Pensions Dashboards Working Group

Data Matching Convention (DMC) Guidance

December 2021

Produced in partnership with itm PASA Experts for Dashboards

Contents

Terminology and Abbreviations

See the Pensions Dashboards Programme (PDP) Glossary for definitions of dashboards terminology. Abbreviations used in this Guidance:

DOB	Date of Birth
GDPR	General Data Protection Regulations
ICO	Information Commissioner's Office
ISP	Integrated Service Provider
NINO	National Insurance Number

1 Acknowledgement

The Guidance has been carefully written with representation and inputs from all pension's professional disciplines. PASA is grateful to the authors of the Guidance and their employers for their active participation in the Pensions Dashboards Working Group.

To read a biography of authors, click their name.

Members:

Kim Gubler (Board Sponsor)	Chris Connelly (Chair)	Richard Smith	Chris Batts
Geraldine Brassett	Phillip Cork	Ian Dack	Rob Dodson
Nick Green	Karl Lidgley	Ian McQuade	Climate Mupfiga
Francesca Parnell	Sonya Purkayastha	Simon Rawson	Amy Regler
David Rich	Ranjit Shoker	Andrew Short	Angela Bell (TPR Observer)

If you have any questions about this guidance, please contact PASA at:
info@pasa-uk.com

This Guidance has kindly been designed by EQ

2 Executive summary

Schemes must decide how to match:

From 2023 onwards, on their staging date all pension schemes and providers must comply with new legislative duties for pensions dashboards. As part of this, schemes and providers need to soon decide how they want to digitally compare and match 'find requests' from users of dashboards against all the records they hold. The detail of the new legislative duties will be published for consultation in early 2022, but you can start to consider the importance of your matching and data quality now.

Guidance to help:

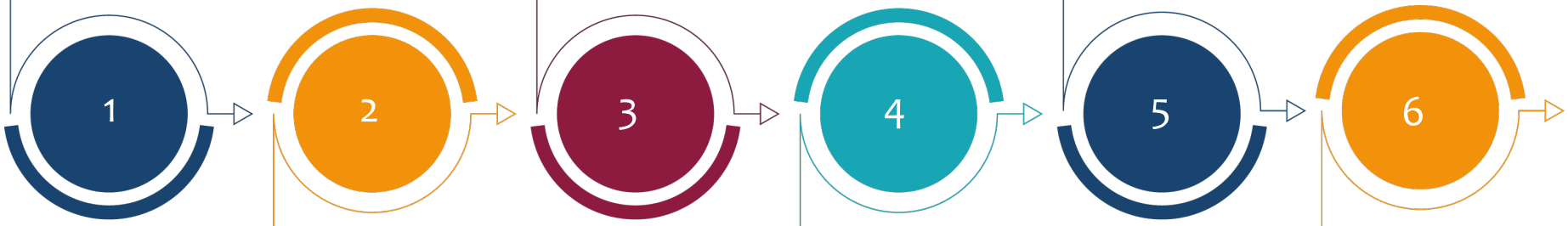
This industry-wide Data Matching Convention (DMC) Guidance is intended to help schemes with making these matching decisions, rather than instructing schemes how they should match.

Data accuracy is key:

Schemes should be working towards having a high level of confidence in the accuracy of Surnames, DOBs and NINOs for deferred and active members. This will involve continual, systematic checking of the accuracy of actual values held in these data elements.

Third parties control data:

Administrators understand schemes and providers have never been able to fully control the accuracy of personal data. Third parties (such as employers, deferred members, previous administrators, etc.) are responsible for data accuracy, both initially on joining and over the whole lifetime of individuals' pension entitlements.



Three core data elements:

Based on current practice, many (but not all) schemes will decide to match on Surname, Date of Birth (DOB) and National Insurance Number (NINO). Some schemes may add the fourth data element of Forename to the three core data elements.

Accurate data enables simple matching:

Only where schemes believe they are successfully validating the accuracy of all Surnames, DOBs and NINOs, will 'simple' matching on these three core data elements produce robust positive matches. In this Guidance, this is referred to as Option 1.

Data assessment and improvement:

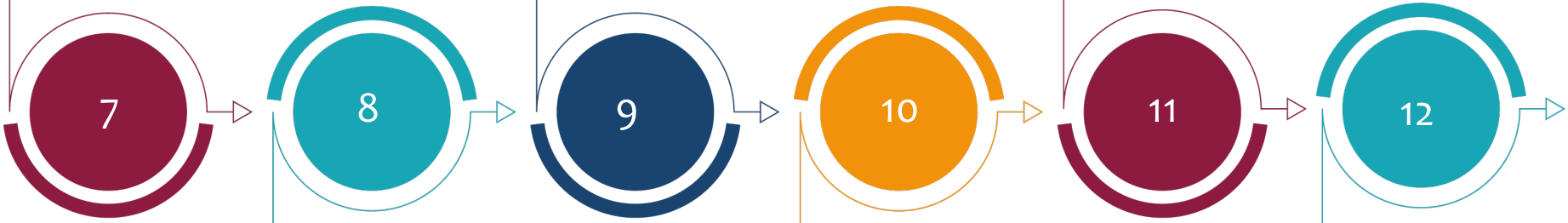
Schemes work hard to continually assess and improve, as far as they can, the accuracy of data such as Surnames, DOBs and NINOs. Many schemes are on a journey to high confidence in this data, but given the reliance on third parties, full accuracy may never be possible.

2 Executive summary

Risks of simple matching:
Where a scheme has yet to reach a high-level of confidence in their personal data, there are risks of choosing simple digital matching on Surname, DOB and NINO. These risks will negatively impact members, administrators, trustees, and others.

More sophisticated matching options:
Based on points 7 and 8, it's likely schemes still on their journey to high confidence in data accuracy will wish to choose more sophisticated matching approaches.

Deciding how to match depends on your scheme's data accuracy:
The key point for schemes to understand is their decision(s) on how to match must be made in the context of their specific scheme's known data accuracy (particularly Surnames, DOBs and NINOs across all deferred and active records).



Competing legislative risks:
Schemes will have to assess their risks against the twin objectives of satisfying the ICO's GDPR requirements as well as complying with their new statutory duties to carry out matching under the forthcoming pensions dashboards legislation.

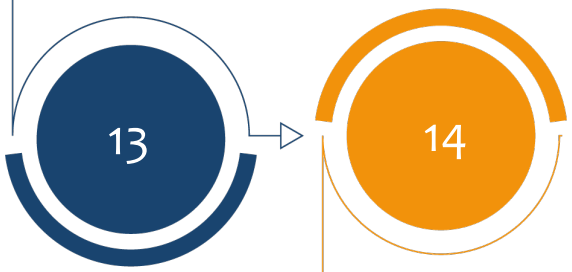
Matching rules and additional data elements:
This Guidance sets out examples of more sophisticated matching, specifically:

- Option 2/Annex A – 'maybe matching' rules where some of Surname, DOB and NINO nearly match
- Option 3/Annex B - using other data elements such as Address Line 1 and Postcode.

Collaboration is key:
Schemes need to engage in detail on this topic soon. They should work collaboratively with their administrators, technology providers and wider data specialists.

2 Executive summary

Supplemental Guidance:
To support schemes' immediate preparations, this first iteration of DMC Guidance is being published in early 2022, but Supplemental Guidance may follow in due course, reflecting further learnings from any testing, or further clarifications.



