



Briefing Note: Learner identity and the National Student Index

To:	Hon. Jan Tinetti, Associate Minister of Education		
Cc:	Hon. Chris Hipkins, Minister of Education		
Date:	21 June 2021	Priority:	Medium
Security Level:	In Confidence	METIS No:	1255292
Drafter:	Alison Anderson	DDI:	04 463 7084
Key contact and number:	Wendy Hamilton	DDI:	04 463 0934
Messaging seen by Communications team:	No	Round Robin:	No

Purpose of Report

The purpose of this paper is for you to:

- Note** that learner identity information, currently stored in the National Student Index, (NSI) is the backbone of the education data system, but its quality is deteriorating due to weaknesses in the way the identification of learners and assignment of National Student Numbers (NSNs) takes place.
- Note** that system initiatives such as digital identity for online learning and assessment and Te Rito will rely on each learner having been properly identified and having just one NSN, so that their full education record can follow the learner throughout their education journey.
- Note** that we are re-building data matching functionality so that we can clean the NSI, but that a more holistic solution would be to stop multiple NSNs being created, by ensuring the way we carry out identification is fit for purpose.

d.

9(2)(f)(iv)

Noted

- Agree** to proactively release this Briefing Note.

Agreed / Not Agreed

Summary

- Learner identity is at the heart of a learner's education journey. If the system cannot rely on each learner having been properly identified and having a unique NSN, it will not be able to deliver the expected benefits of a learner's record following them throughout their education journey.
- We are re-building data matching functionality and obtaining data from DIA so that we can clean the NSI, but this is expensive, inefficient and we cannot make sure the data is up to date.
- It would be preferable to stop poor quality data being created, instead of continually having to fix it "after the event".
- We propose investigating options for making improvements to digital services, the identification of learners and the assignment of National Student Numbers so that we can provide you with preliminary advice. This will include investigation of DIA's Digital Wallet and SmartStart initiatives, as examples of digital services focused on identity.



Alexander Brunt

Deputy Secretary

Evidence, Data and Knowledge

25 / 6 / 2021



Hon Jan Tinetti

Associate Minister of Education

26 / 06 / 2021

What is the NSI?

1. The National Student Index (NSI) holds learner identity information. Its purpose is to link that identity information to a unique National Student Number (NSN) for every student. This identity information forms the backbone of the education data system.
2. Learner identity is at the heart of a learner's education journey. The learner's records – of their needs, progress, achievement, and qualifications as they go through that journey – need to be associated solely and completely with their identity. This means that the assignment of an NSN must be part of a robust identification process, which establishes that the learner is who they say they are and links that identification process to their NSN.
3. Today, most learners are assigned an NSN when they first enrol with an early learning service. Otherwise they are assigned one by a subsequent early learning service provider, kura, school or tertiary provider. The provider is expected to ask the learner or their parents for evidence of identity. The provider then records the learner's identity details in the NSI and assigns them an NSN.
4. Many adults don't have an NSN because it wasn't mandatory when they were at school. However, an increasing number are now getting an NSN for the first time, when they enrol for education or training related to their work.
5. The NSI only holds NSNs and identity information relating to:
 - Family name
 - Given names
 - Alternative names
 - A preferred name indicator
 - Date of birth
 - Date of death (if populated)
 - Gender
 - Residential status

It does not hold any other information that is accessible to sector users. Other information relating to the creation, modification and verification of records is stored for administrative use.

What is the problem?

6. It is evident that our current way of carrying out identification of learners and assigning NSNs is not working. Providers do not always follow the expected identification process. They may make mistakes with recording the name, or even date of birth, in the NSI.
7. Too many NSNs have incorrect identity information. This often leads to the learner getting a new NSN each time they enrol with another provider.
8. Over time, a learner is assigned multiple NSNs, without knowing it. Each of these NSNs has a different part of their education journey recorded against it. However, no one can see the whole record, because no one knows they all belong to one learner. The learner's current provider can only see whatever is attached to their current NSN and has no way to access their previous record or any credits they may have gained.
9. There has been an increasing build-up of multiple NSNs in recent years. The Ministry used to try and fix this by matching NSNs against the birth register, however this "cleaning" process was discontinued in 2015.

10. At that time, the Ministry intended to integrate with DIA's real time identity verification system, RealMe, however this was not initially in a state to be usable. By the time RealMe became usable, the cost of integration was considered too high.
11. As a result, there has been a gap of six years without any identity data in the NSI having been "cleaned".

Why are multiple NSNs a problem?

