



28.02.2012

Mr Timothy Pilgrim
The Privacy Commissioner
Office of the Australian Information Commissioner
GPO Box 5218
SYDNEY NSW 2001

Dear Mr Pilgrim,

The Australian Taxation Office (ATO) appreciates the opportunity to provide comments on the draft *Guidelines for Data-Matching in Australian Government Administration* (the draft Guidelines). The ATO considers that it is in a position to give valuable feedback given its extensive practical experience in applying the existing guidelines. I would also like to take this opportunity to acknowledge your officers who continue to work with the ATO in a professional manner in order to expedite the publication of the data matching program protocols that are becoming increasingly critical to the business of the ATO.

Background

I would like to set the context from which the following comments were prepared. The ATO is keen to continue to conform to the Guidelines and the Privacy Principles while efficiently achieving outcomes for the community and government. Government revenue raised from data matching covered by the Guidelines has significantly increased since 2008 with a reduction in the administrative costs for collecting this revenue. Further, the number of taxpayers and tax agents using the data provided by the ATO through its pre-filling service in order to complete annual income tax returns has significantly increased. Data matching programs play a key role in this work.

The ATO is conscious of the findings of the Australian National Audit Office 2008 report¹ on data matching in the ATO (the Auditor General's report). The Auditor General highlighted the changing nature of the environment that the ATO operates in and the way it has changed its business approaches to capitalise on the emerging opportunities offered by the use of technology to detect and improve non-compliance in the community, as well as offering improved services. I would like to note that since 2008 these trends have continued.

The Auditor General's report specifically highlighted:

- The ATO's improved tax administration from the use of high quality data and the data matching and analytics capability used to provide new and better web based services at the time taxpayers complete their tax returns.
- the expansion of current ATO data matching activities that increase the need for the ATO to retain and use historical data for tax administration purposes. For example in undertaking Capital Gains Tax compliance activities that require an extended timeframe, in order to undertake business-as-usual compliance activities, and prevent repeated requests of data providers.

¹ Audit Report No.30 2007–08: The Australian Taxation Office's Use of Data Matching and Analytics in Tax Administration tabled on 28 April 2008

- changing community perceptions including greater community acknowledgement that data matching is an efficient way to improve tax compliance

In the 2008 report the Auditor General called for the review of the Guidelines recommending the review:

- ‘...usefully focus on procedural and operational matters and the administrative arrangements that apply to data matching, mining and analytics, rather than on the privacy principles...’²
- ‘...while being consistent with privacy principles be more relevant to contemporary public sector administration, including tax administration...’³

Comments

Detailed comments are provided in the attachment. Key comments are outlined below:

1. The Guidelines are well structured and easy to read and there are many areas that the ATO agrees with and will continue to take steps to improve its approach to these activities; for instance:

- the need to prepare and publicise data matching program protocols, where appropriate (Guideline 5, and paragraphs 59 – 64)
- the need to regularly evaluate all data matching programs, including on-going programs (Guideline 9)

2. The ATO is concerned that the review has not addressed recommendations made by the Auditor General that the guidelines take account of the use of data matching in contemporary administration.

3. The Guidelines introduce additional requirements in documentation, reporting and publishing that would significantly and unnecessarily increase administrative costs. The rationale for increasing the compliance burden is difficult to understand given feedback that the ATO currently satisfies a high level of compliance and cooperation.

4. The draft Guidelines were circulated widely through the ATO and the feedback received demonstrated the terms, definitions and language used were open to multiple interpretations. This raises a genuine concern that the ATO’s interpretation of the Guidelines may vary from your intention. We suggest examples (prepared in consultation with us) to demonstrate the interpretation may serve to make it easier to apply the Guidelines. Examples where it is possible for a number of interpretations (and therefore applications) to exist are below:

