

# Data integrity and quality of superannuation data

Discussion Paper

November 2011

Association of Superannuation Funds of Australia

ASFA Secretariat  
Level 6  
66 Clarence Street  
Sydney NSW 2000

PO Box 1485  
Sydney NSW 1005  
Ph: +61 2 9264 9300  
Fax: +61 2 9264 8824

Outside Sydney  
1800 812 798

Website: [www.superannuation.asn.au](http://www.superannuation.asn.au)

The Association of Superannuation Funds of Australia Limited ABN 29 002 786 290 ACN 002 786 290

**Please Note:**

This discussion paper provides general information and is not intended as advice. Although verification of the accuracy of the information contained in this paper has taken place, liability is not accepted for any errors or omissions that may have occurred.

© ASFA 2011

## Contents

1.	Introduction	4
2.	The importance of good data quality	5
3.	Issues contributing to poor data quality	6
4.	Key actions to improve data quality	7
5.	Data that should be checked	8
6.	Industry consultation	9
APPENDIX 1: ASFA-ITM DATA Benchmark		11
ASFA-ITM Core Member Details Benchmark Tests		13

## 1. Introduction

This Discussion Paper concerns the quality and integrity of data that is created and used in the administration of superannuation funds. It is issued by The Association of Superannuation Funds of Australia Limited (ASFA). ASFA is a national, not for profit, non-party political representative organisation for the superannuation industry. The content has been researched and authored by Steve Shoreson, Managing Director of ITM ... the data specialist (ITM), and reviewed by ASFA. Valuable support and input has been received from a number of other parties, including the industry's prudential regulator, the Australian Prudential Regulation Authority (APRA). Key inputs include:

- ITM's actual data integrity work and its experience in the Australian, UK and South African superannuation industries.
- Research undertaken by The Pensions Regulator (TPR) in the UK.
- Discussions with ASFA, APRA, clients of ITM and other industry stakeholders.

The industry needs a consistent methodology for checking the quality of record keeping. The aim being to provide trustees with information to enable them to assess the quality of their data objectively, benchmark it against industry standards, and create a data cleansing plan where required.

Responsibility for ensuring the accuracy of data rests with the trustees. The trustees will depend on the fund's administrators or independent providers to perform the necessary testing and analyse the results. Administrators will need to work with the trustees and other parties (such as employers) to correct any process and data problems that are identified.

It is questionable whether or not a voluntary approach to raising the quality of record keeping will lead to a significant and sustained improvement in data quality. However, initiatives such as implementing data benchmarks will help educate trustees and provide them with some measure of their own administration effectiveness. See Appendix 1 for an overview of the ASFA-ITM Benchmark.

### Consultation - We value your input

ASFA believes the industry wants to move towards better data quality in its record keeping. We would appreciate agreement, comments and/or objections from any stakeholder to the assumptions stated in section 6 of this paper.

Our desired outcome, as an industry representative, is to develop a consistent and widely-accepted approach to the issue of data quality.

We would like you to use the pro-forma in section 6 for your response, if at all possible, and send it to us by email or mail to:

Jon Echevarria  
Senior Policy Adviser  
ASFA  
PO Box 1485  
Sydney NSW 2001  
Email: [jechevarria@superannuation.asn.au](mailto:jechevarria@superannuation.asn.au)

We may share the information you send to us between ASFA and ITM. We may also publish the overall results of the responses received (on a 'no names' basis) as part of this consultation process.

## 2. The importance of good data quality

The need to ensure a high level of data integrity is applicable to all types of superannuation fund. It is relevant to the wider superannuation industry including trustees, secretariats, administrators and advisers. It is also relevant to employers to the extent of their obligation to provide data to the superannuation funds chosen by their employees. Most importantly, members have the right to expect accurate and timely record keeping and reporting of their retirement savings.

