The Auditor-General
Audit Report No.30 2007–08
Performance Audit

The Australian Taxation Office's Use of Data Matching and Analytics in Tax Administration

Australian Taxation Office

Australian National Audit Office

ISSN 1036–7632

ISBN 0 642 81011 7

Australian National
Audit Office

Canberra ACT
24 April 2008

Dear Mr President
Dear Mr Speaker

The Australian National Audit Office has undertaken a performance audit in the Australian Taxation Office in accordance with the authority contained in the *Auditor-General Act 1997*. Pursuant to Senate Standing Order 166 relating to the presentation of documents when the Senate is not sitting, I present the report of this audit and the accompanying brochure. The report is titled *The Australian Taxation Office's Use of Data Matching and Analytics in Tax Administration.*

Following its presentation and receipt, the report will be placed on the Australian National Audit Office's Homepage—http://www.anao.gov.au.

Yours sincerely

Ian McPhee
Auditor-General

The Honourable the President of the Senate
The Honourable the Speaker of the House of Representatives
Parliament House
Canberra ACT

AUDITING FOR AUSTRALIA

The Auditor-General is head of the Australian National Audit Office. The ANAO assists the Auditor-General to carry out his duties under the *Auditor-General Act 1997* to undertake performance audits and financial statement audits of Commonwealth public sector bodies and to provide independent reports and advice for the Parliament, the Government and the community. The aim is to improve Commonwealth public sector administration and accountability.

For further information contact:
The Publications Manager
Australian National Audit Office
GPO Box 707
Canberra ACT 2601

Telephone: **(02) 6203 7505**
Fax: **(02) 6203 7519**
Email: **webmaster@anao.gov.au**

ANAO audit reports and information about the ANAO are available at our internet address:

http://www.anao.gov.au

Audit Team
David Crossley
Richard Mackey
Kirsty Stuart
Budi Cahyono

Contents

Tables

Figures

Abbreviations

ABN	Australian Business Number
AIIR	Annual Investment Income Reports
AML/CTF	Anti-Money Laundering and Counter-Terrorism Financing
ANAO	Australian National Audit Office
ASIC	Australian Securities and Investments Commission
ATOms	Australian Taxation Office matching system
ATRO	Australian Taxation and Revenue Offices
AUSTRAC	Australian Transaction Reports and Analysis Centre
CGT	Capital Gains Tax
CKO	Chief Knowledge Officer
COAG	Council of Australian Governments
DMSC	Data Matching Steering Committee
ECI	Electronic Commerce Interface
ECMP	Easier, Cheaper and More Personalised
EDW	Enterprise Data Warehouse
GST	Goods and Services Tax
HECS/SFSS	Higher Education Contribution Scheme/Student Financial Supplement Scheme
HOTSA	Health of the System Assessment
ICP	Integrated Core Processing

ITAA	Income Tax Assessment Act
ME&I	Micro Enterprises and Individuals
OCKO	Office of the Chief Knowledge Officer
OECD	Organisation for Economic Co-operation and Development
PAYG	Pay As You Go
RFT	Request for Tender
Tax Office	Australian Taxation Office
TFN	Tax File Number

Summary and Recommendations

Summary

Introduction

1. In 2006–07 the Australian Taxation Office's (Tax Office) revenue from income tax, indirect tax, other taxes and excise totalled some $254.8 billion.[1] Australia's taxation system of self assessment places responsibility on taxpayers to declare all of their assessable income and claim only deductions and/or offsets to which they are entitled. In administering the taxation system, the Tax Office has to balance the underlying principles of self assessment with the need to manage risks, such as lack of compliance with legislation, failings in the integrity of the tax system, damage to the Tax Office's reputation, and erosion of the Commonwealth's revenue base, as efficiently and as effectively as possible.

2. In a self-assessment system the Tax Office seeks to achieve high levels of compliance by using the Compliance Model[2] of behaviour to guide the application of risk mitigation strategies. These strategies range from the communication of information to help taxpayers fulfil their obligations to criminal prosecution to stop and deter egregious non-compliance. In terms of resource investments, audits have traditionally constituted the primary form of verification adopted by revenue bodies, including the Tax Office, to check taxpayers' compliance with their obligations. However, there are practical limitations regarding the level of taxpayer and income coverage achievable from audit activities.