- Key terms (pp1-2) – match, data-set, data matching, data matching cycle and data matching program. The ATO is keen to ensure there is clarity about these terms and their interaction. For example, we interpret the interaction of these terms to mean that a data matching cycle has not been completed until the data-set is used to identify a discrepancy; that is, no match occurs when we identify the particular data-set belongs to a particular taxpayer (identity matched). This implies that data-sets that are not identity matched can be kept indefinitely.
- Paragraph 3 (the Guidelines do not apply to data matching where Tax File Numbers are used). The ATO considers that the draft Guidelines should not apply to data matching of legislative data that is covered under requirements in legislation. However, paragraph 9 in the previous Guidelines has not been replicated in this draft. The ATO relied on this paragraph to support the decision not to prepare data matching program protocols in matching data provided under legislation. The ATO has not previously prepared data matching program protocols for this type of data matching and would be concerned if a change was intended.

² Audit Report No. 30 2007-08 paragraph 2.14

³ Audit Report No. 30 2007-08 paragraph 2.14

- Should the 'boxed' Guideline title be used to interpret the subsequent paragraphs? In some instances the paragraphs following a Guideline statement appear inconsistent with that statement. For example, Guideline 7 (destroy information that is no longer required). Does the statement within the Guideline box that starts 'At the conclusion of a data-matching program...' mean that data does not need to be destroyed where data matching programs are on-going (for example the Share data and Property data matching program protocols)? Paragraph 80 of Guideline 7 also seems to support this view as it recommends '... destruction or de-identification of personal information that is no longer required...'
- The definitions and Guidelines do not appear to recognise the use of data in contemporary data modelling processes; for example in using analytic models.

5. The ATO particularly wishes to discuss the requirements in the following two guidelines, in addition to the topics highlighted in the attachment:

- **Guideline 2 (Deciding to carry out or participate in a data matching program)**
The ATO through its past experience knows and understands that data matching has proven to be a cost effective and efficient means of ensuring compliance of taxation obligations. Undertaking any further cost benefit analysis or consideration in finding other ways for determining an individual's compliance with taxation obligations is not an effective use of ATO resources.
- **Guideline 4 (Prepare a technical standards report)** The ATO currently includes information in data matching program protocols and evaluation reporting explaining many of the items listed in paragraph 50 of the draft Guidelines. Drafting technical standard reports in the detail shown in Appendix B would require extensive ATO resources as the specifications for each matching algorithm may change frequently in a data matching program or may not actually be known until the ATO has received the data. We do not believe this is an efficient and effective use of ATO resources.

Conclusion

The Guidelines are an important element in the environment as the ATO moves to implement its strategic direction of supporting people to fulfil their tax obligations at minimal cost and of protecting the community by deterring, detecting and dealing with those who have not complied. Given the sensitivity and future impact of the document I have asked that my officers arrange a discussion with you and your officers to ensure we have an agreed position on all relevant points, before this draft is finalised.

Please contact Assistant Commissioner Andrew Watson on (08) 82081826 or 0414583017 if you wish to discuss further before we meet.

Yours sincerely,

[signed]

Erin Holland
Deputy Commissioner of Taxation
Australian Taxation Office

Australian Taxation Office (ATO) Detailed Feedback to the Office of the Australian Information Commissioner's review of the Guidelines for Data-matching in Commonwealth Administration

The following are ATO detailed comments on the draft Guidelines for Data-matching in Commonwealth Administration (the draft Guidelines) referred to in the covering letter.