Improvements in the quality of data will have benefits to all involved in our industry, including:

- Administration providers will increase their reputation with their clients by reducing re-work, unnecessary member enquiries and providing more accurate reporting. The result being easier, more efficient and cost-effective administration.
- Trustees will have greater certainty that member benefits are more accurately recorded and reported and see a reduction in the number of member complaints.
- Employer costs will reduce through sending correct data the first time and not having to manage re-work and/or disconcerted employees.
- Members will receive better service and more accurate reporting and payment of benefits.

### Data regulations

Trustees have a fiduciary responsibility under basic trust law principles to hold accurate and up-to-date information in relation to all members.

Australia deals with data protection under a variety of laws, mainly around privacy. There is no actual Data Protection Act that is as comprehensive as in other jurisdictions such as the UK.

#### ***UK Data Protection Act 1998***

The Common Law requirement is codified in the Data Protection Act 1998 which came into force on 1 March 2000, classing trustees as 'data controllers', and those working for trustees (eg administrators and insurers) as 'data processors'. This Act makes many demands of data users. There are eight Data Protection Principles, which (amongst other things) include requirements that personal data:

- be adequate, relevant, and not excessive in relation to the purpose or purposes for which they are processed;
- be accurate, and, where necessary, kept up-to-date; and
- processed for any purpose or purposes, shall not be kept for longer than is necessary for that purpose or purposes.

This Act also requires that appropriate technical and organisational measures be taken against unauthorised or unlawful processing of personal data and against accidental loss or destruction of, or damage to, personal data.

### 3. Issues contributing to poor data quality

The impending *Stronger Super* package of reforms, particularly the SuperStream and Governance initiatives, has a focus on assisting the industry to become more efficient. In an efficient system, how can there be an average of three superannuation accounts for every Australian worker? Although there may be valid reasons for some individuals to have more than one account (eg to access generous insurance or pension benefits in a particular fund), there is general agreement that, for a workforce of around 10 million, 30 million accounts is unnecessary. Research and analysis recently undertaken by ASFA indicates that the number of accounts should be able to be reduced to around 19 million through the implementation of the various SuperStream initiatives.

#### **SuperStream commentary on inefficiencies**

“There are no mandated data or transmission requirements for transactions within the superannuation system. This has resulted in funds:

Having different data requirements and formats for processing data and payments; receiving poor quality data leading to members being incorrectly enrolled in funds with a resultant risk of multiple accounts being created.

Receiving insufficient member information preventing the correct allocation of contributions.”

#### Data issues that are evident in our industry include:

- Poor quality legacy data exists: Changes in benefit design, legislative rules, administrator and platform, have gradually deteriorated the quality of historical data records. The current spate of fund mergers has the potential to exacerbate this issue.
- The movement of data reduces data quality: There is evidence of poor quality data transferring between employers and funds, with no real repercussion on employers. Members are often not engaged enough to keep their personal details up-to-date.
- There is no industry standard for data quality: There is no industry-wide agreement on what is good or poor record keeping. In fact, there is very little reporting on data quality at any level.
- Problem resolution is ad hoc: The norm is that data problems are investigated and resolved on a case-by-case basis, rarely considering the wider impact. Administration is considered a back office function by many and therefore gets less attention than it deserves.
- Administration expertise at the trustee level is low: Trustee boards always have expertise in functions such as finance, investments and legal/compliance. However, very few have administration ‘professionals’ on the Board. Data integrity is often only given trustee consideration when a specific member problem is raised.
- Organisation-wide data integrity capability is low: The focus among administration and systems support resources needs to be improved. The current capability for identifying and then subsequently eradicating erroneous data needs to be more evident.

Most would agree with the above issues and that many of them have been in existence for quite some time.

## 4. Key actions to improve data quality

A stated aim of SuperStream is “improving the quality of data in the system”. The current SuperStream timeline has mandatory data standards and use of e-commerce for large and medium employers making contributions by July 2014 and for small employers by July 2015.

