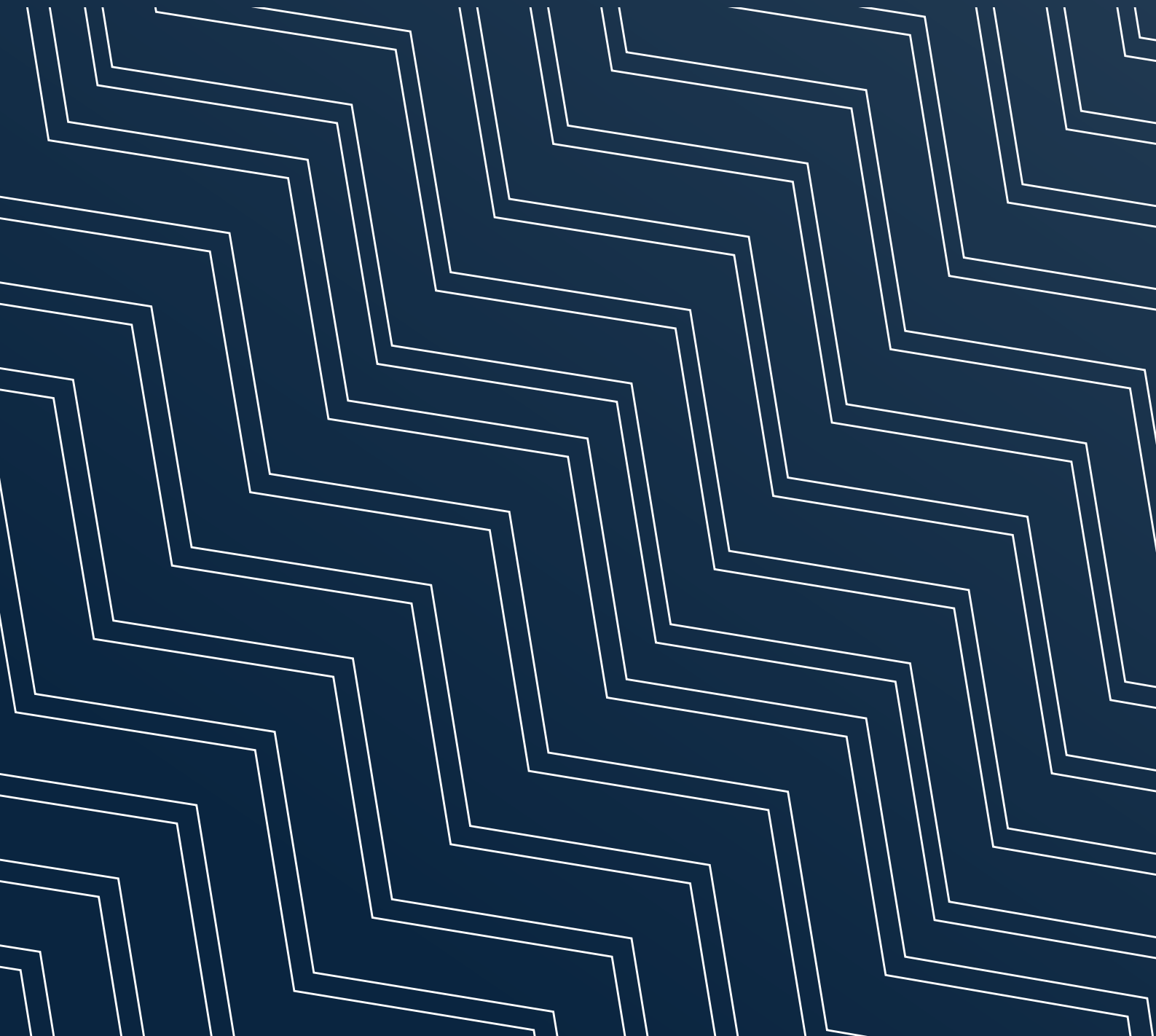
AML Watcher

Whitepaper

# DECODING NAME SCREENING NIGHTMARES IN MENA AML COMPLIANCE



# DECODING NAME SCREENING NIGHTMARES IN MENA AML COMPLIANCE

# Introduction

It was mid-February 2011, and the streets of Egypt, Libya, and Tunisia were roaring with calls for revolution. Widespread demonstrations wreaked havoc in these autocratic regimes that once seemed impossible to shake.

Dictators like Zine El Abidine Ben Ali of Tunisia and Hosni Mubarak of Egypt had already fallen.
However, Libya's Colonel Qaddafi refused to stand down amid widespread anti-government protests and unleashed a crackdown.

Countries around the world watched in horror.

On February 26, the United Nations Security Council (UNSC) moved quickly and adopted Resolution 1970 in the wake of the widespread crackdown. The UNO obligated its 193 member states to freeze assets of Qaddafi, his close associates and members of the Libyan government.

The same day, US President Barack Obama announced direct sanctions against the Libyan government, and ordered to block any transactions involving assets of Col. Qadaffi and his inner circle.

It turned out there was no standard way to write Arabic names into English. For instance, the Arab name “Qaddafi” has a dozen other ways to transliterate into English. It could be "Gathafi," "Gaddafi," "Kadafi," "Gadafy," “Khaddafi," "Gadaffi," or "Khadaffi."

Variations on his birth name included “Muammar,” “Moammar,” “Mu'ammar,” and “Moamar,” and many others. After settling on how to spell his first and last names, the compliance teams had to decide further what Arabic prefix should be used before his last name, it could be “al-” or “el-”.  And then the banks had to further decide whether the prefix should be capitalized.

This multiplicity of spellings created an unholy mess for the international banks.

Regional naming conventions would further complicate the pronunciation, which means there are wild variations in how a single name is spelled. At that time, banks had the practice of letting customers write their names however they preferred to write.

