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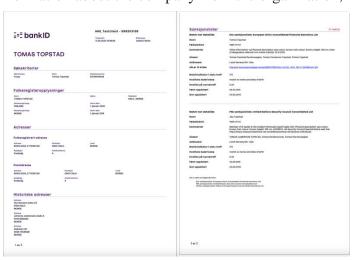
This UserGuide is ment to give the users of BankID AML a better understanding and how to interpret and use the different matching-indicators given in relation to "name similarity" and "Date of birth" similarity in the response, and how to interpret and use these responses.

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
Matching names against PEP and Sanctions lists is done through a "fuzzy matching mode" in the API.

The search result will contain an indicator "NameMatchValue" in which ranges from 0.0 to 1.0. The higher the value is, the better the name similarity for the name in the request and the name in the lists was. The customer can set a limit to how low the NameMatchValue of a search result may be through usage of the request parameter "nameMatchThreshold".

In general, a "nameMatchThreshold=1.0" means that all the names in your query must match the exact spelling of names in in the list that you screen against. When using "nameMatchThreshold=1.0" you will Find that there will be no match even if you have the same Date of birth as a listed person and you have almost the correct/same name.

Understanding of how a <u>match on name</u> is generated

Summary

By default the system is set to "fuzzy matching mode". The search result will contain an indicator "NameMatchValue" in which ranges from 0.0 to 1.0. The higher the value is, the better the name similarity for the name in the request and the name in the lists was. The customer can set a limit to how low the NameMatchValue of a search result may be through usage of the request parameter "nameMatchThreshold".

The default "nameMatchThreshold" is set to 0.85. This threshold can (and should) be adjusted according to the end user's own wish, based on each user's internal routines.

More about "nameMatchThreshold=1.0"

In general, a "nameMatchThreshold=1.0" means that all the names in your query must match the exact spelling of the names in the list that you screen against.

However, there are built in some margins that will help to create a match based on typical limited knowledge of how the names are built up or listed in the sanction lists, even if the threshold is set to 1.0**. A search with "nameMatchThreshold=1.0" will therefore 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 the 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 person. The system will create a match as long as the names used are correctly spelled ("Qurishi Abdul" will match with "Abdul Ghafar Qurishi").

The "NameMatchValue" will then give you an indication of such a situation with displaying a value <1.0 in the response

> E.g.:
> Search done with "Qurishi Abdul" in "nameMatchThreshold=1.0", 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 Andrè Njåstad"

☐ The same will apply for the Swedish letters where o will match with ö

When using "nameMatchThreshold=1.0" 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. For example:
"Shafi Sultan Mohamm<u>a</u>d 01.01.1973" will not match at all with "Shafi Sultan Mohamm<u>e</u>d 01.01.1973".

How to use "fuzzy matching mode"

To ensure that small typos, like mentioned above, do not omit hits against listed units, you need to set the "nameMatchThreshold" a bit lower than 1.0. By default the system is set to 0.85, but this can be changed according the end user's own wish. Please note that this mode will create more false positives, but might be necessary in order to not omit true positives.

When the threshold is lower than 1.0, it is recommended that you supply more than two keywords (name part) and always combine it with a 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	With "nameMatchThreshold=1.0" this will not be a match.
	Gahr	Støre	0,95	With "nameMatchThreshold=1.0" this will be a match
Jonas	G	Støre	0,93	With "nameMatchThreshold=1.0" this will not be a match, because "G" doesn't match with "Gahr"
Jons	Gar	Støre	0,87	With "nameMatchThreshold<=0.87" this will be a match with the shown "nameMatchValue" = 0,87
John		Støre	<0,6	Match with "nameMatchThreshold<=0.6", but low nameMatchValue (below 0,6) ignores the hit = not displayed

Response with Value <0,6 will be ignored (not displayed) unless there is a complete match on national ID number.

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. Our experience is also that values between 0.6 – 0.7 also includes matches with names too far away, so consider where to put your cutoff when using fuzzy mode, according to your own testing, procedures and risk appetite.