We must not lose sight of the fact that there is much room for improvement in the quality of superannuation data. It is unlikely that data quality will improve by itself between now and 2014-15. We need to ensure that data standards and e-commerce do not just become a faster way of getting incorrect data into the system.

To ensure an effective approach to improving data quality, the industry needs to undertake a number of actions.

- Educate and encourage trustees to raise their understanding and expectations of acceptable levels of data quality and related risks.
- Funds, administrators, employers and other service providers must work co-operatively on the introduction of data integrity measures.
- There is core information about every member which is fundamental to proper administration. Trustees should introduce the measurement of this data quality (eg the ASFA-ITM Data Benchmark). Other Benchmarks to be introduced include areas such as insurance, transactional information, fees and charges, etc.
- Trustees must set annual benchmarks for record keeping that will raise standards over time. Once a trustee has measured its data against the industry benchmark, it should determine its own benchmark and ensure it is showing improvement over time. The target for ‘core member details’ (as defined in section 5(i) of this Paper) should be 100 per cent accuracy.
- Trustees should request administrators to report on data integrity issues. Realistic remediation plans would be agreed between trustees and administrators where issues are found.
- The area of data integrity should become subject to regulatory oversight and the regulator should consider providing further guidance on materiality of data errors.
- Minimum acceptable data standards from employers should be mandated and enforced.
- Consideration should be given to including data integrity and data security requirements in any future regulation of administrators and trustees.

Any assessment of data quality must be accompanied by a methodology to enable the prioritisation and root cause analysis of the issues identified. This includes categorising possible data issues by:

- Financial impact and criticality of data issues.
- Issues affecting member benefit statements and reporting.
- Data that raises potential breaches, compliance issues and/or fraud indicators.
- Data issues caused by the underlying administration platform.
- Duplicate members and duplicate data.
- Data issues that are critical to any successful migration of data.

## 5. Data that should be checked

### Data audit and data cleansing

- A data audit is a repeatable and controlled process that interrogates the quality of stored data.
- The data audit identifies incomplete, incorrect, inaccurate and irrelevant data.
- Data cleansing involves modifying or deleting 'dirty' data, either manually or using automated methods.
- Very few Australian superannuation funds undertake data cleansing activities regularly and comprehensively.

To identify the true quality of data, a data integrity audit should be undertaken on three levels. As a minimum, the core member data should be audited on a regular basis.

Administrators and trustees have a responsibility to ensure that any data they hold is complete, consistent and accurate.

#### (i) Core member data fields

There is a core set of member data that is fundamental to the proper administration of funds. An error in any of these core data items is highly likely to render the member's record unusable, raise concerns over the identity of the member, or lead to an incorrect payment or payee. The aim should be that core member data eventually be 100 per cent accurate.

'Core member details' encompasses the actual fields mentioned in SuperStream (full name, date of birth, current address and TFN) plus other member-specific non-transactional fields (date joined fund, beneficiary and gender). The quality of this core data is an excellent indicator, as if a poor Benchmark result is returned by analysing the core member data, trustees could expect that other data will be of a similar poor quality. See Appendix 1 for a list of the recommended tests against these core member data fields.

#### (ii) Other data fields

There is obviously a lot more detailed data required for effective administration of a fund including contributions and other transactional fields, fees and charges, investments, insurances, etc.

The need to audit the quality of specific data fields will vary from fund to fund. For instance, salary may not be important to a fund unless it is defined benefit or benefits such as insurances are calculated using salary as a component. However, almost all funds have fees and charges and they should audit related data for correctness.

#### (iii) Calculated data fields

As well as the above data, most funds have calculated data items and benefits. While the data in a given data field may be correct, where a value is calculated using two or more pieces of data, the calculation routine may not be configured correctly.

Any calculated fields, such as benefits including insurances, should be checked for correctness on a regular basis. This type of audit would be considered essential where a migration (to a new platform or a new administrator) is taking place. The first annual review on the new platform should be accompanied by a Certificate Authority (CA) to ensure the platform was correctly configured.