What compounded this problem the most was the fact that the Arabic language does not use any vowels, so transliteration into English was no less than a guessing game. So, the task of complying with the international watchlists and freezing sanctioned entity assets had just run headfirst into a complex linguistic riddle.

What started as the fight for freedom quickly turned into a fight against spelling.

# Why Name Screening Is An Uphill Battle For Financial Institutions

For any financial service provider obliged to follow KYC/AML regulations, the most important information to confirm in the Know Your Customer (KYC) process is to verify a client's original name. A client's name is their most fundamental piece of identity. No bank wants to make a mistake when verifying a person's basic identity.

Names can be tricky. It becomes trickier when name screening is carried out in diverse languages.

These screenings usually account for multilingual adverse media and government watchlists in diverse languages, plus numerous regional writing systems.

When a bank conducts adverse media screening, it once again requires a correct name to identify risks connected with the person across various media forums.

Media reporting, geography, culture and politics all create complex layers of challenges for effective name matching in due diligence and KYC procedures no matter whether you employ hi-tech solutions to perform the matches.

However, documenting a transliterated name from paper to a digital medium may also introduce errors. So, the question is who audits an entity's name while moving from point Y to point Z? Actually no one and this is why depending on a signature to spell a name can create further problems.

It's not as easy as it seems to tell the difference between two spellings of the same word when they are used in different languages and scripts. Unpredictability always increases by the variety of keyboard templates.

Latin languages that share the same alphabet, such as French and English, may have distinct keyboard layouts, logographic scripts may be susceptible to the "Wubi Effect," in which a population speaking just one tongue uses a number of opposing keyboard layouts.


Zaynab
Zeinab
Zainab
زينب
Kalid
Khalid
Khaled
خالد
Husayn
Hussain
Hussein
حسين
Mohd
Muhamad
Mohamed
محمد
Ahmet
Ahmad
Ahmed
أحمد
Yusuf
Youssef
يوسف
Yousuf
Ali
علي
Aly
Alee
Fatimah
Fatma
Fatima
فاطمة

# How Do Political and Cultural Shifts Lead to Confusing Name Variations for Banks?

Different cultures use completely separate writing systems (e.g., Cyrillic, Latin, Arabic, and Chinese characters) and spoken languages.

Within the same language, too, there are common practices and diverse rules for writing names. Some cultures prefer to write family names first instead of last and some cultures may have different placement for middle names.

Social norms also make a difference in how names are given and spelled. It's something that evolves over generations and isn't really static.

Sometimes the name challenges are inherently about cultural preference or linguistic evolution. Over the course of history we have seen that often political systems and governments can directly dictate how names are translated, spelled and documented in the official records.

It's not the linguistic accuracy that propels such objectives. Such goals are driven by political agendas too.

# How Linguistic Reforms Present an Ongoing Challenge for Name Screening in Banks?

One instance of this happened in the Soviet Union when the central Soviet government and the Communist Party of the Soviet Union (CPSU) leadership in Moscow ordered the enforcement of the Russian language and the Cyrillic alphabet in the Central Asian states such as the Kazakh, Uzbek, Kyrgyz, and Azerbaijani regions which had used Arabic or newly adopted Latin scripts, and the Baltic states, including Estonia, Latvia, and Lithuania which used Latin-based alphabets.

This policy was introduced under the agenda of "indigenization" or "Russification" of all soviet states in an effort to modernize them and build social cohesion and harmonize the diverse ethnic and linguistic groups under a common Soviet identity.

This led to the transliteration of Baltic names (originally written in Latin-based alphabets) and Central Asian names (originally written in Arabic script) into the Cyrillic-based alphabet.

Change in writing systems brought about confusing variations in forms of names, how they were spelled in previous scripts and converted scripts, and their previous and new pronunciations.

However, when the Soviet Union disintegrated in 1991, the newly independent states that previously adopted the Cyrillic script now reverted back to their original writing systems (Latin or Arabic) which resulted in a second new wave of name changes for the same generation of families and individuals.

So, a name isn't just a philological creation. Political influence, state policy and cultural norms also determine how a name is transliterated and recorded, and how it may end up having multiple variations which creates a confusing situation for financial crime compliance that claims to offer accurate name matching.

# What Makes the Arabic Names So Tricky for MENA AML Compliance?

Arabic is a language that is largely spoken in the Middle East and North Africa (MENA) region, which accounts for 6% of the world population. The region's combined GDP is approximately $3.3 trillion every year across 21 diverse countries.

MENA, where ancient civilizations meet modern finance, is known for its complex and multi-layered landscape for Anti-Money Laundering (AML) and Counter-Terrorist Financing (CTF).

The biggest challenge for Middle Eastern banks in AML compliance is the diversity of scripts in watchlist data.

While major international regulatory bodies, such as the U.S. Office of Foreign Assets Control (OFAC) and the UK's Office of Financial Sanctions Implementation (OFSI), primarily issue their sanctions lists with names in Latin-based English (often transliterated from non-Latin originals), other global bodies like the United Nations (UN) also include names in various non-Latin scripts.

This complexity is further compounded as regulatory bodies within Arab countries frequently publish their national watchlists directly in Arabic script.

The core technical hurdle lies in the absence of a universally standardized transliteration system between Arabic and Latin scripts, leading to numerous legitimate spelling variations for the same name.

Consequently, MENA banks, whose customer records are predominantly in Arabic, face the arduous task of accurately matching and screening names across this multifaceted linguistic and script landscape, often resulting in high volumes of false positives requiring extensive manual review, or critically, the risk of false negatives if advanced phonetic and fuzzy matching algorithms are not effectively deployed."

Currently, Arabic is classified into three categories:

- Classical Arabic
- Modern standard Arabic
- Colloquial or dialectal Arabic

Classical Arabic is known to employ classical diacritics in classical poetry and classical books. The modern Arabic language is widely used in government, news channels, media agencies and textbooks, has stems from classical Arabic.

Unlike classical Arabic, modern Arabic's unique feature is that it rarely uses diacritics.