Additional Topics for more complex arrangements:
After this first iteration, there may be other topics to consider. One emerging topic is whether schemes with benefits administered by multiple parties will need to consider which party is sending data to the requesting dashboard. A common example being where a DB scheme has separately administered AVCs. Schemes will need to consider whether all administrators for a scheme's overall benefits are following the same matching guidance and holding the same matching data.

We are grateful for the help and expertise other organisations have given us to ensure we cover the broadest possible spectrum of industry views.

Our thanks go to the PLSA, the ABI, administration software providers and our regulators: TPR and FCA.

Chris Connelly
PASA Working Group Chair

3 Introduction

As data controllers, pension schemes and providers are required to make their members'/customers' pension data available.

Whenever an individual logs in to their chosen pensions dashboard and makes a request to find their pensions, all schemes and providers must compare and match the user's personal details and return their pension information for them to view.

The core objective is to enable individuals to find their pensions, through schemes and providers making as many definitively positive matches as they can.

What is the DMC Guidance and who is it for?

1. This PASA DMC Guidance is intended for a wider audience than PASA's normal core audience of administration service providers. When the Guidance refers to 'schemes', this should be read as meaning all 'pension arrangements' falling under the dashboards compulsion legislation.
2. This Guidance, produced in conjunction with the PLSA and ABI, is intended to help pension schemes determine how to compare and match the 'find requests' from dashboard users to the records they hold. See the Background section for more details.
3. The Guidance considers existing data matching practices across the industry. Then the core Guidance (and Annexes) suggest some specific ways schemes could choose to match.
4. The Guidance contains practical examples and highlights unusual categories of pension scheme members who may require special consideration in terms of how their data is matched.
5. As data controllers, the decision on how to match remains with schemes. This DMC Guidance provides suggestions and examples to help with this decision.

3 Introduction

Why has PASA produced this Guidance?

Dashboards are universal, spanning the entire pension industry.

Rather than every scheme grappling with this challenge on their own, PASA has produced this DMC Guidance to bring together industry practice and support schemes' decision-making process. Applying industry practice will also result in cost efficiencies.

We have brought together the combined expertise of PASA, PLSA, ABI, and the administration software providers. We are also grateful for the support we have received from both regulators, i.e. TPR and FCA.

What should schemes do?

Schemes should review the various matching options set out in this Guidance, including the risks associated with each option, and consider the most appropriate option for them. Working very closely with their administration/technology provider(s), schemes should ask questions such as:

- **How accurate are our Surnames, DOBs and NINOs on our deferred and active records?**
- **Are they accurate enough to adopt the 'simple' Option 1 matching? ***
- **How significant are our risks if we adopt Option 1 and our data isn't accurate? (For example, in terms of potential data breaches, member frustration, increased administration demand, increased costs)**
- **If our Surnames, DOBs and NINOs aren't accurate enough now, what can we do to fix them? ***
- **If we include Forename in the matching, how would this alter the level of risk?**

*** Administrators' answers to these questions will depend on the depth of analysis done on the accuracy of the scheme's data.**

3 Introduction

Wider PASA Guidance on assessing and managing data accuracy

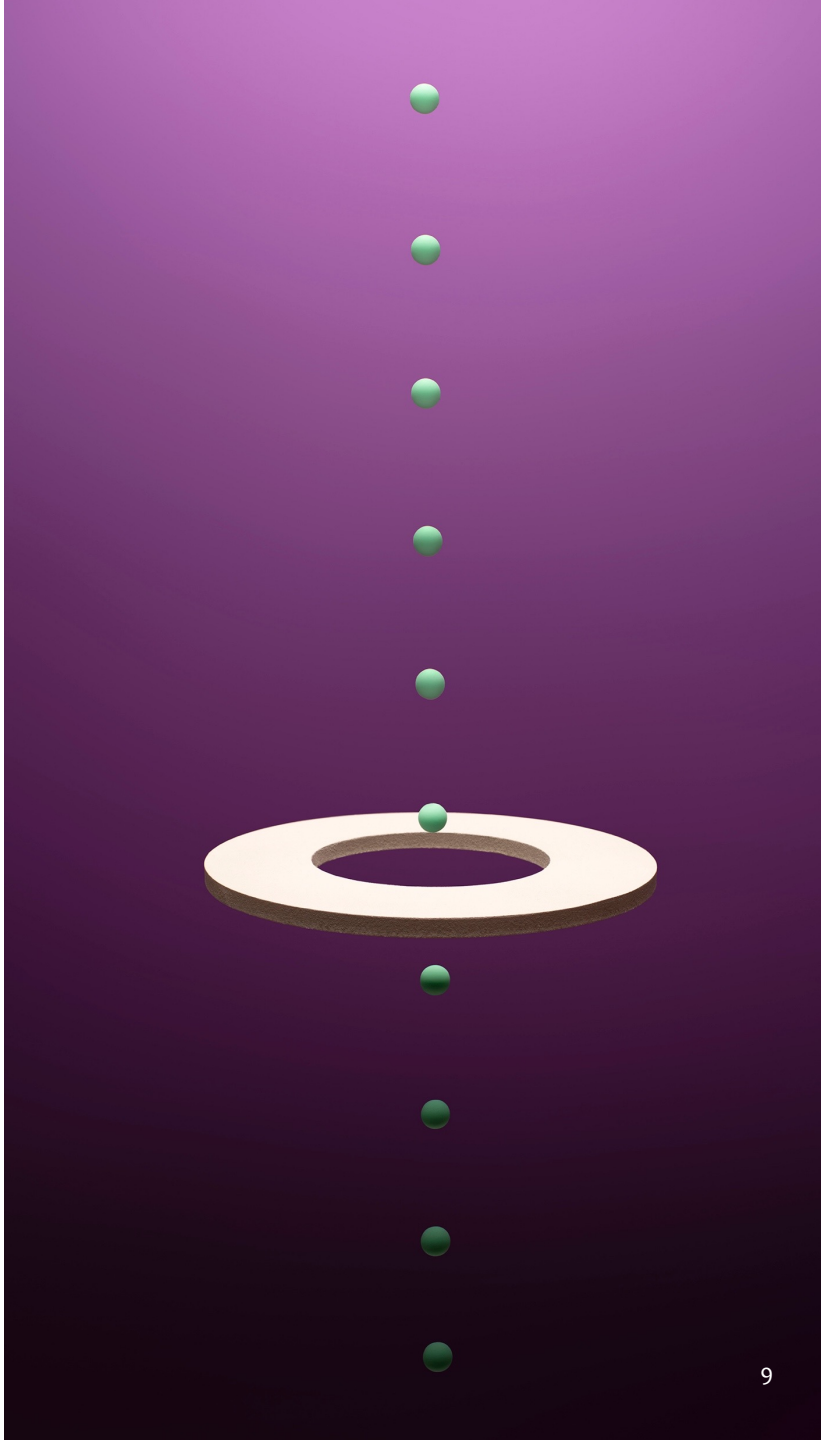
Working collaboratively with their administrator, technology providers and wider data specialist providers, schemes should now be taking steps to analyse, and improve, the accuracy of their Surnames, DOBs and NINOs on their deferred and active records. To support this work and for more information, see the Guidance published by the PASA Data Working Group: https://www.pasa-uk.com/guidance/data/.

Will there be further Guidance?

This Guidance, published in December 2021, will help schemes make initial progress with their dashboards preparations by deciding how to match and improve data accuracy.

As we learn more, supplemental guidance will be published, for example following beta testing of the dashboards ecosystem.