3. Since the 1970s the Tax Office has been developing a computer-based data matching capability which is deployed to assist the mitigation of risks associated with the management of taxpayer compliance. Traditional semi-automated data matching validates the income reported by taxpayers in their returns and discloses income that had not been reported to the Tax Office. In many cases it generates automatic correspondence to taxpayers.

---

1 Australian Taxation Office, Annual Report 2006–07, p. 283.

2 The Compliance Model (see Fig 3.1 of this Report) summarises a considerable body of knowledge about the reasons why people function the way they do in relation to society's institutional arrangements. The Model shows the most cost-effective compliance strategy that the Tax Office should adopt for a particular group of taxpayers. The Model provides a knowledge-based framework for determining the most appropriate strategy to take in relation to a compliance problem, given what is known about taxpayers, their situation, circumstances, and lines of business. The Model was first presented in 1998 in the Second Report of the Cash Economy Task Force, *Improving Tax Compliance in the Cash Economy*; ATO April 1998.

4. More recently, the Tax Office has augmented traditional, semi-automated data matching with project-based data matching and analytics[3] projects. The more focussed data matching projects are framed around categories of compliance risk and involve significant investigations by specialist tax officers. Most of the projects have a large research and development component. The analytics projects generally use advanced statistical and/or mathematical tools to examine a wide range of compliance risks. Once established after testing, some of the analytics projects involve semi-automated processes, which result in administrative action not unlike traditional semi-automated data matching.

5. As the Tax Office gains more experience of the successes and challenges of these projects it is improving an evolving generic capability. The ANAO reviewed 80 data matching and analytics projects all of which indicate the scope and potential of this generic capability. In particular, it enables the bringing together and interrogation of large volumes of data from different sources. The capability has facilitated an increased capacity to make comparisons, link data sets, compare corresponding items, find relationships and patterns and construct descriptive or hypothetical representational and/or functional relationships amongst the data variables relevant to the management of compliance risks.

6. In summary, the Tax Office has expanded its range of computer-based strategies to support the management of compliance risks. These now range from traditional semi-automated data matching to advanced analytics which create sophisticated mathematical models of varieties of serious non-compliance.

7. Within this expanding framework, the Tax Office defines traditional data matching as '*the acquisition of external data used by the Tax Office for a variety*

[3] In this Report 'analytics' is a technical term with a precise meaning. 'Analytics' means *a discipline that identifies patterns, relationships and trends from data, using a variety of mathematically based technologies principally drawn from statistics and data mining. Most broadly, analytics covers what might be called basic analytics, including data exploration and aggregation, and advanced analytics, which uses data mining technology for discovery and model building purposes. Using statistical and data mining technologies, significantly more complex relationships within and between entities (e.g. taxpayers) can be discovered and modelled, based on analyses over very large populations of collected data. Analytics is assisted by the use of good data matching and data linking techniques which improve the quality and value of data inputs available to a data miner. Conversely, analytics can also provide technology to assist data matching and data linking activities.* See further the Glossary at the end of this Report.

of different purposes such as detection of un-disclosed income, non-lodgement of returns and activity statements, and entities outside the tax system.'[4]

8. The Tax Office undertakes two categories of data matching:

- large scale post-issue semi-automated system processing (e.g. the Tax Office's automated Australian Taxation Office matching system (ATOms) automatically matches data, including legislated data[5] against data taxpayers provide directly to the Tax Office, detects discrepancies in stated income, and produces a pool of discrepant cases for compliance activity);[6] and
- data matching projects (e.g. business lines[7] initiate projects that involve acquiring and matching external data to internal Tax Office databases and producing a pool of cases for risk identification or compliance activity. See Appendix 2, Table A5 for a list of the Tax Office's data matching projects since 2000).

9. The two categories of data matching are administered by different areas of the Tax Office. The Micro Enterprises and Individuals business line is responsible for data matching in the first category. However, the Data Matching Steering Committee (DMSC) coordinates and facilitates the initiation and conduct of the data matching projects in the second category.