In some cases, the modern Arabic language employs diacritics to distinguish between similar words. Colloquial or dialectal Arabic has a number of regional and national diversities that make up the everyday spoken language.

There are numerous local variations of this kind of Arabic. Some linguists view them as separate languages because they can occasionally differ sufficiently to be incomprehensible to one another.

In Arabic, words are written in a string of consonants with little or no indication of vowels, known as unvocalized Arabic, which is highly ambiguous. Sometimes, diacritics are used to indicate short vowels; however, their use is very rare.

Lack of vowels in Arabic means any written Arabic word can sound like or have many different meanings. This creates a lot of confusion and makes the transliteration of Arabic names into Latin-based English super hard for compliance officers.

# AML Screening Challenges with Arabic Names & their Multiple Variants

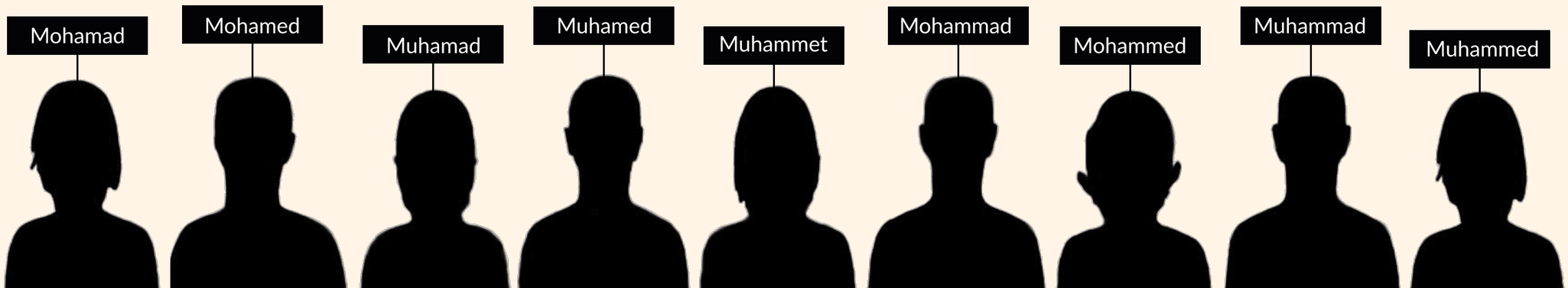
There are arbitrary ways to write Arabic names in English-based writing systems. It's the lack of certain vowels in Arabic that leads to this randomness. The diversity of Latin spellings can be attributed to dialectical variations in vowel sounds.

It is challenging to effectively translate Latin names into Arabic using a rule-based method since a single Arabic name frequently has multiple Latin variations.

The tricky part when it comes to names is that they are not as distinctive as the people who own them.

In the Arab world, for example, the most common Muslim boy' name is Muhammad due to its connection to the founder of Islam. There could be 150 million people who bear the name Muhammad, which varies in spelling from region to region and country to country.

The Arabic name **محمد** can be spelled in multiple ways using the Latin script.
**Alternatives Include:**



# Map Of The Spellings Of Muhammad Around The World

| Example of Naming Variation | Example of Arabic Name |
| --- | --- |
| Common name variations | Mohammed vs Muhammad |
| Phonetic matching | Yusuf vs Yousif |
| Similar names | Hassan vs Hussain |
| Repeated characters | Alee vs Ali |
| Abbreviations/initials | A. R. Rahman vs Abdul Rahman |
| Missing names | Karim vs Abdul Karim |
| Name order | Al-Sayed Ahmed vs Ahmed Al-Sayed |
| Noise | M0HAMMAD vs MOHAMMAD |
| Different naming conventions | Abdul Rehman vs Abd al Rehman |

Banks name-screen their clients using adverse media, PEP lists and watchlists from different regions which includes names in a myriad of scripts in multiple different languages.

Since there are dozens of ways to write a single name, even a slight change in a name can help high risk entities to evade detection.

Some cultures have common family names that go on for generations. Such customs makes name screening explode with a high volume of false positives and leads to a waste of compliance resources.

Considering the broad range of spellings in these alternates, it is evident why a lexically-based method is required to transliterate these names from Latin to Arabic. This meant that rules cannot reproduce the arbitrary character of Arabic name orthography when expressed in Latin letters.

So, it turns out that using a broad list of names and variants can help to limit the number of false matches.

# AML Screening Challenges With the Arabic Dialect:

Then come the dialectal differences.

Numerous national or regional dialects of Arabic that make up the commonly spoken language are referred to as colloquial or dialectal Arabic.

The Arabic letter ج is spoken as "g" in "good" in Egypt but as "j" in "jungle" everywhere else in the Middle East. It gets confusing when translated into English as either "Gamel" or "Jamel."

There are numerous regional variations of this kind of Arabic, some linguists view them as separate languages because they can occasionally differ sufficiently to be incomprehensible to one another.

They can be found in specific types of written media, such as poetry and printed advertisements, as well as frequently in casual spoken media, like talk programs.

# AML Screening Challenges Due to Different Naming Conventions

Another challenge that makes **AML** compliance hard in the MENA region is the segmentation of Arabic names.

There are a number of ways to write the name “Abdul Rahman” or “Abd al-Rahman” or “Abdurrahman.” They may suggest the same name, but they could be different persons or high-risk entities just writing their names differently to evade detection.

This lack of standardization results in high-risk entities evading the AML screening systems. If the screening systems are not advanced enough, criminals can leverage this weakness by conducting suspicious transactions from multiple accounts using slightly different segmentations of their names in different banks.

Banks that offer automated name screening systems may fail to detect such variations caused by segmentations in Arabic names. For example, a watchlist may designate a person named "Abd al-Rahman," but the bank’s customer record may miss it and may have recorded the name as "Abdurrahman."

