

- UserGuide for interpretation and use of the different matching indicators given in the Person resource

This UserGuide is ment to give the users of BankID AML a better understanding and how to interpret and use the different matching-indicators given in relation to “name similarity” and “Date of birth” similarity in the response, and how to interpret and use these responses.

Use-cases for the person resource in a B2C and B2B context

The person resource is helpful to retrieve information about individuals in a B2C context, but is also used in a B2B context, to retrieve information about relevant roles in the organizations (board members, beneficial owners, ceo and further), provided that you already know who they are. The person resource can be successfully combined with the organization resource. In such cases, the merchant will first retrieve role information about the company from the organization, then use the person resource to get information about the provided individuals, often combined with information provided by the customer himself.

Typically, the person resource will be used during or before a transaction or as part of an onboarding process, often combined with acquiring legitimation with BankID. The person resource can also be used in an ad-hoc fashion, when there is reason for suspicion or reason to believe that the information has changed.

A matching process starts with (and demands) a match on names

There are two modes for matching names supported in the person resource; “exact mode” and “fuzzy matching mode”

If the “exact mode” is used, all the names in your query must match the exact spelling of the same names in the list you screen against.

If the “fuzzy matching mode” is active even names that are close to the listed persons names will generate a match.

The higher the “nameMatchValue” is the better the similarity of the name match was.

The screenshot displays the BankID AML search results interface. It is divided into two main sections: a search result for a specific person and a list of matches.

Search Result (Søkeresultat AML):

- Søket er gjennomført av:** 998303168
- Organisasjonsnummer:** IDFY NORGE AS
- Foretaksnavn:** 18.09.2019 13:49:25
- Dato for utført søk:** f29cf111-e5ea-4559-86aa-b120489892
- Referanse:**
- Søkekriterier:**
 - Navn:** KLAVEN-KARI KLAVEN
 - Fødselsnummer:** 01010750241
 - Fuzzy-modus:** Aktivert
- Adresser:**
 - Folkerregistret adresse:**
 - Adresse:** BOKS 6300, ETTERSTAD
 - Postnummer:** 0603
 - Poststed:** Oslo
 - Metadata:** Kilder: NO_NATIONAL_REGISTRY
 - Adresse:**
 - Postnummer:** 0674
 - Poststed:** Oslo
 - Boligtype:** House
 - Antall beboere:** 2

Match mot datakilde (Match with data source):

- Match mot datakilde:** EU_GLOBAL
- Navn:** Abu Abdurrahman the
- Registrerte fødselsdato(e):** 1978-10-19, 1978-10-1
- Aliaser:** Abu Abdurrahman the the Norwegian, Anders Muslim Abu Abdurrah
- Addresser:** Land: Norway NO- Ost
- URL'er til kilder:** http://eur-lex.europa.eu/urn:uj:JOL_2014_293
- Kvalitet på søk:** 165
- Kvalitets beskrivelse:** match on name, year a date of birth).
- Kvalitet på navnetreff:** 0.82959599090909
- Datoleverandør:** TRAPETS
- Ekstern-id:** 7476
- Først oppdatert:** 08.10.2014
- Sist oppdatert:** 03.06.2019
- Match mot datakilde:** UN_CONSOLIDATED
- Navn:** Muslim Abu Abdurrah
- Registrerte fødselsdato(e):** 1978-10-19, 1978-10-1

Understanding of how a match on name is generated.

Exact mode

In the API there is a parameter named: matchMode -> here you choose EXACT or fuzzy.
By default the system are set to “exact mode.

In general this means that all the names in your query must match the exact spelling of the names in the list that you screen against.

However the “exact mode” has built in some margins that will help to create a match based on typical limited knowledge of how the names are build up or listed in the sanction lists

Res. List

United Nations Security Council Consolidated List

Generated on: 8 July 2019

Generated on refers to the date on which the user accessed the list and not the last date of substantive update to the list. Information on the substantive list updates are provided on the Council / Committee's website.