Topic	Guideline Reference	Guideline Reference details	ATO Comment
Application of the Guidelines to legislative data	Background	<p>Paragraph 2 & 3 - Who should use the Guidelines?</p> <p>The draft Guidelines apply to data matching used to determine whether administrative action is warranted.</p> <p>The draft Guidelines do not apply to data matching where Tax File Numbers are used.</p>	<p>The ATO is concerned that the Guidelines may apply more widely than previously.</p> <p>Paragraph 9 of The use of data matching in Commonwealth administration – February 1998, has not been included in the draft Guidelines. The ATO relied on this paragraph to support the decision not to prepare data matching program protocols in matching data provided under legislation. The ATO has not previously prepared data matching program protocols for this type of data matching and would be concerned if a change was intended. An example is Australian Transaction Reports and Analysis Centre data provided to the ATO under legislation.</p> <p>An alternative view is that, at face value, the draft Guidelines would not apply to taxation administration with data matching using Tax File Numbers. The ATO as administrators of Tax File Numbers match data to an entity and the associated unique identifier - the Tax File number.</p>
Application of the Guidelines	What scale of data matching do the Guidelines apply to?	<p>Paragraph 26-28 scale of data matching reporting and notification required</p>	<p>The draft Guidelines apply to data matching programs that include the comparison of data sets containing information about more than 5000 individuals.</p> <p>Data matching reporting requirement for data matching containing records of fewer than 5000 but more than 1000 individuals. Should the reporting requirement be for more than 1000 records to 5000 records rather than fewer than 5000 records?</p> <p>The ATO seeks your view as to the currency of these numbers. It seems the same quantum has been used since 1992.</p>

Topic	Guideline Reference	Guideline Reference details	ATO Comment
Privacy Impact Assessment	Guideline 2	Paragraph 36 The Office of the Australian Information Commissioner encourages agencies conducting a data matching project to carry out a Privacy Impact Assessment	<p>The ATO assesses privacy impacts as part of standard practices and approvals that are already undertaken when using the Commissioner's formal access powers, and in the preparation of the data matching protocols.</p> <p>The ATO would regard a formal Privacy Impact Assessment as a repetition of the privacy assessment processes already undertaken. This would also add unnecessary administrative costs.</p>
Cost Benefit Analysis	Guideline 2 Appendix C	Appendix C Statement of costs and benefits for data matching programs	<p>The ATO would like to discuss the rationale for developing a Statement of Costs and Benefits for each data matching program protocol using the level of detail described in the Appendix C. Government, the ATO, industry and the community continue to consistently recognise data matching as the most cost effective avenue for taxation administration. The Australian National Audit Office 2008 report⁴ on data matching in the ATO reiterated this as well as the benefits of data matching in taxation administration for assisting the community with better services making it easier, cheaper and simpler to comply. Data matching will continue to be the most efficient and effective use of ATO resources.</p> <p>The costs and benefits of ATO data matching are stated in unquantifiable terms in current data matching program protocols and final evaluation reports. The ATO would continue to regard this as sufficient cost and benefits analysis under the new Guidelines.</p>

⁴ Audit Report No.30 2007-08: The Australian Taxation Office's Use of Data Matching and Analytics in Tax Administration tabled on 28 April 2008

Topic	Guideline Reference	Guideline Reference details	ATO Comment
Publishing the Program Protocols	Guideline 3	<p>Guideline 3 3.1(c) Make the program protocol publicly available</p> <p>Paragraph 41 the Office of the Australian Information Commissioner recommends that the primary user agency publish the program protocol for the data matching program on its website</p>	<p>This Guideline requirement aligns with the Australian National Audit Office 2008 report⁵ on data matching in the ATO. The ATO agrees with this requirement that is supported by our current practices:</p> <ul style="list-style-type: none"> ▪ we consistently publish gazette notices of ATO data matching programs in the Commonwealth Gazette ▪ We publicise information about our data matching programs on our website ▪ It is our standard practice to provide the data matching program protocol on request (that have been minimal) as detailed in our gazette notices. ▪ Case by case we consider broader communication strategies for publicising ATO data matching programs <p>The ATO will continue to assess and take steps to improve its approach to these activities.</p>
Technical Standards report	Guideline 4	<p>Guideline 4 4.1 Before commencing a data matching program, the matching agency should prepare detailed technical standards to govern the conduct of the data matching program</p>	<p>The ATO provides the majority of the information recommended for the technical standards report in the data matching program protocol and in our evaluation reports.</p> <p>The ATO would be concerned with:</p> <ul style="list-style-type: none"> ▪ increased reporting requirements because of the added administrative costs. ▪ providing data matching specifications that frequently change. These specifications often can not be defined until the data is acquired and will be refined and improved during the matching process. The ATO would regard changes, of the nature contemplated by paragraph 56, sufficiently captured in our current protocols and evaluation reports.