This slight change in segmentation may bring an explosion of "false positives" (a legitimate person is flagged multiple times due to different spellings of their own name) or "false negatives" (a high-risk person is missed because their name is segmented differently), which ends up in a complete waste of the bank's compliance resources.

# AML Screening Challenges With the Arabic Script vs. Latin Script

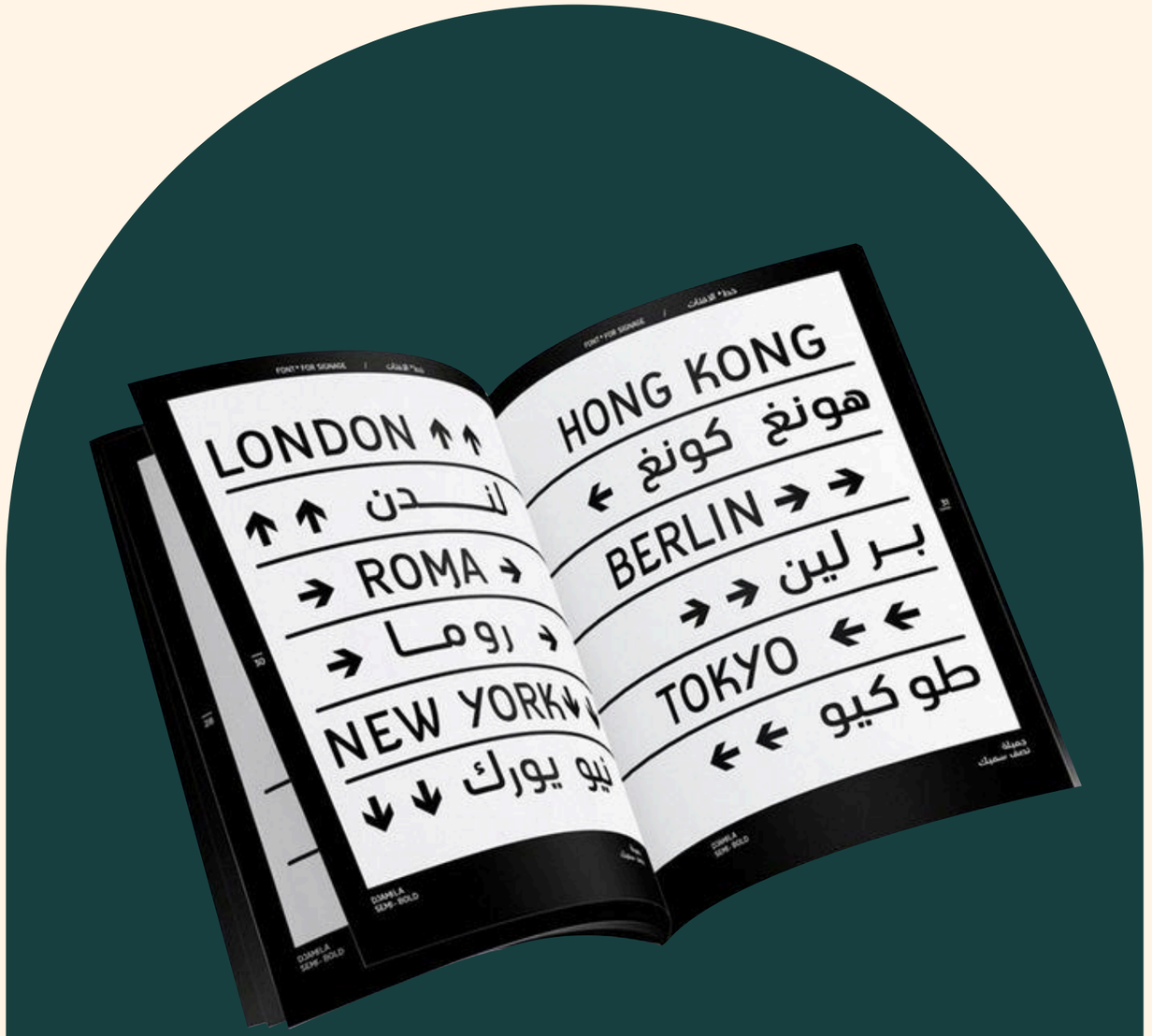
Additionally, the Arabic language has its own script or alphabet, which is very different from most Western languages which use the Latin script as their primary alphabet system. During transliteration, which is a process of writing words using different alphabets, it becomes very difficult to write Arabic-based names into Latin-based English writing system. Some Arabic alphabets don't have their English counterpart to accurately represent the same sound and script in another language.

For instance, in Arabic, the letter "ق" represents a distinct "k" sound. It sounds like the English "k," except it is pronounced in the back of the throat. In English, it is commonly transliterated as "k" or "q."

Another example is ع. In the English writing system, it is known as 'ayin'. It's one of the renowned Arabic "guttural" sounds (those made in the throat). In Latin-based English, it is frequently marked with an apostrophe or left off entirely, which leads to spelling variations.

Other such Arabic consonants that do not exist in the Latin-based European languages include ʔˤ], ض] dˤ], ط] tˤ] and ظ] ðˤ].

Hence, names starting with such Arabic consonants are difficult to pronounce and create lots of confusion.



# AML Screening Challenges with Arabic Pronunciation

In the Arabic language, their Arabic pronunciation may differ slightly from the way it's written. For example, a name like “al-Din” may be pronounced like the following letter, i.e., “ad-Din.” However, it’s always written in Arabic with its letter for “l” (“al” is the definite article).

Different rules apply to different languages while translating them into Arabic. The Arabic letter "ش," for example, is transliterated as "ch" in French and "sh" in English. The problem is that English is a translation of French, for instance.

The problem is that French transliterations frequently appear in English-language documents, particularly when the English text is a translation of the French.

As a result, the variations "Bashir" and "Bachir" are seen.

Matching names as part of an AML procedure presents many more challenges than these transliteration issues for matching names when both are in Latin characters.