TAI.130 Name: 1: ABDUL GHAFAR 2: QURISHI 3: ABDUL GHANI 4: na

Name (original script): عبدالغفار قريشي

Title: Maulavi **Designation:** Representation Attache, Taliban Embassy, Islamabad, Pakistan **DOB:** a) 1970 b) 196

POB: Turshut village, Wursaj District, Takhar Province, Afghanistan **Good quality a.k.a.:** Abdul Ghaffar Qureshi

Low quality a.k.a.: na **Nationality:** Afghanistan **Passport no:** Afghanistan D 000933, issued on 13 Sep. 1998, issued in KABUL **National identification no:** 55130 (tazkira) **Address:** Khairkhana Section Number 3, Kabul, Afghanistan **Listed on:** 25 Jan. 2001 (amended on 3 Sep. 2003, 18 Jul. 2007, 21 Sep. 2007, 29 Nov. 2011, 31 Dec.

Taken from the UN Sanction list.

The “Exact mode” will therefor include that;

- ❑ The order of the names doesn't matter (“Abdul Ghafar” will match “Ghafar Abdul”)
- ❑ Match against alias names are done. Some persons are known as/with another name than their birth name, or they have nick names etc. These names are often listed as “aliases” in the Sanction list, and a match against one of these names will be displayed just as any other name-match with matchindicator.
- ❑ You don't need all the names of a persons. The system will created a match as long as the names used are correct spelled (“Qurishi Abdul” will match with “Abdul Ghafar Qurishi”). The “NameMatchValue” will then give you an indication of such situation with displaying a value <1.0 in the response

E.g;

Search done with “Qurishi Abdul” in exact mode, against the listed;

1: Abdul Ghafar 2: Qurishi

will give you a “match on name” with this “nameMatchValue”;

"nameMatchValue": 0.95 (See “Fuzzy mode” below for value explanation)

- ❑ Noise words are ignored by the algorithm (Anna-Karin will match with Anna Karin).
These symbols are ignored; * () [] { } ~ . , ' + ? \ " ^ \ \ < >
- ❑ The letters; æ, ø, å can be used, and some rewriting of these letters will also be accepted;
“Bent Hoie” will match with “Bent Høie”
“Monica Maland” will match with “Monica Mæland”
“Helge Andre Njastad” will match with “Helge André Njåstad”
- ❑ The same will apply for the Swedish letters where o will match with ö

When using “exact mode” you will find that there will be no match if you almost have the correct/same name and even have the same Date of birth as a listed person.

“Shafi Sultan Mohammad 01.01.1973” will not match at all with;
 “Shafi Sultan Mohammed 01.01.1973”

Fuzzy mode

To ensure that small typos, like mentioned above, do not omit hits against listed units, you need to activate “fuzzy match mode” Mark that this mode will create more false positives.

We do not recommend to use this mode if you only have one name in your query and even two names will create several hits if one of the names are similar to the listed person.

How to activate and use Fuzzy mode

In the API there is a parameter named: matchMode -> here you choose “fuzzy”

When “fuzzy” is enabled it is recommended that you supply more than two (always more than one) keyword (name part) and always combine it with Date Of Birth to prevent that the number of false positives becomes overwhelming.

The response will contain a “nameMatchValue” in which ranges from 0.0 to 1.0. The higher the value is, the better the name similarity for the names in the request and the names on the listed person was. Some examples;

Query examples			“nameMatchValue”	Comment
Jonas	Gahr	Støre	1.00	exact match on all the names in your query
Jonas	Gar	Støre	0,94	
	Gahr	Støre	0,95	In “exact modus” this will also give a match
Jonas	G	Støre	0,93	In “exact modus” this will not give a match because "G" don't match with "Gahr"
Jons	Gar	Støre	0,87	
John		Støre	<0,6	No match even in fuzzy modus

The calculation of the “nameMatchValue” is done in several steps:

1. Distance calculations are made for each searched name part and the closest corresponding name part in the list item. Distance is calculated using the [Levenshtein distance algorithm](#).
2. The distances for each name part are weighted by the number of letters in each part. The distance for a longer name part has greater impact on the total distance than the distance for a shorter name part.
3. The total distance is reduced with a penalty of 5% if the list item contains one more part than the searched number of name parts.
4. The total distance is reduced further with a penalty of 3% if the list item contains two more parts than the searched number of name parts.
5. The total distance is reduced further with a penalty of 2% if the list item contains three more parts than the searched number of name parts.