10. There are similarly two broad categories of analytics projects. Advanced analytics projects, largely based on data mining and modelling, are administered by the analytic group in the Office of the Chief Knowledge Officer. Basic analytics projects, using more conventional tools, are distributed broadly across most Tax Office business lines. Generically, analytics activities involve discovering relationships, patterns and trends in datasets, building,

---

4 Tax Office, External Data Matching PS CM 2004/17.

5 Legislated data refers to data third parties provide to the Tax Office, as a result of specific legislated requirements (e.g. Annual Investment Income Reports (AIIR) provided by financial institutions see Audit Report No.48 2003–04, *The Australian Taxation Office's Management and Use of Annual Investment Income Reports)*. Legislated data sets cover major income streams such as salary and wages and interest and dividends, but do not cover other income streams such as capital gains, net rents, and income from self employment. See Appendix 2, Table A2 for a list of legislated datasets the Tax Office regularly receives, and the legislation under which the data is provided.

6 The investigation of discrepant cases, which is critical to the effectiveness of ATOms relies on an investigatory effort in a way that is similar to some analytics projects. See Audit Report No.48 2003–04, *The Australian Taxation Office's Management and Use of Annual Investment Income Reports*, Chapter 5.

7 Business lines are the market segment compliance divisions and include Large Business and International, Small Business Enterprises, Micro Enterprises and Individuals, Superannuation, Serious Non Compliance, Debt, Goods and Services Tax and Excise.

testing and operating data models using data provided directly to the Tax Office, internally generated data (such as the results of compliance activities) and externally sourced datasets. (See Glossary for definitions of *advanced analytics*, *basic analytics*, *data modelling*, and *data mining*).

11. Traditionally, Tax Office data matching has had a compliance focus, with external data being used for compliance verification purposes. However, the Tax Office has also recently begun to address in a substantial way the potential of data matching and analytics to support taxpayer service needs (and to reduce taxpayers' compliance burden, perhaps significantly). The spread and power of IT systems used by business and Government, in conjunction with more secure facilities on the World Wide Web, provides significantly increased capacity for timely electronic exchanges of data and opportunities for new and improved services to tax payers. This expanding use of high quality data, especially for the provision of web-based services at the time taxpayers complete their tax returns, represents a new paradigm in tax administration.

Audit objective and scope

12. The objective of the audit was to evaluate the Tax Office's corporate management of data matching, including analytics.

13. The ANAO examined the Tax Office's strategic goals and governance arrangements for data matching and analytics, its compliance with privacy requirements and whether the Tax Office is achieving intended results, which include revenue collection, optimised compliance and provision of improved services to taxpayers.

14. Tax Office executives have been increasingly drawing on the interrelationships and conceptual commonalities of Tax Office data matching and analytics activity.[8] Accordingly, the audit included these relationships and conceptual commonalities within the scope of the audit. The audit was guided,

[8] For example, *"As our Integrated Core Processing matures, we will be able to develop more refined risk models with wider data warehousing, analytics, and data mining and matching capabilities. These capabilities will mean that we will be able to better differentiate between those taxpayers who are trying to do the right thing and those that are not; and by integrating data matching work with our risk profiling we are making sure that they are working hand-in-glove. This work happens without the Tax Office contacting people and wasting their time."* And *"In the future we will be developing more sophisticated multi-factor data matching, not just using, for example, TFN or name matches. We will match simultaneously across multiple data sources and more attributes, such as birth date and address. This wider range of factors will provide progressively higher degrees of reliability."* Michael D'Ascenzo, Commissioner of Taxation, *Simplifying Tax Administration in a Complex World: The Challenge of Infinite Variety*, Australian Tax Teachers' Association Conference, University of Queensland 22–24 January 2007.

therefore, by a broader definition of 'data matching': meaning 'finding relationships and patterns in large volumes of data'. This includes the more traditional idea of data matching as 'bringing together data from different sources and comparing it'.

15. At the conceptual level of this broader definition, traditional semi-automated Tax Office data matching is increasingly becoming an integrated element within Tax Office's more recently initiated analytics work. Accordingly, in this report, the phrase 'data matching and analytics capability' includes all the computer-based methodologies that the Tax Office applies to bring together large volumes of data from different sources, making comparisons, linking data sets, comparing corresponding items, finding relationships and patterns and constructing descriptive or hypothetical representational and/or functional relationships.

16. The audit focussed on the conduct of data matching and analytics projects by Tax Office business lines, the significant synergies and interrelationships between these projects and other Tax Office activities, such as identity matching, Australian Taxation Office matching system (ATOms) and analytics projects.