## 6. Industry consultation

Please provide your comments with regard to the following statements:

	Agree Y/N	Additional comments
1. Core member data that must be audited includes full name, date of birth, gender, current address and TFN.		
2. 100 per cent data accuracy may not be achievable in the short-term for the majority of funds. Nevertheless, for core member data, the benchmark target must be 100 per cent.		
3. Other data fields should be audited and reported to trustees including contributions, fees and charges, investments and insurances.		
4. Calculated fields, such as benefits and insurances, should be audited to ensure the calculations and member benefits are correct.		
5. Trustees should develop a plan for auditing and, where necessary, cleansing their data. The audit checking period should be at least annually and formal reports should be produced.		
6. In addition to annual audits, data integrity audits should be undertaken for significant events, such as fund mergers, changes of administrators and changes of administration platform.		
7. If legacy data integrity checks reveal shortcomings, the trustee should prioritise data issues and the administrator should rectify agreed issues. Prioritisation may result in an acceptance that some data will not be cleansed.		
8. Trustees and providers should, where necessary, aim to complete data cleansing within 12 months of identification.		
9. Data audit and cleansing activity has a cost. Trustees need to ensure the data identification and cleansing is performed and the costs for the exercise fairly distributed between the fund and administrator.		
10. Data from employers can be suspect. Legislative requirements on employers to provide minimum standards of data correctness are essential to enable a high level of quality for new data received by funds.		
11. Whatever the cause, any shortcomings in the quality of data received, from any source, should demonstrate inadequate processes and controls that need to be addressed.		
12. Providers, administrators, employers and trustees will benefit from using standardised processes for transferring data. The industry needs to continue to agree and develop standardisation in relevant areas.		

	Agree Y/N	Additional comments
13. At some future date, once the industry is better informed of its responsibilities, data quality should be regulated.		
14. The Australian superannuation industry needs more specific data quality rules such as in the UK Data Protection Act 1998 (See section 2 of this Paper).		
<b>Use this section to provide additional commentary or to expand upon any of the above issues.</b>		

Please send your responses to section 6 of this paper by email or mail to:

Jon Echevarria  
Senior Policy Adviser  
ASFA  
PO Box 1485  
Sydney NSW 2001  
Email: [jechevarria@superannuation.asn.au](mailto:jechevarria@superannuation.asn.au)

