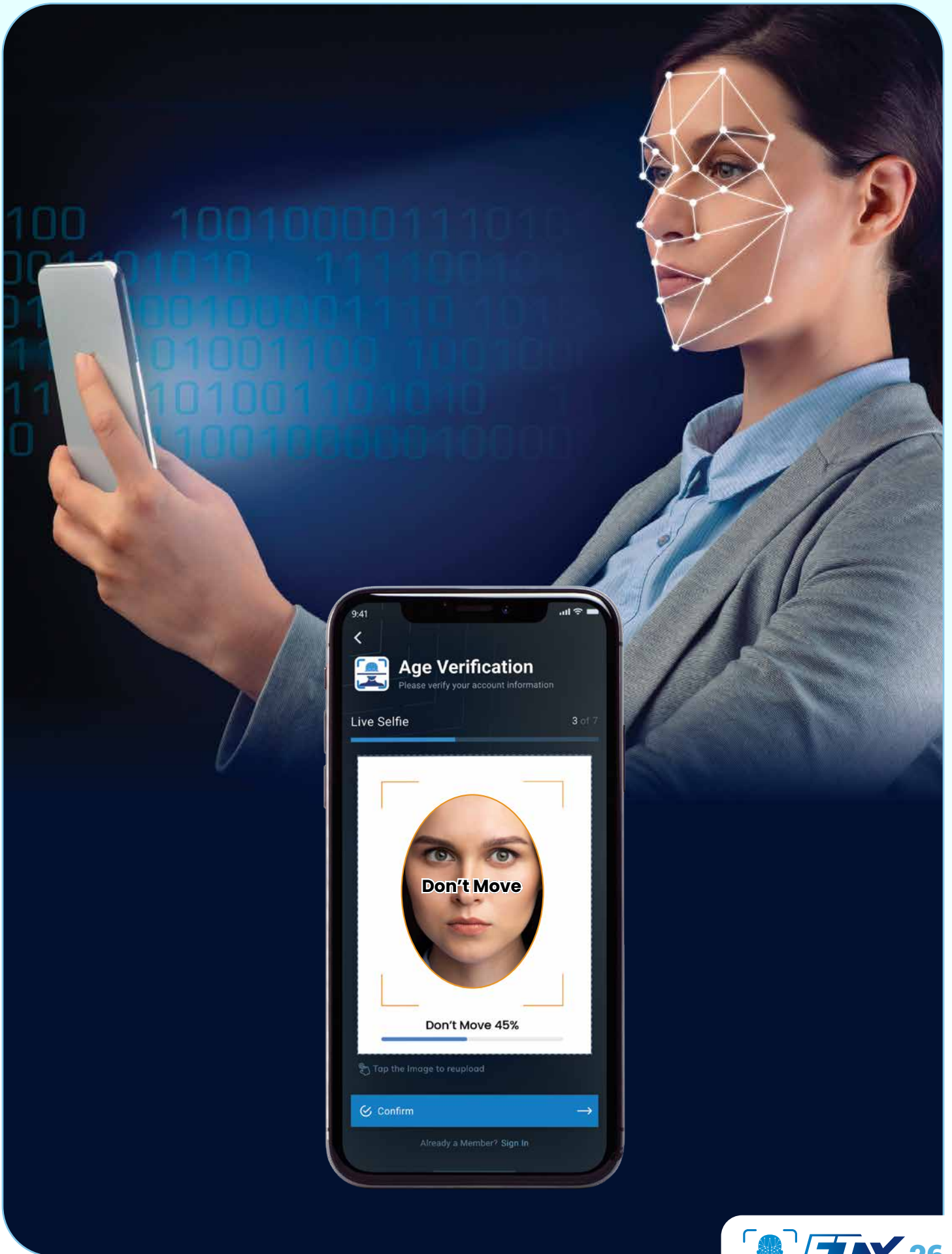
Each of the listed procedures of verifying an identity involves judgment calls. For instance, whether two faces belong to the same identity or whether the MRZ's font is the right one. These decisions are powered by machine learning, which uses a digital image as an input and produces a decision with a level of confidence. Images increase the intricacy of this procedure. Because of this, FTx Identity's technology also incorporates computer vision components.

Images present a challenge due to their great dimensionality and low signal-to-noise ratio. Because the object of interest frequently only makes up a small part of the image, only a small percentage of the pixels in an image will have a significant impact on the decision. This effect is magnified while developing a fraud detection system because we frequently search for small differences. The system must be trained to accept images of various quality and distortions since users capture images in an uncontrolled environment, similar to the human brain.

The technology of FTx Identity makes use of cutting-edge deep learning methods. One significant distinction is that in a deep learning system, we can learn the picture features rather than using ones that were manually created. By doing this, we can model more complex functions and the model is given more expressive power. Massive volumes of data, both internal and external, can be utilized thanks to these models. Better outcomes derive from more data.

Our technology can quickly and broadly detect items. Our technology looks for irregularities in the data it receives to identify fraud. The amount of "regular" data will outnumber the amount of abnormal data in an anomaly detection scenario. In contrast, the majority of datasets used in academia are balanced, which means that each class is fairly well-represented.

Only a few instances of fraud will often be found in a dataset from an online company that requires online identity verification. This impacts how the decision-making process is structured. Decision-making must be based on how much a new fraudulent submission deviates from the model's norm if we can only model the normal class.



The Machine Learning Technology of FTx Identity



Focus Areas

Facial biometrics and identification documents have been the two problem domains that FTx Identity has heavily prioritized. The development of proprietary models trained on owned data has greatly enhanced system performance on both. Data extraction from documents is one instance. OCR is already highly commoditized, with top-notch products from companies like ABBYY, Google Vision, and Microsoft Cognitive Services. We can significantly improve performance by creating a custom OCR system that is tailored for matching photos on IDs.

Similar to this, FTx Identity's facial recognition technology has been trained to perform far better than cutting-edge general purpose systems. We can model the true class and tune the parameters using the fraudulent examples by recasting the issue as an anomaly detection one. This implies that far more data from the true class can be used. For instance, when we use this method for MRZ font verification, the quantity of false positives produced by the model can be cut in half.



The Future of Identity is Here

Technology that verifies an identity is now accessible to everyone, not only higher risk financial institutions, because of advancements in smartphone ownership, machine learning, privacy awareness, and cloud architecture. Additionally, the move away from centralized, sensitive PII databases necessitates a new structure in terms of both technology and regulation. By protecting their identities from scammers, verifying an identity safeguards customers. Customers desire reliable identity proofing just as much as companies do since we all detest having our accounts compromised.

However, in order for a system to be widely used, it must also be convenient so that people won't avoid it. The key to this will be the capability of identity porting, or the secure and safe transfer of identification between providers. Customers experience less friction, and businesses spend less money doing it. Our goal at FTx Identity is to create an open universe where identity serves as the key to access. Some amazing new possibilities for the next chapter of identity services can emerge by connecting offline identities (government-issued IDs) with a new, digital credential established upon verification of an identity claim.

Decentralized Identity

Nobody wants to make it possible for a single business to store and control all users' identity information worldwide and digitally today. These systems are impossible to keep secure, and once they are, the trust in the entire system is broken. A better system would be dispersed, owned by the end-user, and housed in that manner. Unless the user decided there should be a single owner, there would not be one. The technological foundation is already in place thanks to blockchain and distributed ledgers. This system may come to pass after the relying parties' legal structures are established. These and other advancements have made digital identity a rapidly growing field. And FTx Identity is promoting this transformation with its constantly developing technology. Experience the best identity verification platform today!