Conclusion

17. The Tax Office is making use of its data matching and analytical capabilities in a more corporate and strategic way. This has contributed to the Tax Office reporting improved compliance, better services and the more efficient and effective use of resources. It has also enabled the Tax Office to better understand risks. Procedural arrangements have enabled the Tax Office to improve the use of data matching and analytics and comply with Privacy Commissioner *Guidelines*. Contact with a number of national revenue authorities of member states of the Organisation for Economic Co-operation and Development (OECD) suggest that the Tax Office is well regarded for its work in the data matching and analytics field.

18. The analytics group in the Office of the Chief Knowledge Officer has made a valuable contribution to tax administration during the three years since its establishment. It has pioneered many new sophisticated mathematical and statistical methodologies (see Glossary) resulting in improved compliance, improved services, and the more efficient and effective use of resources.

19. As a result of these achievements the Tax Office has developed a *data matching and analytics capability* which encompasses traditional data matching and its integration into the broader area of analytics.

20. The progressive implementation of the *Easier, Cheaper and More Personalised Change Program* (*ECMP Change Program*) now under way provides the Tax Office with the opportunity to get much more from the capability than previously possible. The *ECMP Change Program* provides for the first time, an administrative and systems environment which will enable the Tax Office to better manage the data matching and analytics capability strategically and corporately.

21. Although traditional semi-automated data-matching has been a feature of tax administration since the 1970s, the Tax Office has only recently developed the more comprehensive data matching and analytics capability. There is scope to improve its corporate and strategic management of this emerging capability. The development of a strategic plan for acquiring and using external data over a three to five year period would be a key step in this direction.

22. A high level stock take of achievements and lessons learned from the deployment of the capability would also inform the identification of opportunities for the improved management of the capability at the corporate and strategic levels.

23. Additionally, there is scope to improve the design and use of the Tax Office's identity matching systems especially in two areas. One concerns non-individual identity matching, particularly in relation to companies, trusts and superannuation funds. The other concerns the implementation of a negative search[9] facility in relation to the registration of individual tax payers and the administration of superannuation. A negative search facility provides increased assurance that individuals do not exist in the data sets, thus improving the integrity of the data.

24. The Tax Office is moving into an environment which is placing a premium on the fully effective use of the data matching and analytics capability. An example of this is the pre-filling initiative. It represents a paradigm shift for the Tax Office's use of data matching and analytics. This is because it is pioneering a shift from bulk post-assessment income data

[9] A negative search facility can match and report a 'No Match' as a success; that is, establish with a specific level of confidence, that the individual does not exist in the data being searched.

matching to the pre-filling of electronic tax returns with data that would have been used in semi-automated post-assessment data matching.

25. The longer term success of pre-filling depends on the continued assistance by the suppliers of legislated data.[10] However, to be fully effective the reporting deadlines for legislated data may have to be brought forward. Also, the Tax Office's use of third party data, necessary for improved administration and achieving a higher integrity tax system, would be improved if legislation enabled attachment of the TFN to the data in a manner similar to the arrangements that apply to legislated data.

26. The legislative development of the TFN reflects the tension of maintaining an appropriate balance between individual privacy and improving the efficiency and effectiveness of public administration. Given that the determination of where the balance of this tension lies is a matter for the Government and the Parliament, the ANAO has proposed that the Tax Office develop options for legislative amendment for discussion with the Treasury in the first instance.

27. In light of the Tax Office's stronger data matching and analytics capability, there is scope for its increased exploratory use. There is scope to more efficiently and more quickly identify a range of compliance risks, including those to the tax system's integrity and to the revenue base, thereby improving the Tax Office's understanding of these risks.

28. Most compliance risks in respect of the majority of taxpayers are addressed, in line with the Compliance Model, by the provision of services such as education, assistance, information, and making compliance easier, cheaper and more personalised. Compliance risks of the most egregious type, which are depicted in the Compliance Model as being towards the top of the compliance pyramid (Fig 3.1), are considered by the Tax Office as being best addressed by personalised, swift, forceful and, if at all possible, pre-emptive, action.