A Middle Eastern bank might maintain a database with names in Arabic script, however, lists such as OFAC are primarily in Latin characters. And this is where name screening poses the challenges.

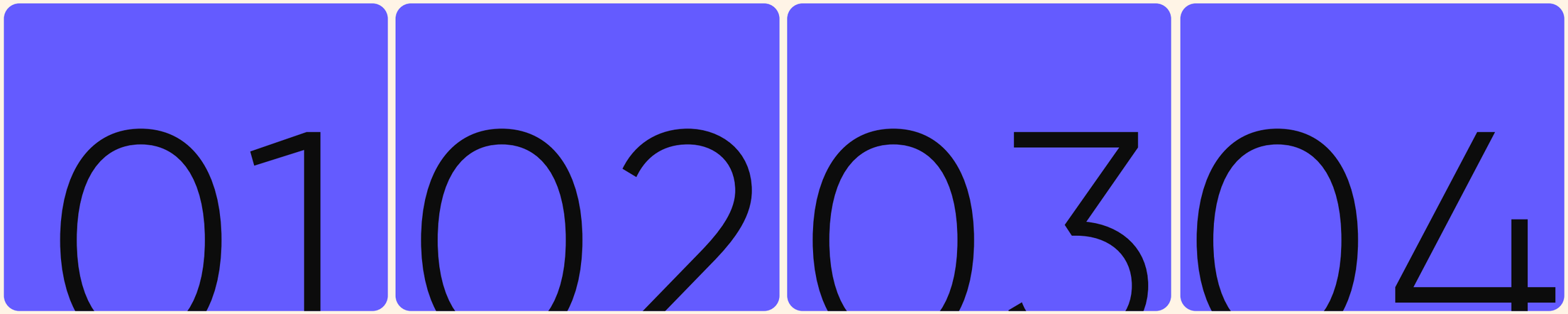
# How Can AML Watcher Help FIs Overcome Arabic Name Screening Challenges?

Arabic names present one of the most complex blockades in Anti-Money Laundering (AML) screening notably for banks and financial institutions operating in the MENA region.

Variations in spelling, transliteration, name order, and prefix/suffix usage can all cause missed matches or false positives.

AML Watcher is a solution forged from the frustrations of AML compliance.

It deeply innovates the AML Data layer by intelligently mining regulatory data in line with jurisdiction-specific regulations, and its truly innovative application layer provides an exceptionally efficient name screening process.



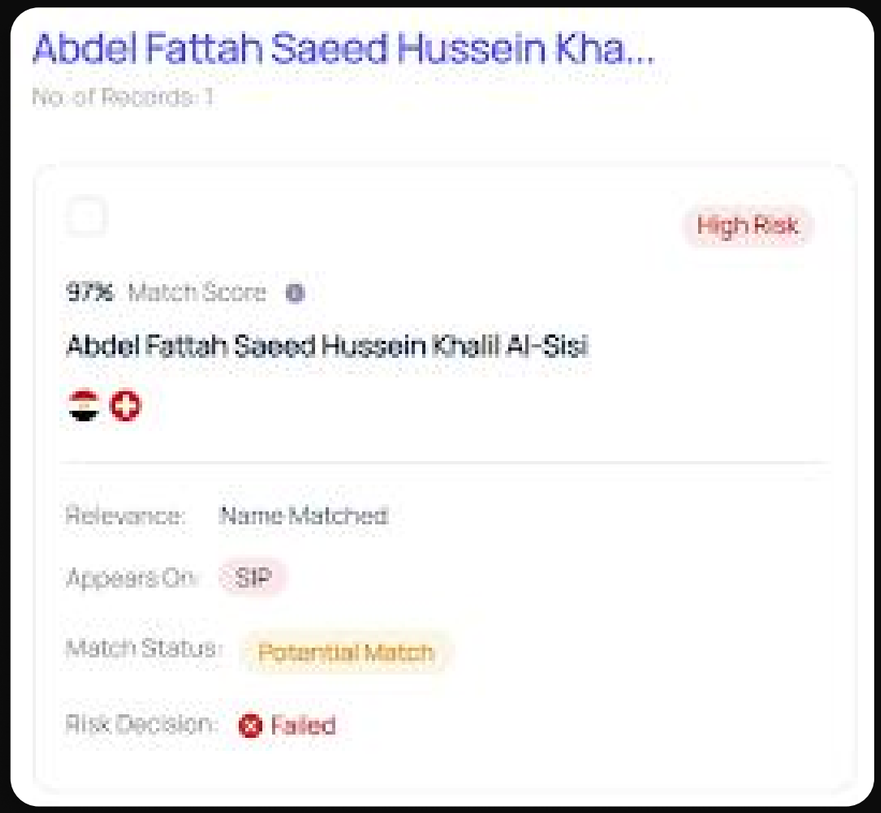
# 1. Matching Arabic Names with Multiple Transliteration Variants

## Challenge:

Arabic names can be written in numerous ways when transliterated into Latin script due to the absence of standard vowel usage, dialectical pronunciation, and regional spelling preferences.

## AML Watcher Solution:

AML Watcher uses multi-variant matching algorithms and transliteration normalization to recognize alternate versions of the same name across watchlists, databases, and customer records. Its proprietary system can map dozens of spelling variants to a single entity with high precision.




Example:

- Searched Name: Abdel Fattah Saeed Hussein Khalil El-Sisi
- Recognized Variants:

1. Abdul Fattah Said Husayn Khalil Al-Sisi
2. Abd Al-Fattah Sa'id Husayn Khalil El-Sisi

- Result: Successfully matched across variants with zero false positives, ensuring that sanctioned individuals are never missed due to spelling differences.

# 2. Handling Long Arabic Names Without Truncation

## Challenge:

Arabic names can be lengthy, often including given names, family names, tribal identifiers, and geographical origins, which increases the risk of truncation or mismatches during automated screening.