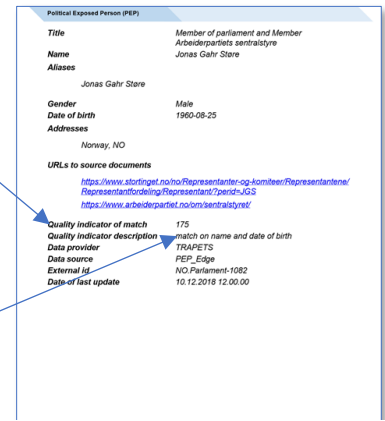
Note that all matches with nameMatchValue with value < 0.6 will be ignored (unless there is a match by national id) and not shown as a match at all.

If the name (including Alias names), and only the name, matches with a listed person (in exact or in fuzzy mode) you will have a match with a “matchindicator DoB match” = 125 (Match on name only).

Further; The similarity of the persons “Date of birth” gives you a possibility to have a stronger and more secure “matchindicator DoB match” an then a higher value than 125.

How to understand and use the “Matchindicator DoB match”.

A screening results with match on name can very well be a false match; it is not unlikely that some of your business relations or customers has the same or closely resembling name as a listed person. A high matchindicator is a stronger indication that you have a positive match (not a false positive) and together with other different inputs; like the persons address information, birthplace and further, a match can be confirmed or disproved



The API and the PDF-reports will deliver indicators telling you how strong the match against the sanction and/or pep list is. The Match on DoB is delivered with a number and a short description explaining the quality of the match.

The different Matchindicators for DoB match are;

Matchindicator DoB match	Quality Description
125	match on name only
140	Not in use yet – related to match against a range of Birth years
145	Not in use yet – related to match against a range of Dates of birth
155	match on name and year of birth
165	match on name, year and month of birth (not complete date of birth)
175	match on name and date of birth (complete date of birth)
500	match on national ID number

Table 1.1

For a deeper understanding on how to use this information we have some examples, and remarked situation to be aware of, listed in tabel 1.2 below.

Match indicator	Quality description, and examples
125	<p>Match on name only.</p> <p>Remark that both of the two examples below gives MI = 125;</p> <p>A) Sanction listing; Son Jong Hyok, DOB: 20 May 1980 Search data; Son Jong Hyok, DOB: 01 January 1988</p> <p>B) Sanction listing; Abdul Manan Agha, DOB: na Search data; Abdul Mana Agha, DOB: 01 january 1988</p> <p>In example A, a person who has a registered date of birth in the sanction list, and this date is different from the date on which is applied: This is very likely to be a false positive. In example B, the person is listed without a date of birth in the sanction list and it is therefore difficult to confirm/deny that this is a positive of false positive match. For some users the example B could entail the need for further investigations while example A could be described as a confirmed false positive.</p> <p>Important; Pay attention to the info given in the field «birthDateInfo» (delivered in the API and in the PDF report). If this field is empty it means that the person is listed without any date of birth in the list. If the person is listed with one or several dates of birth these will all be returned in the field “birthDateInfo”</p>
140 Not implemented	Match on name, and the year of birth is within a <u>range</u> of several listed “year of birth” in the sanction list.
145 Not implemented	Match on name, and complete date of birth is within a <u>range</u> of several listed “Dates of birth” in the sanction list.
155	<p>Match on name and year of birth</p> <p>Remark that all 3 examples below gives MI = 155;</p>