29. The implementation of pre-filling, other on-line, web-based services and the *ECMP Change Program*, provides the Tax Office with the potential to

[10] Legislated data refers to data third parties provide to the Tax Office, as a result of specific legislated requirements (e.g. Annual Investment Income Reports (AIIR) provided by financial institutions - see Audit Report No.48 2003–04, *The Australian Taxation Office's Management and Use of Annual Investment Income Reports)*. Legislated data sets cover major income streams such as salary and wages and interest and dividends, but do not cover other income streams such as capital gains, net rents, and income from self employment. See Appendix 2, Table 2 for a list of legislated datasets the Tax Office regularly receives, and the legislation under which the data is provided.

reduce post-assessment processing compliance activity and provide the education, assistance and information services more cost-effectively.

30. Developments reviewed in this Report have the potential to enable the Tax Office to respond to the challenge of optimising compliance in a way that is more efficient and effective than previously possible. These developments include the continuing roll out of on-line and web-based services supported by the technological innovations of the *ECMP Change Program* and the data matching and analytics capability.

Key findings by chapter

Chapter 2: Legislative and Policy Framework

31. Chapter 2 identifies the key features of the legislative framework for the Tax Office's data matching activities, and reviews the Tax Office's adherence to privacy legislation.

32. External data the Tax Office uses for data matching and analytics (and other activities) is acquired via four main avenues, namely:

- datasets provided by an external party to the Tax Office as a result of specific legislative requirements (e.g. Annual Investment Income Reports provided by financial institutions). These legislated data sets cover major income streams such as salary and wages and interest and dividends, but do not cover other income streams such as capital gains, net rents, and income from self employment;
- datasets provided by a government agency to the Tax Office, under memorandums of understanding (e.g. state revenue offices);
- commercially available datasets purchased by the Tax Office (e.g. electoral roll data); and
- datasets requisitioned by the Tax Office using legislative authority, on a non-routine basis, for data matching and analytics projects (e.g. data is requisitioned from sellers of luxury vehicles to assist in detecting undeclared income and identify other tax risks).

33. In the past, the Tax Office sometimes used different approaches to decide which of the Commissioner of Taxation's powers were to be used in procuring data. Enhanced consultation with all business lines and the Office of the Chief Knowledge Officer during the project planning stage would assist in ensuring that data is acquired under the appropriate legislative authority, the

compliance burden on data providers is minimised, and enable the optimal use of the data by the Tax Office.

34. The Office of the Privacy Commissioner has published *The Use of Data Matching in Commonwealth Administration – Guidelines* (the *Guidelines*). These contain a number of procedures to ensure data matching activities are designed and conducted in accordance with sound privacy principles. The Tax Office has, as required by the *Guidelines*, submitted to the Privacy Commissioner *Data Matching Program Protocols* for its data matching projects, announced its intention to conduct authorised data matching projects through the *Commonwealth Gazette*, and submitted project evaluations to the Privacy Commissioner. Table 1 provides details of the Tax Office's submissions to the Privacy Commissioner, and announcements in the Commonwealth Gazette since inception of the DMSC in 2003. The ANAO notes that as compliance with the *Guidelines* is voluntary, the Privacy Commission reviews the Tax Office's submissions, providing the Tax Office with comment, not approval.

Table 1

Tax Office submissions to the Privacy Commissioner, 2003–04 to 2006–07

Tax Office data matching projects	2002–03	2003–04	2004–05	2005–06	2006–07
Data matching privacy protocols reviewed by the Privacy Commissioner	1	12	9	10	12
Notices of intention to conduct data matching published in the Commonwealth Gazette Notices	1	12	9	10	13
Requests for exemptions from data matching privacy protocols reviewed by the Privacy Commissioner	0	0	1	0	4
Evaluations of data matching projects submitted to the Privacy Commissioner	1	12	0	0	0

Source: Tax Office

35. The Tax Office could broaden the announcement of data matching projects by directly advising tax agents[11] of gazetted data matching projects as well as communicating these projects through the Tax Office's marketing and education channels.

[11] Tax agents prepare about 73 per cent of individual, and over 95 per cent of business tax returns.

36. A number of recent developments have a bearing on the applicability of the *Guidelines* to contemporary data matching practices in the Tax Office. These include rapid advances in technology, changing community perceptions, and expansion of data matching activity.