## AML Watcher Solution:

To support full-name parsing, AML Watcher extends character limits in entity fields, implements phonetic search algorithms, and optimizes multi-part name tokenization. This allows the system to maintain integrity even with extended name structures.

Example:

Searched Name: Amr Mohamed Zaki Mohamed Abdel Aal

Matched Against: عمرو محمد زكي محمد عبد العال (Arabic script)

Result: Matched with 80%+ confidence score and no false positives, despite the complex name order and script differences.

# 3. Multiple Name Order Variations

## Challenge:

In Arabic culture, name components can appear in varying sequences, such as given name before or after the father's or grandfather's name. This non-linear structure confuses most rule-based name screening systems.

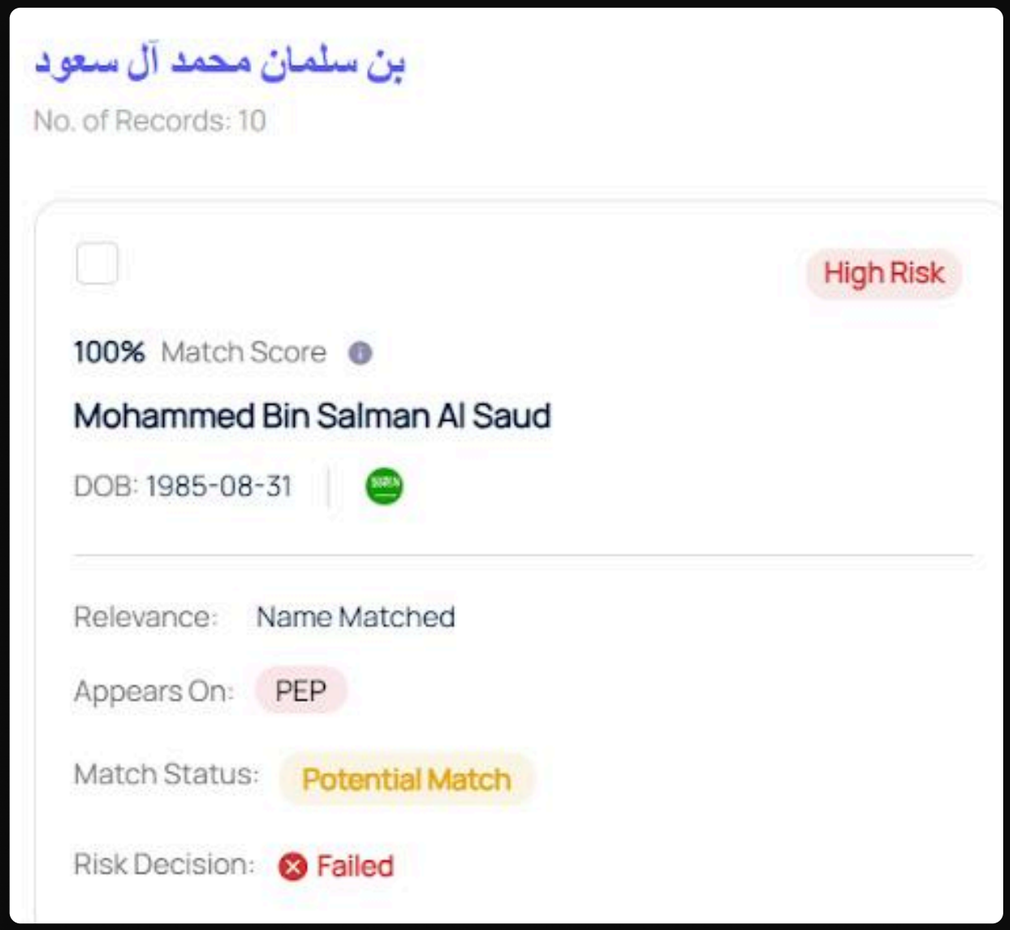
## AML Watcher Solution:

AML Watcher leverages a dynamic name order recognition model, combining name permutation logic and fuzzy matching algorithms to identify reordered or rearranged names, even if the data structure is inconsistent across systems.

Example:

- Name Searched: "محمد بن سلمان آل سعود" .vs "بن سلمان محمد آل سعود"

**Result:** Successfully matched despite order differences, with no false positives.



# 4. Identifying Prefixes & Suffixes:

## Challenge:

Arabic names often include prefixes like “Al-,” “El-,” “Abu,” “Ibn,” or suffixes like “al-Din” that may or may not be written in formal records. These additions can throw off traditional name-matching systems.