A very important topic is where some, but not all, of the personal data elements match, meaning the dashboard user ***may*** have a pension with the scheme, but it's not certain because it's not a definitive positive match. In October 2020, PDP reported in its summary of Responses to the Call for Input on data standards many respondents specifically requested guidance in this area. We call these 'maybe matches'. During beta testing, we'll start to learn how real data flows through dashboards and we'll share what we discover.



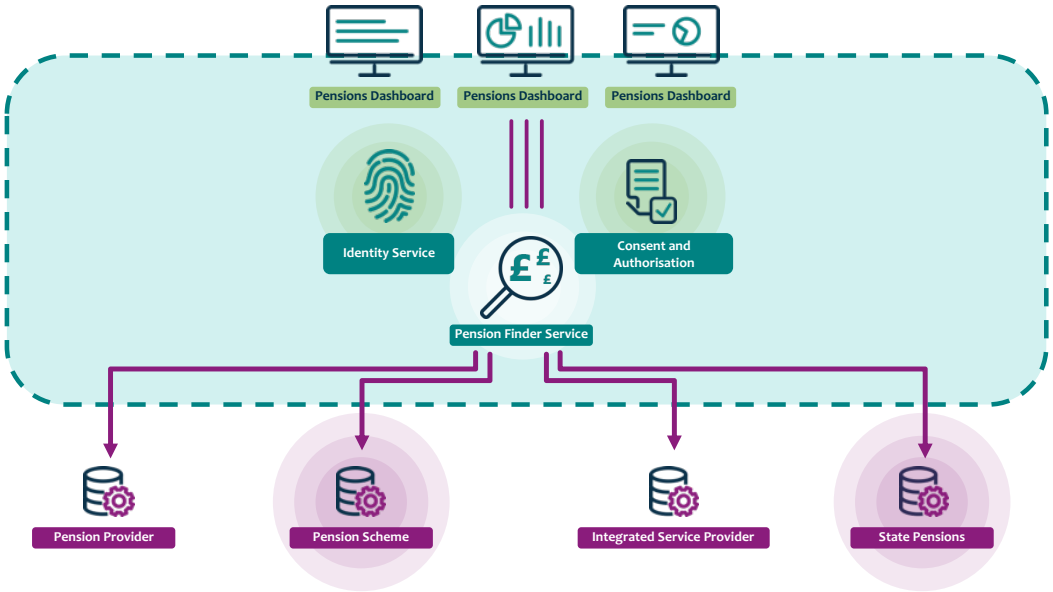
4 Background

In the pensions dashboards ecosystem, thousands of 'find requests' will be sent from the central Pension Finder Service (PFS) every day to all connected data providers.

Data providers will compare personal data elements in each incoming find request against the personal data elements they hold across their records: see the diagram below, courtesy of PDP.

Data providers have indicated they would welcome guidance on how they should make these data comparisons.

For example: what personal data elements should they compare against? And what should they do if only some, but not all, of the data elements they hold match?



Importance of industry-wide applicability

Pensions dashboards are aiming to help savers reconnect with their pensions, but this will only happen if a saver's pensions can be found digitally.

From a saver's perspective, they could have pensions held with several, or all, of the different types of organisation which administer pensions. These include: DWP for state pension, public service pension schemes, private sector schemes (both in-house and third-party administered), master trusts, and various different types of commercial pension and buy-out providers.

Working with their administrator(s)/integrated service provider(s) (see the PDP Glossary for an explanation), all these different types of pension arrangements need to determine the particular combination of personal data elements to be compared and matched (e.g. Surname, DOB, NINO, etc).

If every scheme did this independently, aside from being highly inefficient for industry, it would result in a very poor experience for dashboard users. For example, a user's details might match with one scheme but not another. This difference could confuse users raising questions such as: 'Is one of my schemes less trustworthy than the other? Is the system itself trustworthy?'

A consistent approach across different pension arrangements will deliver a consistent user experience.

5 Existing practice

In putting together this DMC Guidance for pensions dashboards, we first looked at existing data matching practices across the industry.

In current practice, schemes often answer two questions simultaneously:

> **Can I positively match the enquirer to a pension record we hold ?**

> **Has the enquirer adequately proven their identity to us ?**

When a saver telephones an administrator with a question about their pension, administrators ask for information when locating a record. Some information is geared towards finding their record, whereas other information is about proving the caller is who they say they are.

In the context of dashboards, we're only concerned with finding the record. Identity verification will already have been dealt with by the central digital Identity Service within the central dashboards digital architecture. Scheme data controllers can be comfortable the personal details for dashboard users passed down to them have been verified as belonging to the individual in question.

What data elements do schemes/administrators/master trusts use today to find records?

Whether by telephone, correspondence or automated matching, most schemes check the following data elements:

Name

- Some schemes match on just Surname, but most match on First name and Surname.
- On a telephone enquiry, any question about the precise spelling of the First name and/or Surname can be resolved there and then between the service centre operator and the member. When dealing with incoming correspondence, there is manual intervention where any minor spelling issues can be intelligently investigated and resolved.
- In an automated-matching context, there is no manual intervention. Spellings are likely to need to be 100% accurate to ensure correct matches.

DOB

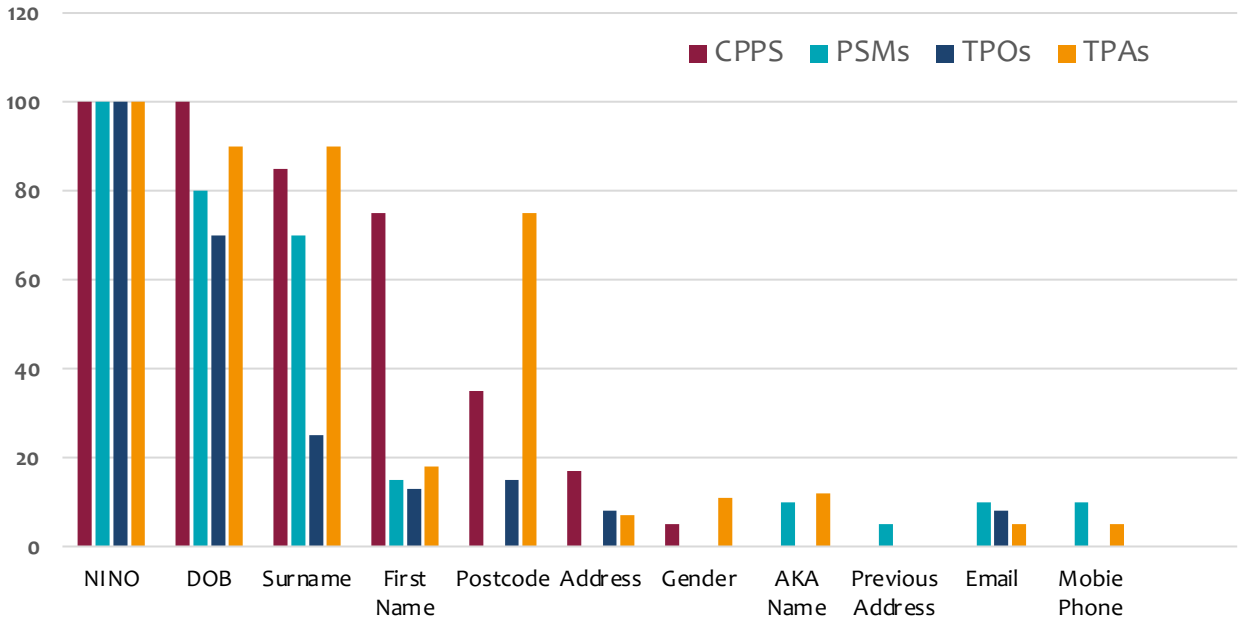
- All parts of the date must match: DD, MM and YYYY.