12. Having multiple NSNs can often affect a learner when they apply for entitlements or a tertiary course with entry requirements. A common scenario is that some of the learner's credits are recorded against their current NSN, while the rest are recorded against another NSN they are unaware of. When the provider checks, using the current NSN, they find the learner has insufficient credits to enrol in the course. The provider has to work with the learner to find the other NSN under a different or misspelled name.
13. The next task is to assemble the evidence that both NSNs rightfully belong to that learner. This can be difficult, as each provider will have recorded them as a different identity, with a different name and sometimes different gender or birth date. The learner has to assemble sufficient evidence that each of these identities is theirs. The provider then applies to the Ministry to get the NSNs merged.
14. Merging multiple NSNs requires a lot of time and effort from Ministry teams, including the Service Desk. They often have to contact providers who had the learner on their roll several years ago and ask them to supply copies of their records. Last year the Service Desk completed 1,577 merges, plus 5,709 requests to amend a date of birth or name.
15. NZQA also completes merges of multiple NSNs, working with the Ministry when required. In 2019, it completed 1,915 merges, taking the equivalent of 8 weeks work for 1 FTE. In 2020, it completed 1,136 merges. The process can be convoluted and time-consuming when learners do not have the identification documentation to support the merge.
16. Learners having multiple NSNs is an issue for TEC in relation to funding eligibility. A learner's current NSN is used as the identifier in determining their eligibility for the likes of Fees Free and the Targeted Training and Apprenticeship Fund. Part of determining eligibility is assessing what the learner has been enrolled in previously. For example, if they have been recorded as enrolled in 60 credits or more at level 3 or above on the New Zealand Qualifications Framework, they are not entitled to Fees Free. If a learner has multiple NSNs, then all or part of their previous study will have been recorded against NSNs that are different from their current NSN. That previous study will not be included in the assessment and they can be incorrectly deemed eligible for Fees Free.
17. Poor quality identity data in the NSI will affect the Education Work Programme. If the system cannot rely on each learner having a unique NSN, it will not be able to fulfil our commitment to "the learner at the centre" and will not deliver the expected benefits of a learner's record following them throughout their education journey. For example, learners with learning support needs are more likely to experience gaps or delays in receiving the support they need if they have multiple NSNs.

How are we cleaning the NSI?

18. We are re-building the data matching functionality in the NSI. This is an automated function that will compare NSI data against official identity data from DIA. It will automatically verify exact matches and provide information about partial matches that will enable potential multiples to be identified, investigated and merged. The aim is to detect and merge multiples *before* they impact the student. We are also using deaths data to record date of death against matching NSNs, to prevent potential fraudulent use of NSNs.

19. We have an Information Matching Agreement (IMA) with DIA and have already obtained births and deaths data to cover the gap since 2014, so that we can match and clean the NSNs that were missed. This has cost us \$330,000. Under the IMA we can obtain data for new births and deaths monthly, so we can match NSNs for young children soon after they are created. This will cost us \$40,000 a year. In addition, we are arranging to obtain data to cover adults up to the age of 53 who did not get an NSN when they were at school and who have had one assigned by a tertiary provider, and deaths data to cover another gap. This will cost us \$421,000.
20. The benefits of this data matching against DIA data should be that the number of multiples, and merges required, reduces over time. 9(2)(f)(iv)

Why is cleaning the NSI with DIA data not a long-term solution?

21. Fixing poor quality data by matching and cleaning "after the event" is not optimal, because it does not address the cause of the problem. It is also expensive, inefficient and we cannot make sure the data is up to date.
22. Costs are incurred because of the fees DIA requires other government agencies to pay for the data it provides under Information Matching Agreements (IMAs). Under the Privacy Act 2020, fees may be agreed between the parties to an IMA. These fees can be increased by the Registrar-General with one month's notice, as happened in June 2014.
23. The Secretary for Education has recently written to the Chief Executive of DIA about DIA's approach. Paying fees for the work to extract and transfer the data may be understandable, however DIA also charges by volume. This is harder to understand, as the work is the same regardless of how much data is in the file.
24. If we were able to reduce or remove the cost of purchasing data, we could be able to re-direct investment towards more FTEs to work on the manual matching. This would achieve a material increase in data quality, faster. A meeting of officials is now being arranged to discuss a way forward.
25. We have to keep data matching because that is the only way currently available to us. It would be preferable to find a way to prevent poor quality and multiple NSNs from continually being created.

How can we do better?

26. If we are to do better, we need to look at the way we carry out identification of learners and assignment of NSNs, and at new digital service solutions. Having NSNs assigned by providers is not working well, and may no longer be necessary, given digital advances.
27. Application Processing Interface (API) technology could offer a way to create a single NSN per learner, based on verified identity data from other organisations, right from the start, and keep it up to date. APIs could provide real-time data from official identity sources, to ensure that the quality of learner identity data and NSNs was assured and maintained.

28. DIA is currently working towards making new APIs available within the next few years, although it is not yet clear what they will offer and what charges there will be. We have offered to support DIA in any way to help get the APIs get up and running.
29. If DIA cannot provide an API service that meets our needs, or not within the time frame we need it, third party providers could be an option worth investigating. Third party providers verify the identity of an individual using APIs that connect to government data and other sources. However, they only confirm if the identity information you submit to them is correct, they do not provide updated information. They charge for their services.
30. Digital services currently being explored by DIA, such as SmartStart and the Digital Wallet application, offer another opportunity. In future, an individual may be able to use a digital service to assert their identity, already verified by DIA, and have an NSN created for them. This could potentially be done as part of enrolment, as a self-service process. We have been working with DIA on a Proof of Concept for a Digital Wallet which is aimed at enabling a person to assert their identity and share attributes such as their NSN and qualifications, when enrolling with a provider such as a university. We also intend to investigate SmartStart as a possible option for early learning and schooling.
31. Some of the ideas for potential changes signalled above may require legislative amendment, including stakeholder consultation. The Education and Training Amendment Bill (No. 2), currently proposed for introduction in late 2022, is likely to be the earliest legislative vehicle that would be suitable for this purpose.
32. **9(2)(f)(iv)**
33. This means that options for operational fixes should also be investigated, because one of them may be worth implementing as an intermediate solution.

Next steps

34. If you agree, we will provide further advice on options for making improvements to digital services, the identification of learners and the assignment of NSNs, and the timing and process for progressing any legislative amendments required, by August 2021.

Proactive Release

35. We recommend that this Briefing is proactively released as per your expectation that information be released as soon as possible. Any information which may need to be withheld will be done so in line with the provisions of the Official Information Act 1982.