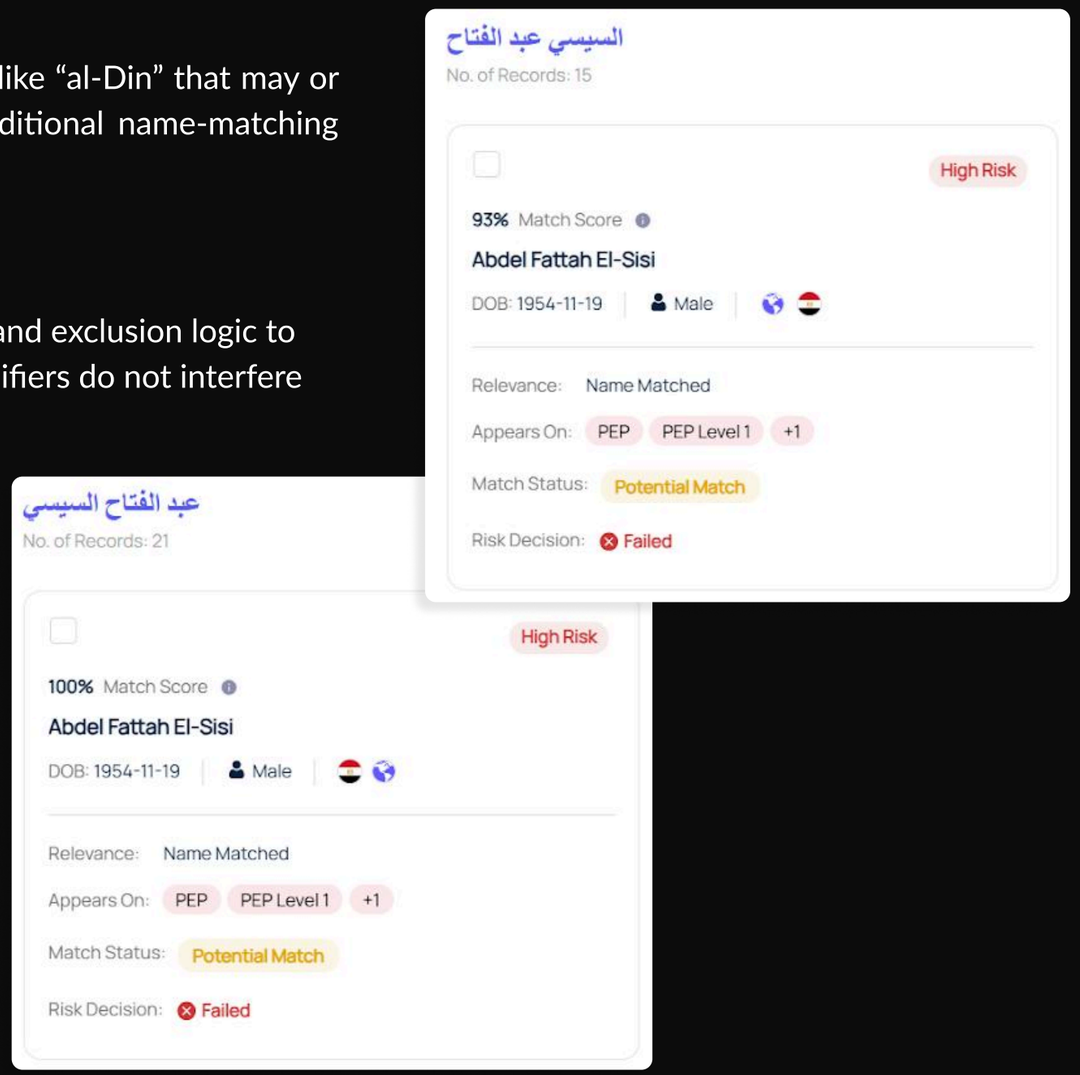
## AML Watcher Solution:

AML Watcher employs intelligent name parsing using rule-based inclusion and exclusion logic to detect and adjust for optional prefixes and suffixes, ensuring that such modifiers do not interfere with accurate matching.

Example:

- Name Searched: "عبد الفتاح السيسي" .vs "السيسي عبد الفتاح"

Result: Successfully matched, regardless of prefix position, with 100% accuracy and no noise from irrelevant alerts.



# Advanced Entity Resolution for Lower False Positives

AML Watcher significantly reduces false positives through a multi-dimensional entity resolution framework that extends far beyond conventional name matching. The platform combines biometric verification, contextual data correlation, and AI-powered natural language processing to create a robust identity resolution system. By cross-referencing multiple data points and applying intelligent matching logic, AML Watcher ensures high-confidence alerts while minimizing noise in compliance workflows.

## Biometric AML

AML Watcher's biometric screening solution addresses the above-mentioned challenges in name screening by augmenting traditional text-based searches with facial recognition technology.

Financial institutions can now submit entity images alongside names, which massively reduces false positives and improves match accuracy. This dual-factor verification enhances due diligence by minimizing manual intervention and simultaneously confirming compliance with global AML/CFT regulations.

With real-time access to over 3,500 watchlists and sanctions data across 235 jurisdictions, the solution enables FIs to conduct faster, more reliable risk assessments without compromising regulatory adherence.

The platform's strength lies in its comprehensive and dynamic data infrastructure, which aggregates Politically Exposed Persons (PEPs) lists, sanctions registers, and adverse media from 100,000+ sources updated every 15 minutes to reflect the latest risk intelligence.

The platform's strength lies in its comprehensive and dynamic data infrastructure, which aggregates Politically Exposed Persons (PEPs) lists, sanctions registers, and adverse media from 100,000+ sources updated every 15 minutes to reflect the latest risk intelligence.

Advanced features such as phonetic matching and transliteration further mitigate false negatives caused by linguistic variations, ensuring robust screening across 80+ languages.

A case study involving a KYC provider demonstrated a 44% reduction in false alerts, which validates the effectiveness of integrating biometric verification with conventional name-matching methodologies.

This not only streamlines compliance workflows but also enhances risk detection capabilities.

# Advanced Contextual Data Correlation

AML Watcher employs a multi-dimensional contextual matching system that goes beyond simple name screening by integrating biometric, demographic, and documentary verification to resolve entity identities with high accuracy. Below is a detailed breakdown of its methodology:

The solution uses multi-attribute identity resolution that employs cross-references to multiple data points to build a comprehensive identity profile, including:

**Core Identifiers:**

1. Full name (including alternative spellings and transliterations)
2. Date of birth (DOB)
3. Nationality and place of birth
4. Government-issued ID numbers (passport, national ID, tax ID)

**Advanced Matching Logic:**

Deterministic Linking: Ensures that a match is only flagged if multiple identifiers align (e.g., name + DOB + passport number).

| Case | Example |
|---|---|
| Simple case (no affixes) | احمد محمود<br>Ahmad Mahmoud |
| Prefix case<br>{عبدAbd, ابو Abou, بن Bin, ال Al, …etc} | عبد العزيز ال سعود<br>Abdulaziz Al Saud |
| Double prefix case<br>{ابو عبدAbou Abd,بن عبد Bin Abd, … etc} | سلطان**بن عبد** العزيز ال سعود<br>Sultan bin Abdulaziz Al Saud |
| Embedded noun case<br>{الدين El-Deen,الله Allah, …etc} | هيردي نور **الدين**<br>Herdi Noor Al-Din |
| Complex name<br>(prefix + embedded noun) | تقي **الدين** محمد **بن** معروف الشامي<br>Taqi al-Din Muhammad Ibn Ma'ruf al-Shami |

**Impact:**

- Reduces false positives by distinguishing between individuals with similar names but different birthdates or IDs.
- Enhances accuracy in cross-border screenings, where naming conventions vary.