NINO

- At least the first eight characters of NINO must match.

5 Existing practice

Fields used for identity matching (by organisation type)



Note: Today, many schemes also ask the member for their home postal address, but this may not (and often doesn't) match the address held on the scheme's records, although the individual's previous address may match what's held.

In this graph of responses, note the lower prevalence of First Name (see Annex C for more details). Key to respondent types:

CPPs	**Commercial Pension Providers (14 responses)**
PSMs	**Pension Schemes and Master trusts (11 responses)**
TPAs	**Third party Pension Administrators (10 responses)**
TPOs	**Technology Providers/bodies and Others (10 responses)**

This table reflects responses to the PDP Call for Input on data standards which asked:

> **Which data items do you anticipate could be used to definitively match individuals to their pension entitlements?**

6 Core guidance

Schemes must decide which data elements they wish to compare against the pension records they hold. Their first step should be to fully understand what data elements will be passed to all schemes from the PFS. This is known as the 'Find data' described in the PDP Data standards guide (DSG). The PDP DSG was published in December 2020. For the most up-to-date version go to PDP's Data providers hub.

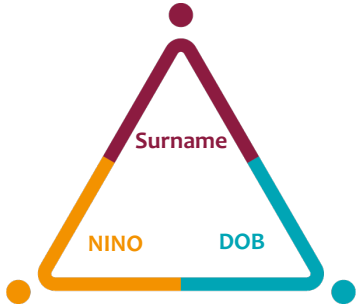
Primary data elements in the Find data to be used for matching

Many schemes will wish to match these few key data elements in the Find data, described in the PDP DSG (note: the DSG explicitly excludes any middle names from Data element 1.001):

Reference	Data element	Description	Optionality
1.001	Given name	Given name/Forename	Mandatory
1.002	Name	Surname of the individual	Mandatory
1.003	Date of birth	Date of birth of the individual	Mandatory
1.004	NI number	National Insurance number of the individual	Mandatory

Three core data elements

This core Guidance proposes the three core data elements for matching are Surname, DOB and NINO. (As set out in Annex C, Forename is unlikely to strengthen, and could reduce, confidence in already strong matches on Surname, DOB and NINO).



It will be **up to individual schemes to decide whether they wish to add Forename to these three core data elements,** based on the Forename/Initials data they hold, and their views on whether this would increase confidence in strong matches without reducing the positive matches they make.

Data accuracy of the three core data elements and real world administration challenges

Schemes should ideally already have a high level of confidence in the accuracy of data held in the three key data elements of Surname, DOB and NINO. However, this is a challenge, as identified in PDP's schemes research: 'Data accuracy relies on the third parties providing the information in the first place and keeping it up to date'. These third-parties include **employers (for active members), deferred members themselves, previous administrators,** and others.

Where schemes have a **high level of confidence** in their Surnames, DOBs and NINOs, across all their deferred and actives, they can be confident choosing 'simple' matching Option 1 as described below.

However, for schemes which **do not yet have completely accurate data in the three core data elements,** there are risks associated with matching Option 1 (see the table below).

Call to action on data accuracy

See Section 10 for practical help from the PASA Data Working Group, on steps that schemes can take to get to the position where they are confident to choose the simple Option 1.

6 Core guidance

Matching Option 1 – Surname, DOB and NINO

Where a pension should be found, but isn't, due to a non-perfect match on all of Surname, DOB and NINO, trustees will not be meeting their dashboards and legislative duties.
In this scenario, trustees will wish to understand to what extent they are liable for pensions not being found – it may be possible to publish supplemental Guidance on this in due course when we better understand the dashboards' liability model.

			Impacts from different perspectives of choosing Option 1			
Option	**Description**	**Confidence in your scheme's personal data accuracy**	**Individual's Experience of this option**	**ISP Technical Operation of this option**	**Administration Implications of this option**	**Trustee Risk of this option**
1	If Surname, DOB and NINO all match perfectly, then do a positive return, otherwise do nothing.	**High confidence** in the accuracy of Surname, DOB and NINO on all deferred and active records.	Pensions are found for the individual to view on their chosen dashboard.	Simple match regime to implement and operate.	High number of found pensions, increased demand for quotations.	High likelihood of correct matching.
		Less than high confidence in the accuracy of Surname, DOB and NINO on all deferred and active records.	Some pensions won't be found, so the individual won't see them on their chosen dashboard.	Simple match regime to implement and operate.	'Failure demand' from individuals not seeing pensions they are expecting to see.	Higher likelihood of false negatives, i.e. not discharging dashboards duties.

6 Core guidance

There are two very common situations where the data provided by the dashboard user may not match with the data held by schemes: Ninth character on NINO does not match, and Surname does not match because the Surname held by the scheme is the user's previous Surname. An enhanced version of Matching Option 1 could largely mitigate against these two common issues if it is defined as follows:

Enhanced Matching Option 1 – Surname, DOB, NINO (chars 1-8) and Previous Surname provided by user

			Impacts from different perspectives of choosing Enhanced option 1			
Option	**Description**	**Confidence in your scheme's personal data accuracy**	**Individual's Experience of this option**	**ISP Technical Operation of this option**	**Admin Implications of this option**	**Trustee Risk of this option**
1	If 1.002 Surname 1.003 DOB and 1.004 NINO (chars 1-8) from the PFS all match perfectly with those held by the scheme, then do a **positive return**, otherwise if 1.007 Alternate name 1.003 DOB and 1.004 NINO (1-8) from the PFS all match perfectly with the Surname, DOB and NINO (1-8) held by the scheme, then do a **positive return** otherwise do nothing.	**High confidence** in the accuracy of Surname, DOB and NINO on all deferred and active records.	Pensions are found for the individual to view on their chosen dashboard. 	Simple match regime to implement and operate. 	High number of found pensions, increased demand for quotations. 	High likelihood of correct matching. 
		Less than high confidence in the accuracy of DOB and NINO (characters 1-8) and less than high confidence that the Surname held by the scheme is either the member's accurate current or previous Surname on all deferred and active records.	Some pensions won't be found, meaning the individual won't see them on their chosen dashboard, but less than with the unenhanced Option 1. 	Simple match regime to implement and operate. 	'Failure demand' from individuals not seeing pensions they are expecting to see, but less than with unenhanced Option 1. 	High likelihood of false negatives, i.e. not discharging dashboards duties, but less so than with unenhanced Option 1. 

6 Core guidance

Other matching options: Matching Options 2 and 3

Without high confidence in the accuracy of the three core data elements of Surname, DOB and NINO there are significant risks in schemes choosing matching Option 1.

Schemes may be on their way to reaching high levels of confidence in their Surnames, DOBs and NINOs (where Option 1 would be appropriate), but may be reluctant to adopt Option 1 until they reach this point.

Option 1, with perfect Surnames, DOBs and NINOs on all deferred and active records is what all schemes should be aiming for. But not all schemes are there yet, so in the interim, other matching options may be required.

Schemes will need to engage in detail with their ISPs and their administration teams about how they require more sophisticated matching to work. Some possibilities are depicted in the tables on the following pages (and Annexes A and B):

Note that neither Option 2 nor Option 3 is as optimal (green) as Option 1 with perfect data.

6 Core guidance

Other matching options: Matching Options 2 and 3

Option 2 still focuses on the three key data elements of Surname, DOB and NINO but also employs some rules to generate 'maybe matches'

Schemes will need to engage in detail with the ISPs about how they require this more sophisticated matching method to work. Some possibilities are depicted in Annex A.