If the name (including Alias names), and only the name, matches with a listed person you will have a match with a "match indicator DoB match" = 125 (Match on name only).
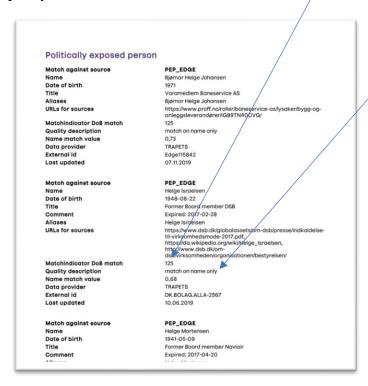
Furthermore: The similarity of the person's "Date of birth" gives you a possibility to have a stronger and more secure "match indicator DoB match" and then a higher value than 125.

How to understand and use the "Match indicator DoB match".

A screening results with match on name only 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 match indicator is a stronger indication that you have a positive match (not a false positive), and together with other different inputs: Like the person's 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 Date of Birth is delivered with a number and a short <u>description</u> explaining the quality of the match:

The different Match indicators for DoB matches are:

Matchindicator DoB match	Quality Description [Table 1.1]
125	match on name only
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

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

Match indicator	Quality description, and examples [Table 1.2]
125	**Match on name only.** Remark that both of the two examples below gives MI = 125; A) Sanction listing; Son Jong Hyok, DOB: 20 May 1980 Search data; Son Jong Hyok, DOB: 01 January 1988 B) Sanction listing; Abdul Manan Agha, DOB: N/A Search data; Abdul Mana Agha, DOB: 01 january 1988 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. 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".
155	**Match on name and year of birth** Remark that all 3 examples below gives MI = 155; A) Sanction listing; Son Jong Hyok, DOB: 20 May 1980 Search data; Son Jong Hyok, DOB: 20 January 1980

Match indicator	Quality description, and examples [Table 1.2]
	B) Sanction listing; Adil Abdallah, DOB: 1945 Search data; Adil Abdallah, DOB: 01 January 1945 C) Sanction listing; Tayeb Nail, DOB: Approximately 1972 Search data; Tayeb Nail, DOB: 01 January 1972 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 a higher match than 155. 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. 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 searches with the given year, and +/- 3 years from the year mentioned in the sanction list, will make a 155 indicator. In JSON and in the report it is only the year itself, and not the text "approximately" that is displayed. 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. 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"
165	**Match on name, year and month of birth (not complete date of birth)** The following example gives MI = 165: A) Sanction listing; Malik Noorzai, DOB: a) 1957 b) 1960 c) 1 jan 1963 Search data; Malik Noorzai, DOB: 08 jan 1963

Match indicator	Quality description, and examples [Table 1.2]
	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 match indicator will be displayed (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. 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.
175	**Match on name, year, month and date of birth (complete date of birth)** The following example gives MI = 175: A) Sanction listing; Malik Noorzai, DOB: a) 1957 b) 1960 c) 1 Jan 1963 Search data; Malik Noorzai, DOB: 1 Jan 1963 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). 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 order of the listed dates does not indicate any assumptions about which one is the correct date.
500	**match on national ID number** A match on national ID number is a strong indicator of a true positive.

Important last remarks

Be aware of the quality differences in the listings in **sanctions list** and the Nordic **pep_edge list**. In general, a hit against sanction lists requires more attention around the background of the match indicators than a hit against the pep_edge list, since the quality of the pep_edge list is considered to be generally better.

NOTE that **sanction** matches often will be a match on name only (because of lack of Birth date information in the list screened against). In cases where you actually have a hit against elements in the date of birth (MI higher that 125), do always include the info displayed in the field "birthdateInfo" in your assessment (it can contain more than one registered birthdates). For **PEP-matches** a match is almost always a match against both name and a complete date of birth (since the pep_edge list are close to 100% covered by complete date of birth for the listed persons), and there is only <u>one</u> registered date of birth per person.

Risk assessment – "The mother of all obligations"
We strongly recommend that you properly document which considerations that have been made for your specific use of the "fuzzy" search mode this tool provides, and how your workflow and use of the matchindicators and "nameMatchValue" correlates with your company`s risk assessment and your risk-based approach. Finanstilsynet also expects you to have internal routines and procedures for the use of any support tools such as this one, and to make sure that the tools used are adapted to the nature and size of your business, cf. Finanstilsynet's "veileder til hvitvaskingsloven- rundskriv 8/2019".