Chapter 3: data matching and analytics capability in the Tax Office

37. In this report *data matching and analytics capability* includes a range of computer-based methodologies that have the common feature of bringing together and comparing large volumes of data from different sources. The capability includes all of the computer-based methodologies that the Tax Office applies to:

- make comparisons;
- link data sets;
- compare corresponding items;
- find relationships and patterns; and
- construct descriptive or hypothetical representational and/or functional relationships of variables of which the data is composed.[12]

38. Whilst there are considerable technical differences between the algorithms which make up the methodologies used, and the particular uses to which the methodologies are put, all the methodologies at the corporate and strategic level have the common feature of bringing together and comparing large volumes of data from different sources.

39. The Tax Office applies these methodologies for a range of compliance improvement and revenue collection initiatives, to improve services, to better understand compliance risks and behaviour, and to better manage those risks.

40. Data matching and analytics capability not only consists of these computer-based methodologies but also includes the advanced professional skills of the staff who develop the software and interpret the results generated by the software, and specialised information technology software and hardware.

41. Data matching and analytics capability is a key competency or resource that the Tax Office has developed, beginning with the traditional computer-

[12] Tax Office executives have been increasingly drawing on the interrelationships and conceptual commonalities of Tax Office data matching and analytics activity.

based data matching that began several decades ago.[13] The capability supports the Tax Office's management of compliance risk having regard to the understanding of compliance behaviour summarised in the Compliance Model. The previous Commissioner of Taxation stated:

> It is clear that data matching is an efficient and effective way to help ensure the integrity of our revenue system. For this reason we are continuing to explore ways to improve our data matching capability and to focus data matching exercises on genuine risks to the community's revenue base.[14]

42. The Tax Office's data matching and analytics capability extends beyond its traditional notion of data matching (e.g. identity matching, income matching using the ATOms and DMSC facilitated data matching compliance projects) and encompasses the full spectrum of Tax Office data matching and analytics activity. The Tax Office's Compliance Model emphasises that the provision of services to taxpayers is the most appropriate Tax Office response to compliance risks presented by the majority of taxpayers. The ANAO considered there was potential for the 'provision of services' to be included in the criteria for data matching and analytics projects.

43. Recent speeches by the Tax Commissioner also take a broad view of the capability:

> As our Integrated Core Processing matures, we will be able to develop more refined risk models with wider data warehousing, analytics, and data mining and matching capabilities. These capabilities will mean that we will be able to better differentiate between those taxpayers who are trying to do the right thing and those that are not.[15]

44. Legislation sets out that quotation of the Tax File Number (TFN) is voluntary in tax administration, and the use of a unique numeric identifier is

---

[13] *Audit Report No.1 of 1995–96 Income Matching System*, reported on the Tax Office's computer based system which identified discrepancies between information in tax returns and external data, noting that the system had been in operation since late 1987.

[14] Michael Carmody, *Data Matching Improves Compliance*, 28 September 2005, <www.ato.gov.au> [accessed 20 February 2007].

[15] Michael D'Ascenzo, *Simplifying Tax Administration in a Complex World: The Challenge of Infinite Variety*, Australasian Tax Teachers Association Conference, University of Queensland, 22–24 January 2007.

not required in most non taxation situations.[16] The Tax Office has developed its data matching and analytics capability partly to address this challenge. Nevertheless, partly because of shortcomings in the quality of third party data sets, the Tax Office is sometimes unable to associate a unique identifier to all relevant records and/or to link satisfactorily particular data items in third party data sets to entries in tax returns that should correspond. Additional forensic investigations are sometimes necessary. The ANAO noted these features do not apply in a number of sovereign states that are members of the OECD.

45. The Tax Office is moving into an environment which will increasingly depend on making better use of the data matching and analytics capability. An example is the pre-filling initiative. The Tax Office began piloting the pre-filling concept using e-tax for 2004–05. This pilot was limited to pre-filling data from Medicare Australia and Centrelink. In 2005–06 the Tax Office expanded its pre-filling pilot to include the 30 per cent childcare rebate, and bank interest and managed fund information from selected financial institutions. The ANAO noted the limited range of fields and data available for pre filling in 2005–06 restricted the relevance of the pre filling pilot facility to a relatively small proportion of e-tax users. In 2006–07, approximately 1.9 million individuals used e-tax to lodge their return, and of these, approximately 1.1 million used pre-filling. A further 1.9 million pre-filled reports for individuals were downloaded by tax agents through the Tax Agent Portal.