Schemes should also engage with their administration teams. Option 2 may result in 'maybe matches' (see Section 7 below), which is better for individuals than Option 1 (where Surname, DOB and NINO do not match perfectly).

However, this is a more complex match process for ISPs to operate and could lead to increased demand on your administrators to resolve the maybe matches.

For balance, by notifying maybe matches, it also reduces trustees' risk of returning false negatives.

		Impacts from different perspectives of choosing Option 2			
Option	**Description**	**Individual's Experience of this option**	**ISP Technical Operation of this option**	**Administration Implications of this option**	**Trustee Risk of this option**
2	If Surname, DOB and NINO nearly match, then do a **maybe return**, otherwise do nothing. See Annex A for some suggested 'nearly' scenarios.	Maybe matches will be notified to the individual, i.e. better than Option 1 where a scheme's data is not perfect.	Moderately complex match regime to implement and operate.	Demand from individuals for resolution of maybe matches.	Reduced risk of false negatives through notification of maybe matches.

Note that neither Option 2 nor Option 3 is as optimal (green) as Option 1 with perfect data.

6 Core guidance

Other matching options: Matching Options 2 and 3

Option 3 broadens the matching to consider other data elements.

You may have noted that Enhanced Option 1 is actually one very specific example of Option 3 because it looks at 1.007 Alternate Name when Surname does not match, and it looks at NINO chars 1-8 instead of all 9 characters.

Option 3 could be preferable to Option 2 because it may 'upgrade' some maybe matches to positive returns, based on additional data elements. This could improve the individual's experience by reducing the number of maybe matches, but is a more complicated match regime for ISPs to build and operate. There will still be demand on administrators to resolve maybe matches, but arguably lower than under Option 2.

However, consider these extra fields with caution. Are you any more certain of the quality of the data in these fields than you are in the main three fields?

You may be just as likely to turn a maybe match into a 'No' if these fields are added. If you want to use more fields, you need to check the accuracy of the data in those fields too.

By comparison to Dates of Birth and NINOs, these other fields typically have fewer format constraints and so matching on the plain text in these fields can have poorer outcomes.

Schemes will need to engage in detail with their ISPs about how they require this more sophisticated matching method to work. They will also need to discuss the approach to maybe matches with their administration teams.

		Impacts from different perspectives of choosing Option 3			
Option	**Description**	**Individual's Experience of this option**	**ISP Technical Operation of this option**	**Admin Implications of this option**	**Trustee Risk of this option**
3	If Surname, DOB and NINO nearly match, then compare with other data elements to check whether doing a **positive return** is appropriate See Annex B some suggested 'nearly' scenarios together with other potential data elements to be considered, which could include: • 1.001 Forename 1.007 Alternate name (e.g. Maiden name) • 1.010 Address Line 1 • 1.015 Postcode	Number of maybe matches may be less than with Option 2.	Complex match regime to implement and operate.	Demand from individuals for resolution of maybe matches may be less than with Option 2. But demand for retirement quotations on found pensions may be higher.	Reduced risk of false negatives through notification of maybe matches, but potentially not as many as with Option 2

Note that neither Option 2 nor Option 3 is as optimal (green) as Option 1 with perfect data.

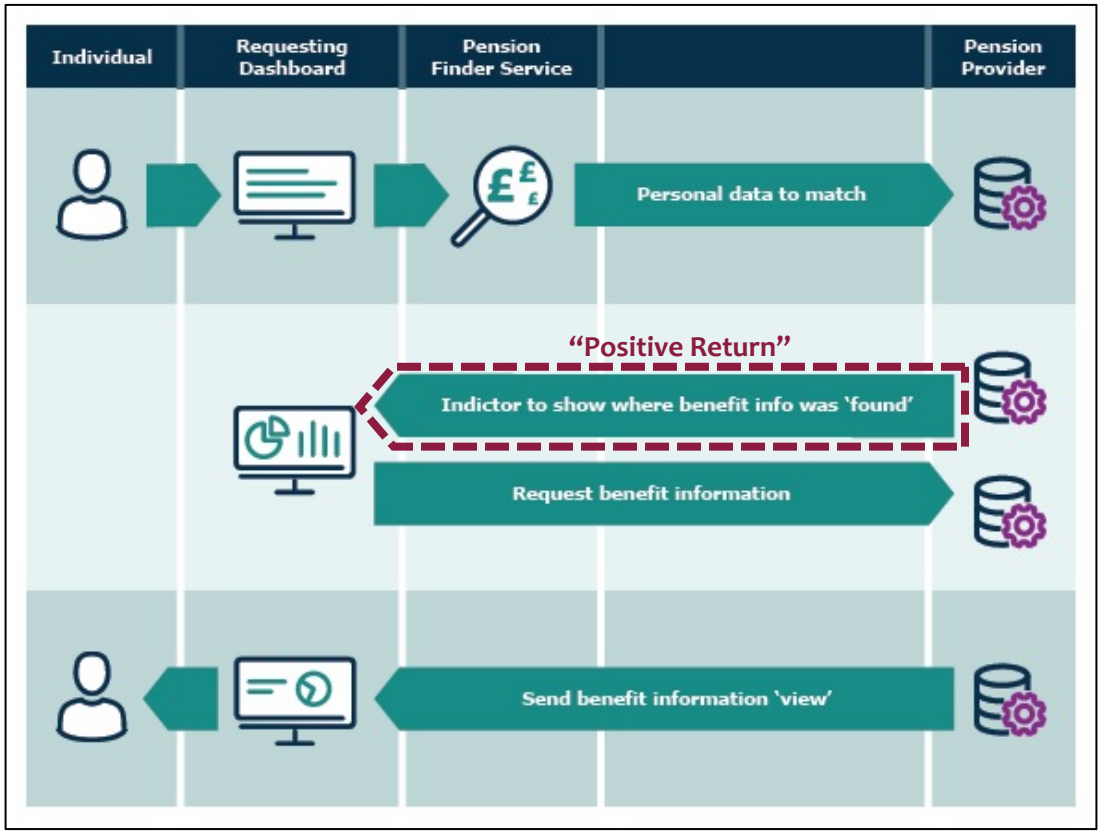
7 Maybe matches

Where a scheme is confident they have a pension record for a dashboard user (for example under matching Option 1 with perfect data), they will indicate this back to the user's chosen dashboard. This process is highlighted on the PDP process diagram to the left:

Throughout we refer to this 'pension found' message sent from a scheme as a '**Positive Return**'

There will be circumstances, however, where schemes think they may have a pension for a user, but their key personal data elements do not match perfectly. For example, see matching Scenario 20210 in Annex A (where NINO and DOB match perfectly but Surname doesn't quite match perfectly).

In these circumstances, schemes may wish to indicate to the dashboard user they **may** have a pension with them. The message the scheme would send back to the user through their chosen dashboard when they make such a maybe match is referred to as a '**Maybe Return**'. Important: The process flow and data for a 'Maybe Return' has not yet been defined by PDP. Once the situation is clearer, we will assess to determine if supplementary Guidance is needed.



8 Practical case study examples

This section provides case study examples illustrating how schemes might use this Guidance to help them decide how to match, working closely with their administrator and technology providers.

Scheme 1:

High confidence in data accuracy of Surname, DOB and NINO on all deferred and active member records

Scheme 1:

Working with a specialist data services provider, this scheme has undertaken a range of data review and cleanse activities with their members, with their participating employers and with HMRC. They also operate thorough ongoing data checking procedures. As a result, the scheme has a high confidence it holds up-to-date and accurately spelt Surnames, DOB and NINOs on all of its deferred and active member records.