Match indicator	Quality description, and examples
	<p>A) Sanction listing; Son Jong Hyok, DOB: 20 May 1980 Search data; Son Jong Hyok, DOB: 20 January 1980</p> <p>B) Sanction listing; Adil Abdallah, DOB: 1945 Search data; Adil Abdallah, DOB: 01 January 1945</p> <p>C) Sanction listing; Tayeb Nail, DOB: Approximately 1972 Search data; Tayeb Nail, DOB: 01 January 1972</p> <p>Example A, a person who has a registered date of birth in the sanction list, but this date is different from the date (ddmm) applied for, but the year is the same. This is likely to be a confirmed false positive. There fact that the listed date (dd) is the same as in the query does not give any degree of an higher match than 155.</p> <p>Example B, the person is listed with only the year of birth in the sanction list and it is therefore not possible to confirm/deny that this is a positive or false positive match only on the basis of the DOB in the search.</p> <p>Example C, the person is listed only with a year of birth that <u>estimates</u> the person's age (the entry can also be like this; "ca 1972"). In such cases, only searches with the given year mentioned in the sanction list will make a 155 indicator. Searching with a nearby year (eg 1971) will give a 125 result. In JSON and in the report it is only the year itself, and not the text "approximately" that is displayed, but work is being done to include this info as well.</p> <p>All 3 examples may entail the need for further investigations to confirm / deny the match, but example A is a strong indication of a false positive.</p> <p>Important; Pay attention to the info given in the field «birthDateInfo» (delivered in the API and in the reports). If this field only contain a year it means that the person is listed without a full date of birth in the list. If the person is listed with several years of birth these will all be returned in the field "birthDateInfo"</p>

Match indicator	Quality description, and examples
165	<p>Match on name, year and month of birth (not complete date of birth)</p> <p>The following example gives MI = 165;</p> <p>A) Sanction listing; Malik Noorzai, DOB: a) 1957 b) 1960 c) 1 jan 1963 Search data; Malik Noorzai, DOB: 08 jan 1963</p> <p>A person can be listed with multiple birthdates and with different degrees of completeness. Your search is made against all entries of the DOB in the list and the hit that gives you the highest math indicator will be the display (but you will find all listed dates in the field “birthDateInfo”). The actual order of the entries gives no indication itself as to which date is considered to be correct. The a) entry is not assumed to be more correct than the c) entry.</p> <p>If we search with Malik Noorzai, DOB: 08 jan 1960 we would have received a 155 hit since the b) entry in the sanction list gives us hits on name and year.</p>
175	<p>Match on name, year, month and date of birth (complete date of birth)</p> <p>The following example gives MI = 175;</p> <p>A) Sanction listing; Malik Noorzai, DOB: a) 1957 b) 1960 c) 1 Jan 1963 Search data; Malik Noorzai, DOB: 1 Jan 1963</p> <p>MI = 175 requires that the person you are searching for is listed with at least one complete date of birth. Both the name and the complete date of birth match the entry. Note that there is no guarantee that the dates in the sanction list are correct. A person who has listed several complete birth dates in the sanction list (in PEP_edge there will only be one date of birth) will generate a 175 hit as long as you hit one of those dates (although it is then obvious that these listed dates are of uncertain nature)</p> <p>Important; Pay attention to the info given in the field «birthDateInfo» (delivered in the API and in the reports) even for a 175 match against sanction lists. If this field only contain one full date of birth you will in general have a stronger match than if the field contains several complete birth dates. If the person is listed with several Dates of birth these will all be returned in the field “birthDateInfo” The ordre of the listed dates does not</p>

Match indicator	Quality description, and examples
	indicate any assumptions about which one is the correct date.
500	match on national ID number

Important last remarks

Be aware of the quality difference in the listings in [sanctions list](#) and the [pep list](#).

In general, a hit against sanction lists requires some more attention around the background of the match indicators than a hit against the pep_edge list.

NOTE that [sanction](#) matches often will be a match on name only (because of lack of Birth date information in the list screened against), and in those cases where you actually have a hit against elements in the date of birth (MI higher than 125) do always include the info displayed in the field "birthdateInfo" in your assessment. For [PEP-matches](#) a match is almost always a match against both name and complete date of birth (since the pep list used in BankID AML are close to 100% covered by date of birth for the listed persons) and there is only one registered date of birth per person.

Risk assessment – “The mother of all obligations”

The BankID AML team strongly recommend that your company document which considerations that have been made for using “exact mode” or “fuzzy mode” and how your workflow and use of the matchindicators and “nameMatchValue” correlates with your company`s risk assessment and your risk-based approach.