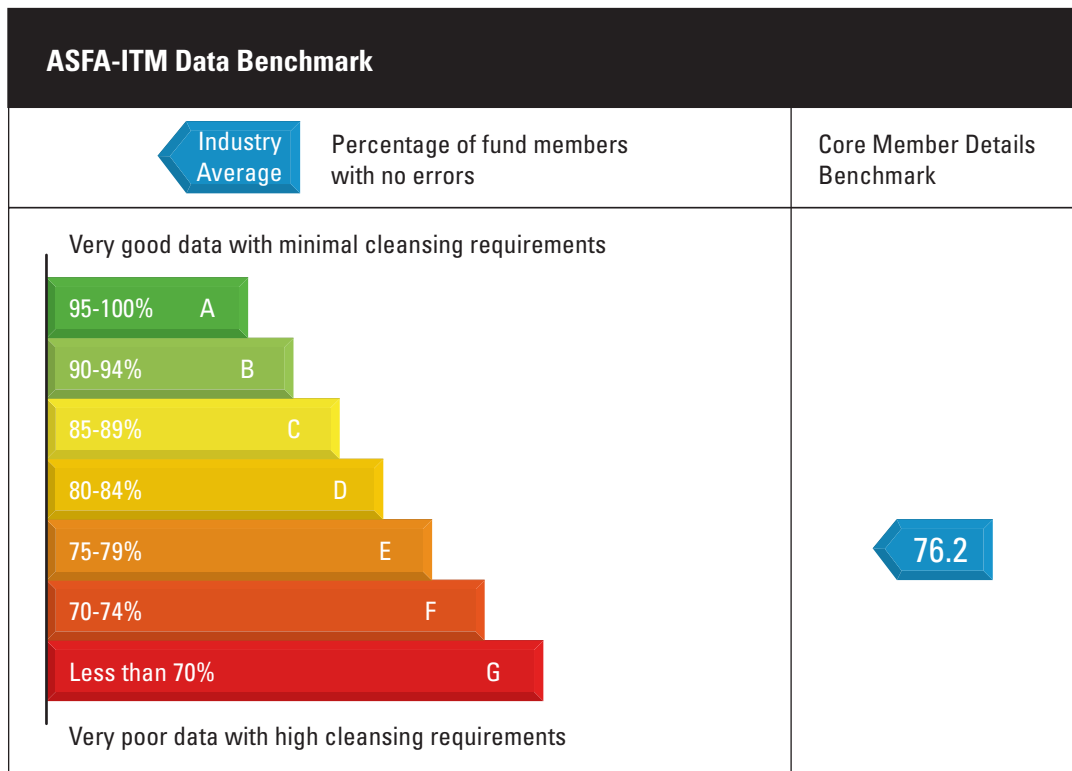
## APPENDIX 1: ASFA-ITM DATA Benchmark

ASFA has partnered with ITM to lead the industry in the push for world’s best practice. The ASFA-ITM Data Benchmark provides actual data integrity benchmarking of each fund against its peers.

There are a number of static member-related data fields that are captured for almost all superannuation funds. ASFA and ITM have used these fields to derive a data benchmark on core member data. In addition to the reports run for the benchmark, there are another 50+ reports run using the same core member data. Most clients undertake a full data audit on 100 per cent of their member and employer data and the results relating to the Core Member Details Benchmark are extracted. Funds may analyse only the core member data and if significant issues are found, they may consider undertaking a full data audit.

‘Core member details’ encompasses the actual fields mentioned in SuperStream (full name, date of birth, current address and tax file number (TFN)) plus other member-specific non-transactional fields (date joined fund, beneficiary and gender).

Over the past two years, 2.5 million active members have been analysed across 34 funds, with over 16 million test results included in the Core Member Details Benchmark. Many millions more tests have been run (including on exited member data) which returned results. These are not included in the Benchmark however since they are seen as lower criticality issues. ITM also runs audit tests in areas including fees and charges, contributions, insurances, investments, unit pricing, taxation and exited members.



The blue arrow on the Benchmark indicates that 76.2% of the 2.5 million active members tested by ITM have no errors. Alternatively, over 23 per cent have an error in one or more of the audited data fields. Trends have indicated that larger funds have a higher ratio of errors.

## Breakdown of the core member data audited

The chart below provides a breakdown of the actual clean members by each data field tested. For example 98.8 for DOB (date of birth) indicates that 1.2 per cent of members tested have issues with this data field.

Most industry funds, as would be expected, have issues with collection of correct TFNs and addresses, often as a result of poor data from employers and members.

ITM believes that 98 per cent 'clean' members (i.e. two per cent or less with errors) should be a minimum acceptable starting point for data quality. However, the data integrity analysis shows that 24 of the 34 data audit assignments fall below this level for the Core Member Details Benchmark.