Looking at the DMC Guidance, the scheme trustees are comfortable there is low risk in instructing their ISP to match in accordance with Option 1, i.e. a Positive Return where Surname, DOB and NINO match perfectly, but otherwise doing nothing.

Improving Scheme 1's data even further: Scheme 1 has invested significantly in ensuring at least Surname, DOB and NINO are accurate on all their deferred and active records. As a high performing scheme, all opportunities to check and improve the accuracy of their data is taken.

One opportunity might be, where the ISP makes a perfect match on Surname, DOB and NINO, to also check the other data elements provided from the dashboards ecosystem. For example, does the address provided match the address Scheme 1 is holding. If not, and because the scheme is sure it is the right person from the perfect match on Surname, DOB and NINO, Scheme 1 might wish to update its records with the address provided by the dashboards ecosystem. This goes beyond matching, turning the user's dashboard use into a potential data cleanse opportunity. This could also potentially extend to checking / recording the email address and mobile numbers provided by the ecosystem.

It may be possible to publish supplemental Guidance in due course on the extent to which schemes can use the address, email and mobile data provided from the PFS to update their records.

8 Practical case study examples

Scheme 2:

Some known data accuracy issues with Surname, DOB and NINO

Scheme 2:

Like Scheme 1, Scheme 2 has been working to improve the accuracy of the Surnames, DOBs and NINOs it holds, but has some known data accuracy issues with these data elements which remain unresolved.

The scheme doesn't have high levels of confidence in the address data it holds either, particularly for deferred members, where many are labelled as 'gone aways' (i.e. the home postal address the scheme holds is known to be out of date).

From reviewing the DMC Guidance, the Scheme 2 trustees are reluctant to instruct their ISP to match in accordance with Option 1 because they realise this would likely lead to false negatives (i.e. where they do not do a Positive Return when they probably should because Surname, DOB and NINO do not match perfectly for particular users). The trustees recognise this would be bad for members, and bad for their administrator, and also present significant risks for them as trustees.

The trustees are also not confident about using additional data elements in the match process because of the known issues with the addresses they hold.

The trustees instruct their ISP to match in accordance with Option 2, sending a Positive Return where Surname, DOB and NINO all match perfectly, but a Maybe Return where they nearly match (in accordance with the scenarios in Annex A).

The trustees recognise Option 2 is suboptimal to Option 1 from several perspectives. Maybe matches are less than ideal for members, and it's more costly for the ISP/administrator to operate Option 2 (costs which will be passed on to trustees or members). The trustees may determine to continue improving the accuracy of the Surnames, DOBs and NINOs they hold with the aim of being able to consider instructing their ISP to switch to matching Option 1 in due course.

8 Practical case study examples

Scheme 3:

Significant issues exist around data accuracy of Surname, DOB and NINO

Scheme 3:

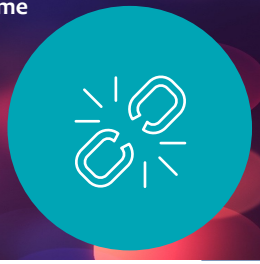
Scheme 3 has known issues with the accuracy of some of the Surnames, DOBs and NINOs it holds, so the trustees know it would not be appropriate to choose matching Option 1.

The trustees decide they want to maximise the number of Positive Returns they can, by looking beyond the three core data elements of Surname, DOB and NINO and they instruct their ISP to match in accordance with Option 3/Annex B scenarios.

Like Scheme 2, the trustees of Scheme 3 realise it's more expensive for their ISP to operate matching Option 3 than it would be to do simple Option 1 matching. They determine to investigate the costs of improving their Surname, DOB and NINO accuracy to see if it might be more cost effective to make these improvements and ask their ISP to switch to Option 1 matching at some point in the future.

Scheme 4:

Mixture of accuracy – i.e. high accuracy in main section, but low accuracy in a smaller section of the scheme



Scheme 4:

The main section of Scheme 4 has high accuracy of Surnames, DOBs and NINOs, but a smaller section, added when the sponsoring employer acquired a subsidiary, has lower quality data (which was received as part of a bulk transfer exercise from a different pensions administrator).

After considering the DMC Guidance and thoroughly investigating the data accuracy in their main section and the subsidiary's section, the Scheme 4 trustees instruct their ISP to match in accordance with Option 1 for the main section, but Option 3 for the subsidiary's section (until such time as they can improve that section's data accuracy up to the level of the main section).

9 Special categories

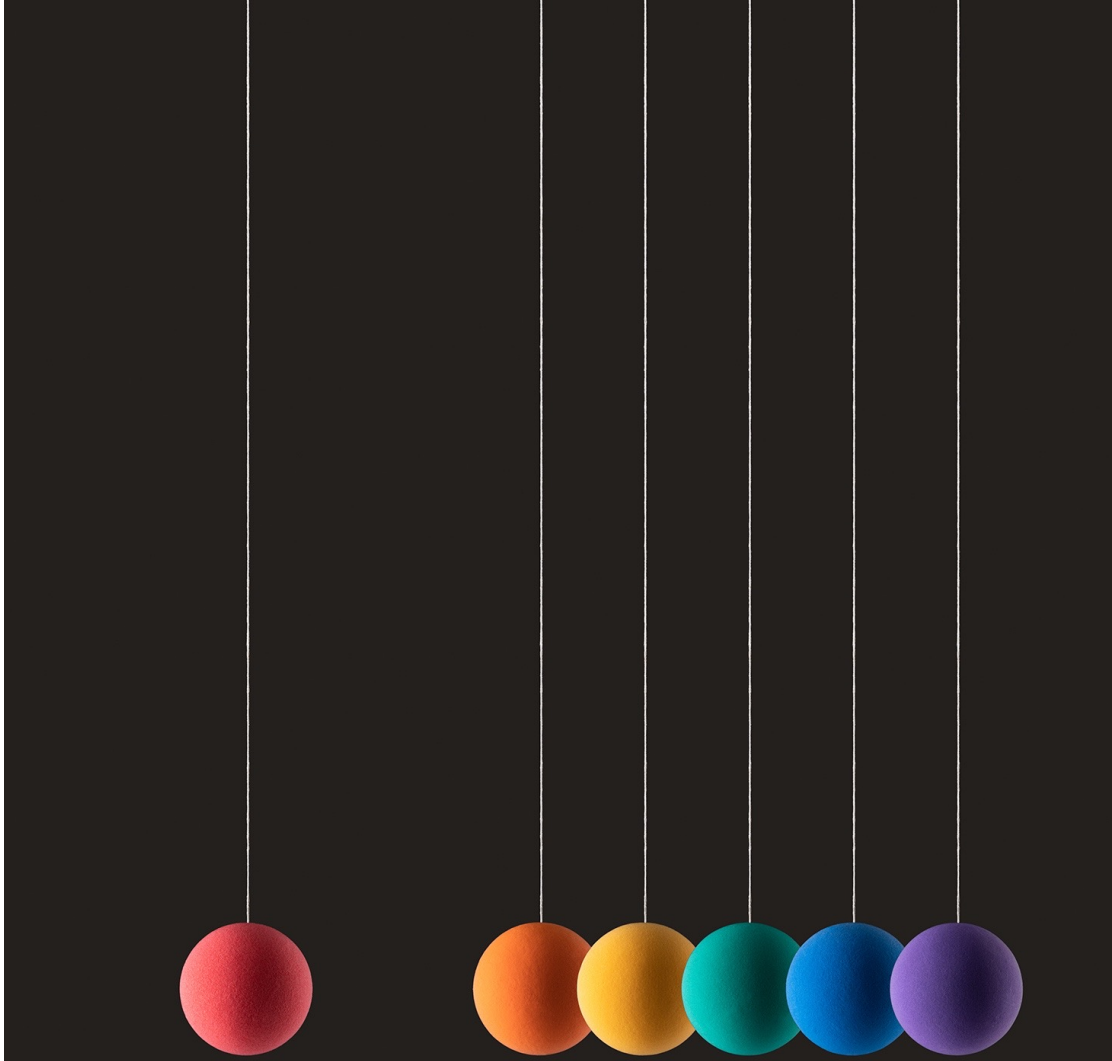
This section covers situations which require special consideration:

Members without a NINO, for example overseas individuals

In the PDP DSG, NINO is a mandatory data element, stating a value of 'N/A' will be provided by the PFS 'for those individuals who do not have a NINO'.