46. Introduction of pre-filling represents a shift for the Tax Office's use of data matching and analytics. It involves, for example:

- advising the taxpayer at the time tax returns are being completed the information the Tax Office holds relevant to the completion of their tax returns;
- a shift from bulk post-assessment income data matching to pre-filling of electronic returns;[17] and

[16] A taxpayer (that is, a person, company, partnership, trust or superannuation fund) is not compelled to quote their TFN in any of their dealings with the Tax Office. This makes Australia unusual amongst nations with TFN type identifiers. Since 1990 legislation governing the use of the TFN for the receipt of most Commonwealth income support payments has required that people claiming, or receiving, this assistance have to provide a TFN as a condition of receiving such payments. TFN withholding taxes are imposed (subject to several exemptions) for not quoting a TFN for specific financial transactions (such as the interest earned on investment accounts). The TFN withholding tax imposed is at the top marginal tax rate (plus the Medicare levy) and applied to the interest earned above a minimum threshold of $120 per annum for interest earned and to each $1 earned as dividends. The top marginal tax rate (plus the Medicare levy) is imposed on the actual salary and wage income being paid for not quoting the TFN on the Employment Declaration Form regardless of the actual income. For further details see Audit Report No.37 1998-99 *Management of Tax File Numbers Australian Taxation Office*.

[17] Commissioner's online update, May 2005.

- the provision of web-based services that can only be enabled by data matching and analytics using multiple, high integrity data bases.

47. The ANAO reviewed the Tax Office's use of data acquired by different avenues: datasets provided by the Tax Office as a result of legislative requirements; datasets provided by Government agencies under memorandums of understanding; commercially available datasets; and datasets requisitioned by the Tax Office using its legislative authority. The Australian Tax Office's matching system (ATOms) automatically matches data from legislated data sources against data taxpayers provide directly to the Tax Office. ATOms detects discrepancies in stated income, and produces a pool of cases that may be selected for compliance activity. Processing statistics and amounts of additional revenue collected from automated data matching are collated and in some years have been published in the Tax Office's annual compliance program report. Table 2 summarises the Tax Office's automated data matching using ATOms for the last three available financial years (details for 2006–07 are not yet available).

Table 2

Automated data matching using Annual Investment Income Report (AIIR) and Dividend and Interest Schedule data, Pay-As-You-Go (PAYG) data (e.g. salary, lump sums, fringe benefits), Centrelink data (e.g. various benefits and allowances, pensions), and Medicare Australia data (e.g. private health insurance premiums, Medicare levy surcharge data)

Automated data matching using the income matching system	2003–04	2004–05	2005–06
- Number of individual records identity matched	33,341,168	33,772,052	33,645,672
- Number of individual records identity matched at a high level of confidence	28,222,567	28,640,996	28,603,493
- Number of individual records subject to income matching as a tax return had been lodged	17,306,225	17,643,022	17,582,283
- Number of potentially discrepant cases	320,211	345,274	132,243
- Number of discrepant cases selected for actioning	276,019	332,738	80,297
- Liabilities raised ($ millions)	141.9	203.3	80.3

Notes:

1. The numbers refer to numbers of records, not numbers of individual persons, as individuals can be matched for more than one data type.

2. Processing for 2005–06 has not yet been completed. See additional table notes in the Report.

Source: Tax Office.

Chapter 4: Governance and Coordination Arrangements

48. The administration of the data matching and analytics capability is necessarily distributed across the Tax Office as the capability is integral to tax administration. The Tax Office established the DMSC to pursue a range of objectives. These include the coordination and facilitation of best corporate results from the data matching component of the capability; supporting other coordination and consultative arrangements; encouraging best practice and intra-business line consultation. These arrangements have generally worked well and the DMSC has been pivotal to results the Tax Office has achieved from the deployment of the capability. In the future, increased coordination between the OCKO's Analytics group and the DMSC would help ensure the two areas proceed jointly and efficiently with respect to strategic planning, prioritisation of projects and community consultation.