# Phonetic Analysis to Minimize False Negatives

AML Watcher's phonetic matching enhances name screening by detecting variations that sound alike but are spelled differently (e.g., "Mohamed" vs. "Muhammad"). It uses algorithms like Soundex and Metaphone to analyze pronunciation across 80+ languages, for **example:**

- Transliteration differences (e.g., "Gaddafi" vs. "Qaddafi")
- Regional naming conventions (e.g., Arabic "Al-" prefixes)
- Common typos and nicknames (e.g., "Bill" for "William")

This reduces false negatives while maintaining precision, helping FIs:

1. Capture sanctioned entities despite spelling variations
2. Cut manual review of minor discrepancies
3. Screen global clients effectively across scripts and dialects

# Using Advanced Fuzzy Matching with Unique Identifiers to Improve Accuracy

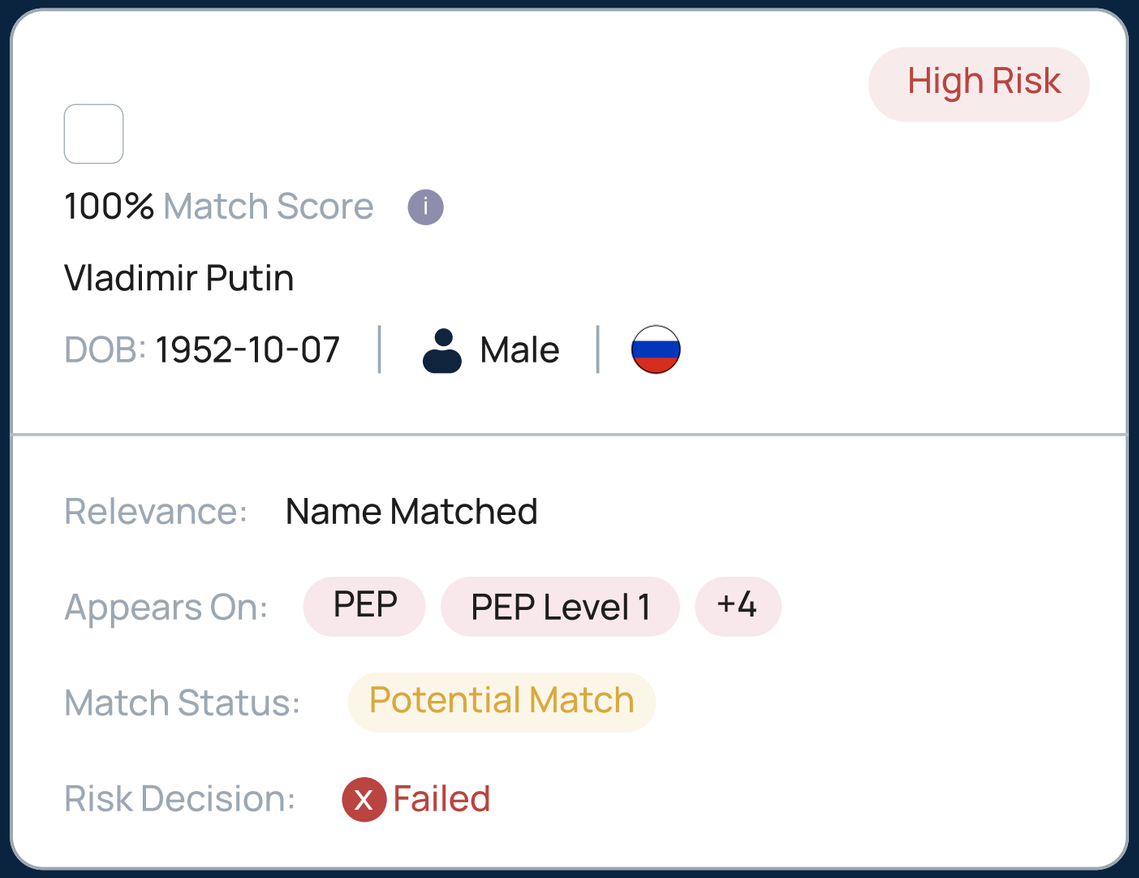
AML Watcher's advanced fuzzy matching technology enables precise name screening across global watchlists, sanctions lists, and PEP databases by:

- Detecting non-exact matches (accounting for typos, aliases, and name variations)
- Supporting multiple matching algorithms (including phonetic and character-based methods)
- Maintaining high precision with unique identifiers while minimzing false positives

**Key benefits for financial institutions include:**

1. Configurable matching thresholds to meet specific risk appetites
2. Reduced manual review workload through intelligent match scoring
3. Comprehensive coverage of name variants across languages and regions

The system's customizable parameters allow compliance teams to balance detection sensitivity with operational efficiency according to their institution's risk profile.



# MENA-Tailored Data Coverage

AML Watcher reduces false positives and false negatives through a balanced approach that brings together smart technology and reliable data.

At the application level, it uses flexible fuzzy matching along with unique identifiers like date of birth, passport number, and national ID to improve the accuracy of name matching.

Behind this, the data layer is constantly refreshed to capture the latest updates in sanctions, watchlists, and PEP statuses. This ensures that organizations are screening against current and compliant data, supporting a risk-based approach that aligns with regulatory expectations.

# ABOUT US

At **AML Watcher,** we aim to support more than 10,000 businesses in their fight against rising FinCrime by creating a secure and compliant financial world where they can thrive.

# CONNECT WITH US:

For more information, visit:


Info@amlwatcher.com

www.amlwatcher.com