All respondents to the PDP Call for Input on Data Standards said they would need to match on NINO. Where no NINO, or rather a NINO of N/A, is provided by the PFS, schemes need guidance on what to do.

When more information becomes available, supplemental guidance on this and other special cases will be published.



10 Call to action, next steps and links to other Guidance

1

All schemes should now be engaging on **this important topic** of matching. To make this manageable and cost effective for trustees, the pensions administration community wants to liaise with schemes and work with them to devise solutions which work for all parties.

2

Working very collaboratively with their administrator, technology providers and wider data specialist providers, schemes should seek to understand the level of confidence in the accuracy of the Surnames, DOBs and NINOs on all deferred and active member records.

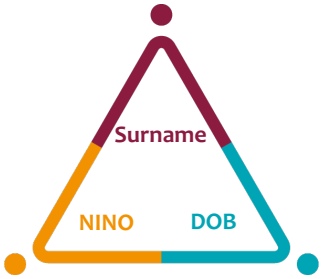
3

Schemes should take steps to improve the accuracy of this personal data. For more on this, see the Guidance published by the PASA Data Working Group: *https://www.pasa-uk.com/guidance/data/*

4

Having understood the accuracy of Surnames, DOBs and NINOs held, and improved it, this DMC Guidance should be reviewed to determine how your scheme's ISP should match:

- The optimal approach is Option 1, i.e. simple matching on Surname, DOB and NINO. But a high degree of confidence in the complete accuracy for all deferreds and actives of the data held within these three data elements is a prerequisite. You may also wish to consider enhanced Option 1, comparing against the dashboard user's Previous Surname (if provided by the PFS) and ignoring the ninth character of NINO.

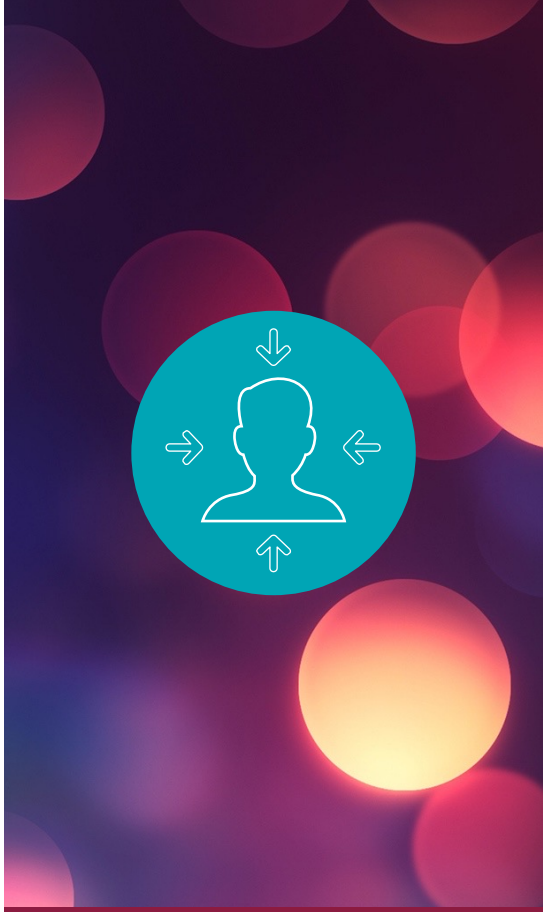


5

Schemes need to carry out a personal data accuracy investigation (and improvements), and make matching decisions, well in advance of the scheme's specific dashboards staging date .

6

PDP is currently in its alpha test phase. Following alpha, beta testing with real data will likely bring to light learnings which will be highly relevant to data matching.



The staged introduction of dashboards compulsion on schemes is expected to commence from April 2023. This first wave will impact large master trusts and pension providers. Live data matching, at scale, for these early 'stagers' may give rise to further important learnings which can be shared with the industry through supplemental guidance.

Pensions Dashboards Working Group

Annexes

December 2021

Annex A – matching option 2 nearly match scenarios

This Annex applies where schemes are adopting Option 2 matching, i.e. using just the three core data elements of Surname, DOB and NINO (like Option 1) but then going on to adopt some matching rules for 'maybe matches' where some of these three elements nearly match. The below are just examples schemes could use.

Scenario code	1.002 Surname	1.003 DOB	1.004 NINO	Result/Data provider action
20210 **(Surname nearly matches, e.g. just one character of Surname does not match)**	(×) Not a 100% match but, for example, only one character does not match	✓ 100% match of all the YYYY, MM and DD parts	✓ 100% match of all nine characters	***Example:*** *1.002 Surname from the PFS is Macdonald but the scheme holds Mcdonald* Some schemes may be comfortable treating this as if it was a perfect match and doing a **Positive Return**, especially if their ISP uses sophisticated tools, such as analysing the Levenshtein distance of the near match. Others may wish to a do **Maybe Return** so, if appropriate, the Surname they hold can be updated.
20220 **(Surname does not match at all)**	× Not a 100% match and more than one character is different	✓ 100% match of all the YYYY, MM and DD parts	✓ 100% match of all nine characters	***Example:*** *1.002 Surname from the PFS is Macdonald but the scheme holds Anand* Some schemes may wish to do a **Maybe Return** so, if appropriate, the Surname they hold can be corrected.
20310 **(DOB does not match)**	✓ 100% match of every character (max 35)	× Not a 100% match of all YYYY, MM and DD parts	✓ 100% match of all nine characters	***Example:*** *1.003 DOB from the PFS is 1987-07-18 but the scheme holds 1987-07-19, i.e. the DD part differs.* Some schemes may wish to do a **Maybe Return** so, if appropriate, the DOB they hold can be corrected.

Annex A – matching option 2 nearly match scenarios

Scenario code	1.002 Surname	1.003 DOB	1.004 NINO	Result/Data provider action
20420 **(NINO does not match at all)**	✓ 100% match of every character (max 35)	✓ 100% match of all the YYYY, MM and DD parts	× Not a 100% match and more than one character is different	***Example:*** *1.004 NINO from the PFS is NZ134289D but the scheme holds AC639273E* Schemes will most likely consider this a negative match, despite Surname and DOB matching. Option 3 matching on additional data elements may help in this scenario.
20510 **(Surname nearly marches but DOB does not match)**	(×) Not a 100% match but, for example, only one character does not match	× Not a 100% match of all the YYYY, MM and DD parts	✓ 100% match of all nine characters	***Example:*** *1.002 Surname from the PFS is Macdonald but the scheme holds Mcdonald. 1.003 DOB from the PFS is 1987-07-18 but the scheme holds 1987-07-19, i.e. the DD part differs.* Some schemes may wish to a do **Maybe Return** so, if appropriate, the Surname and DOB they hold can be updated.
20520 **(Surname and DOB do not match)**	× Not a 100% match with >1 character different	× Not a 100% match of all YYYY, MM and DD parts	✓ 100% match of all nine characters	***Example:*** *1.002 Surname from the PFS is Macdonald but the scheme holds Anand. 1.003 DOB from the PFS is 1987-07-18 but the scheme holds 1987-07-19, i.e. the DD part differs.* Some schemes may wish to a do **Maybe Return** so, if appropriate, the Surname and DOB they hold can be updated.