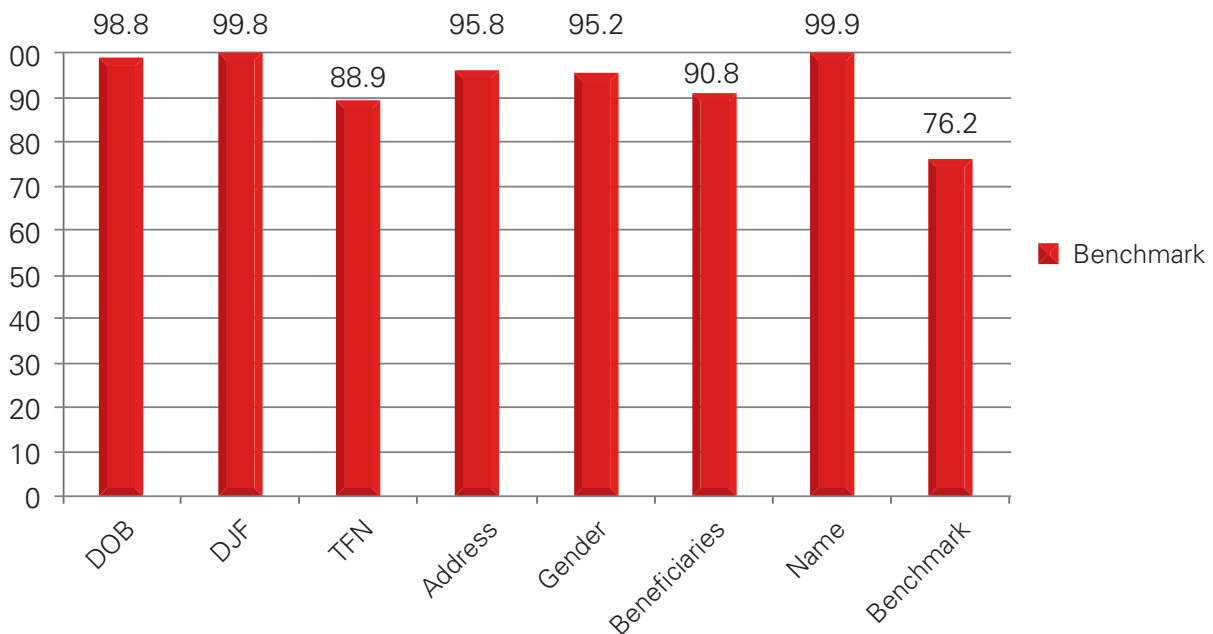
Error rates for date of birth, date joined fund and name are low.

The use of the TFN for agreed purposes has recently been accepted. Twelve funds requested ITM to check TFNs, with 12 of these having potential errors in greater than three per cent of members.

The quality of addresses information varies markedly from fund-to-fund. While it's accepted that a valid residential address is kept for each member, many funds still have a PO Box or the employer address as the only contact details. Almost 20 per cent of funds have five per cent or more members with address issues.

Date of birth is a critical field needed to confirm the correct identity of members and to ensure that benefits are correctly calculated (eg insurances, defined benefits) and paid at the right time (eg retirement, conditions of release).

Core Member Details



## **ASFA-ITM Core Member Details Benchmark Tests**

### **Address errors**

- First two lines missing where suburb, state or postcode is entered
- Invalid suburb/state/postcode combination

### **Date joined fund (DJF) errors**

- DJF is not recorded
- DJF is before fund commencement date
- DJF is after analysis date

### **Date of birth (DOB) errors**

- DOB is not recorded
- DOB is below minimum age at analysis date
- DOB is above maximum age at analysis date
- DOB is greater than DJF

### **Gender errors**

- Gender is not M or F
- Gender is F and title is Mr, Master or Sir
- Gender is M and title is Miss, Ms, Mrs, Dame or Lady

### **Name errors**

- Surname or first name not recorded
- Surname and first name contain non-literal characters

### **Tax File Number (TFN) errors**

- TFN is missing
- TFN is duplicated for members of different surnames or DOB's
- TFN is not numeric, is not eight or nine characters, or fails the Australian Taxation Office TFN algorithm
- TFN is '111 111 111' and member has a DJF more than three months ago
- TFN is '333 333 333' and the member is over age 16 or does not have an Under 16 Exemption

### **Beneficiary errors**

- Total Allocation not 100 per cent
- Nomination type not binding or non-binding
- Binding nomination expiry dates before analysis date
- Beneficiary DOB is incorrect