⁵ Audit Report No.30 2007–08: The Australian Taxation Office’s Use of Data Matching and Analytics in Tax Administration tabled on 28 April 2008

Topic	Guideline Reference	Guideline Reference details	ATO Comment
Notify the public	Guideline 5	Paragraph 58.3 List the categories of information	It is not clear what is intended by 'categories of information'. For example, is what the ATO presently publish in gazette notices sufficient? The data categories in the gazette notice do not provide sufficient context for general understanding. The ATO would consider the information in the data matching protocols provide a better context for people in the community to understand.
Destroy information that is no longer required	Guideline 7 and relevant Definitions	<p>Paragraph 80 to 84 refer to 'key terms' and definitions for when to destroy data</p> <p>Key Terms:</p> <p>Data matching – means the bringing together of data sets that contain personal information</p> <p>Data matching Cycle – one completion of all steps of a data matching program</p> <p>Data matching Program – is data matching in accordance with clearly defined steps and criteria to achieve an objective – it may involve more than one data matching cycle</p> <p>Data-set means a discrete ordered collection of data.</p>	<p>The combination of key terms means there is not a clear point in time to apply data matching destruction time frames.</p> <p>The ATO view in relation to this Guideline is outlined below:</p> <ol style="list-style-type: none"> 1. An identity match is not a completed 'data matching cycle'. ATO automated identity matching (ie matching data to our unique identifier – the Tax File Number) that results in data not matched to a Tax File Number is not considered unmatched until it can be established that the taxpayer is operating in the tax system and meeting their tax obligations. The 'data matching cycle' is only completed at this time. The data that can not be matched to an individual will be destroyed within the 90 day period. 2. According to the definition a 'data matching cycle' is one completion of all the steps of a 'Data matching Program'. Arguably, ongoing data matching programs such as the ATO's Capital Gains data matching (shares and property) will not be complete until all the data relevant to the data matching cycle provides sufficient information to determine whether there is a match for Capital Gains Tax assessment purposes. Paragraph 80 of guideline 7 also supports a view to retain the data as it recommends destruction or de-identification of personal information when the information is no longer required. Capital Gains Tax data is still required to enable the completion of the 'data matching cycle'. 3. The terms used do not relate to more contemporary use of data. <p>The ATO is providing more assistance to the community by providing prior year information to individuals and tax agents. The expectations and requests for assistance in reconstructing records from prior years is growing.</p>

Topic	Guideline Reference	Guideline Reference details	ATO Comment
Regularly evaluate data matching programs	Guideline 9	Paragraph 90 When data matching programs are conducted over several years, agencies should periodically confirm the reasons for participating in the data matching program are still valid.	The ATO interprets the periodical review requirement to be the evaluation report to be provided every three years for ongoing data matching programs.
Data matching with entities other than agencies	Guideline 11	Paragraph 104 The agency should require that the entity adopt these Guidelines in respect of the data matching program.	<p>The ATO would like to understand what the procedure in guideline 11 is about. Is it intended to achieve something more than was required under paragraph 25 of The use of data matching in Commonwealth administration – February 1998?</p> <p>It is ATO standard practice to inform the data providers that the requirements of the Guidelines are to be adopted; however we have no formal powers to ‘require’ a data provider to adopt the guidelines.</p>
Minor Edits		Guideline 3 Paragraph 38.2.5 - a statement of costs and benefits and refers to Appendix B	This should be Appendix C that deals with costs and benefits
		Appendix C Introduction Guideline 4 calls for agencies to prepare statements of costs and benefits	Guideline 2 is the applicable guideline for the Statement of Costs and Benefits
		Guideline 6 Paragraph 74.1 dot point 2	Insert ‘initial’ before the word ‘conclusions’
		Guideline 10 Paragraph 100 Freedom of Information	Information is spelt incorrectly