Annex B – matching option 3 scenarios

This Annex applies where schemes are adopting Option 3 matching, i.e., using the three core data elements of Surname, DOB and NINO (like Option 1) but then going on to look at some additional data elements where some of the three core data elements nearly match. Below are just examples schemes could use.

Scenario code	1.002 Surname	1.003 DOB	1.004 NINO	1.001 Forename	1.007 Alternate name	1.010 Address Line 1	1.015 Postcode	Result/Data provider action
30210 **(Surname nearly matches but Forename and Address match)**	(×) Not a 100% match but, for example, only one character does not match	✓ 100% match of YYYY, MM and DD parts	✓ 100% match of all nine characters	✓ 100% match or match on first character	Not compared	✓ 100% match	✓ 100% match	***Example:*** *1.002 Surname from the PFS is Macdonald but the scheme holds Mcdonald* Under Option 2, this scenario might have led to a Maybe Return. But the fact Forename (or first initial) and Address Line 1/Postcode match might mean some schemes are comfortable doing a **Positive Return**. Some schemes may wish to use Postcode only.
30310 **(DOB does not match but Forename and Address match)**	✓ 100% match of every character (max 35)	× Not a 100% match of all the YYYY, MM and DD parts	✓ 100% match of all nine characters	✓ 100% match or match on first character	Not compared	✓ 100% match	✓ 100% match	***Example:*** *1.003 DOB from the PFS is 1987-07-18 but the scheme holds 1987-07-19, i.e. the DD part differs* Under Option 2, this scenario might have led to a Maybe Return. But the fact Forename (or first initial) and Address Line 1/Postcode match might mean some schemes are comfortable doing a **Positive Return**. Some schemes may wish to use Postcode only.

Annex B – matching option 3 scenarios

Scenario code	1.002 Surname	1.003 DOB	1.004 NINO	1.001 Forename	1.007 Alternate name	1.010 Address Line 1	1.015 Postcode	Result/Data provider action
30410 **(NINO nearly matches but Forename and Address match)**	✓ 100% match of every character (max 35)	✓ 100% match of all the YYYY, MM and DD parts	(×) All characters match except for the ninth character	✓ 100% match or match on first character	Not compared	✓ 100% match	✓ 100% match	***Example:*** *1.004 NINO from the PFS is NZ134289D but the scheme holds NZ124289* Under Option 2, this scenario might have led to a Maybe Return. But the fact Forename (or first initial) and Address Line 1/Postcode match might mean some schemes are comfortable doing a **Positive Return**. Some schemes may wish to use Postcode only.
30420 **(NINO does not match at all but Forename and Address match)**	✓ 100% match of every character (max 35)	✓ 100% match of all the YYYY, MM and DD parts	× Not a 100% match and more than one character is different	✓ 100% match or match on first character	Not compared	✓ 100% match	✓ 100% match	***Example:*** *1.004 NINO from the PFS is NZ134289D but the scheme holds AC639273E* Under Option 2, this scenario might have been considered a negative match, but the fact that Surname, DOB, Forename and Address match might mean some schemes wish to do a **Maybe Match**, leading (if appropriate) to the correction of the NINO held by the scheme.

Annex B – matching option 3 scenarios

Scenario code	1.002 Surname	1.003 DOB	1.004 NINO	1.001 Forename	1.007 Alternate name	1.010 Address Line 1	1.015 Postcode	Result/Data provider action
30510 **(Surname nearly marches, DOB does not match, but Forename and Address match)**	(×) Not a 100% match but, for example, only one character does not match	× Not a 100% match of all the YYYY, MM and DD parts	✓ 100% match of all nine characters	✓ 100% match or match on first character	Not compared	✓ 100% match	✓ 100% match	***Example:*** *1.002 Surname from the PFS is Macdonald but the scheme holds Mcdonald. 1.003 DOB from the PFS is 1987-07-18 but the scheme holds 1987-07-19.* Given that NINO, Forename and Address match, some schemes may wish to a do **Maybe Return** so, if appropriate, the Surname and DOB they hold can be updated.
30520 **(Surname and DOB do not match, but Forename, Maiden surname and Address match)**	× Not a 100% match and more than one character is different	× Not a 100% match of all the YYYY, MM and DD parts	✓ 100% match of all nine characters	✓ 100% match or match on first character	✓ 100% match of all characters (and 1.006 Alternate name type = M for Maiden)	✓ 100% match	✓ 100% match	***Example:*** *1.002 Surname from the PFS is Macdonald but the scheme holds Anand. 1.003 DOB from the PFS is 1987-07-18 but the scheme holds 1987-07-19. 1.007 Alternate name from the PFS is Anand.* Given that NINO, Forename, Maiden surname and Address match, some schemes may wish to a do **Maybe Return** so, if appropriate, the Surname and DOB they hold can be updated.

Annex C – given name/forename challenges

What do schemes hold?

The first part of an individual's name (i.e. everything except their Surname) is held in many ways by different schemes. The existing TPR Record-keeping guidance specifies schemes should hold 'either forename or initials' (second bullet point under Common data).

For a member whose full name is, say, Richard William Edwin Smith, schemes might hold: R or RW or R[space]W or RWE or Richard or Richard William or Richard William Edwin, or other values such as Dick or William (for example if the individual is usually known by a shortened or middle name).

What do schemes match on today?

Some schemes match (with human intervention, e.g. when the member phones in) on just Surname, but most match on 'Full name', which usually means matching on First name (i.e. Richard) or Initial(s) where this is all they hold (i.e. R or RW or RWE), and Surname. This human intervention deals with any lack of precision in the spelling of names – something not possible in automated digital matching.

What will PFS provide to all data providers?

Forename will be collected by dashboards and 'passed down' by the PFS to all data providers (as data element 1.001). For example, for a user called Richard William Edwin Smith, the name data elements 'passed down' by the PFS will be Richard (1.001) and Smith (1.002). Middle names (William and Edwin) are explicitly excluded.

What additional confidence would be added by comparing Forename?

A straightforward example is where the NINO, DOB and Surname passed down from the PFS match perfectly with those held by a scheme giving a high level of confidence this is a positive match.

Would this high level of confidence be increased further if Forename also matches? Many schemes hold only initials (which would mean, for example, matching the first character of Richard from the PFS against RW held by the scheme). There's a greater chance of variations in Forename than with Surname (for example, Dick might be passed from the PFS compared to Richard or R held by the scheme, or vice versa), which might reduce the strong confidence in the match.

Where NINO and DOB match perfectly, but Surname doesn't (for example if the member has married), this could be a 'maybe match'. In this circumstance, a perfect match on Forename might strengthen the confidence this is a maybe, but would it increase the confidence enough to turn this maybe into a positive match? A perfect match against Alternate name (1.007) from the PFS (which could be the member's Maiden surname) would be more likely to upgrade a maybe to a positive match.

Get in touch:

info@pasa-uk.com www.pasa-uk.com

PASA is a Community Interest Company, and our full name is Pensions Administration Standards Association CIC.
Company number: 6597097