49. The Tax Office's data matching practices are well established. The DMSC has facilitated a more corporate and strategic use of data matching as an integral component of tax administration. This has resulted in improved compliance, increased revenue, improved services and the more efficient and effective use of resources. It has also enabled the Tax Office to better understand risks. The procedural arrangements established by the DMSC have enabled the Tax Office to improve the use of data matching and comply with Privacy Commissioner *Guidelines*.

50. The Tax Office would benefit from the consolidation at the corporate level of knowledge of the scope, achievements, lessons learnt and potential of data matching and analytics. It is timely to do this because of the:

- increased incidence of data matching and analytics work over the last five years;
- planned introduction of the *ECMP Change Program*; and
- demonstrable synergies between data matching and analytics methodologies.

Recommendations

51. The ANAO made six recommendations directed at improving the Tax Office's corporate and strategic management of its data matching and analytics capability.

Summary of agency response

52. The Tax Office welcomes the Australian National Audit Office's recommendations in relation to its use of data matching and analytics in tax administration.

53. Data matching and analytics play a key role in modern tax administration. As the report acknowledges, the Tax Office is recognised as a leader in the area by tax authorities internationally.

54. The Tax Office accepts the six recommendations.

55. Developments in technology are allowing the Tax Office to move to increasingly sophisticated approaches in the use of data. The Tax Office will continue to explore opportunities for the increased use of data to improve the operation of the tax system.

56. It is encouraging to note that the Australian National Audit Office supports the shift the Tax Office is making to use data matching to support a range of service initiatives. As noted in the report, these include the expanded pre-filling initiative and the extended use of third party information (for example, real property data) to assist taxpayers in correctly preparing income tax returns, rather than using this data exclusively to support post issue compliance checks.

57. The Tax Office's full response can be found in Appendix 1.

Recommendations

Recommendation No.1

Para 2.15

The ANAO recommends the Tax Office, as the primary user of the Privacy Commissioner's Data Matching Guidelines:

- initiate policy-level discussion with the Office of the Privacy Commissioner with a view to updating the 1998 publication, "*The Use of Data Matching in Commonwealth Administration – Guidelines*"; and
- as appropriate, subsequently revise the Tax Office's data matching protocol, to reflect any changes.

ATO Response: *Agreed*

Recommendation No.2

Para 3.41

The ANAO recommends that in order to achieve additional efficiencies to tax administration and improve the integrity of key tax data bases, the Tax Office improve non-individual identity matching, give further consideration to the negative search proof-of-concept facility and continue to review options to improve its key identity matching facility.

ATO Response: *Agreed*

Recommendation No.3

Para 3.94

The ANAO recommends that to simplify income tax returns by providing pre-filling for increased numbers of taxpayers the Tax Office discuss with the Treasury options in relation to the pre-filling initiative, namely:

- bringing the dates forward for the provision of requisite third party data;
- mandating electronic transmission in a range of areas; and
- including the TFN on some additional data sets, having regard to the need to balance privacy concerns and improving the efficiency and effectiveness of public administration.

ATO Response: *Agreed*

Recommendation No.4

Para 3.121

The ANAO recommends that the Tax Office develop a three to five year corporate and strategic plan for the acquisition and use of external data.

ATO Response: *Agreed*

Recommendation No.5

Para 4.7

To minimise duplication and close gaps in responsibilities and streamline co-ordination arrangements, the ANAO recommends the Tax Office clarify the roles of the various committees relevant to the management of the data matching and analytics capability, so their individual responsibilities are integrated more closely.

ATO Response: *Agreed*

Recommendation No.6

Para 4.21

To improve the efficiency and effectiveness of the corporate and strategic management of the data matching and analytics capability and inform the further development of the capability, the ANAO recommends that the Tax Office complete a high level 'stock take' of achievements and lessons learned from its deployment.

ATO Response: *Agreed*

Audit Findings and Conclusions

1. Background and Context

This chapter provides an overview of data matching, and the Tax Office's use of data matching in tax administration, and outlines the audit objective and methodology.

The Australian Taxation Office

1.1 The Australian Taxation Office (Tax Office) is responsible for effectively managing and shaping administrative systems that support and fund services for Australians, and give effect to social and economic policy through the tax, superannuation, excise and other related systems. In 2006–07, it collected $254.8 billion (net) in tax, superannuation and excise revenue.[18] Its operating expenditure budget was $2661.8 million, and as at 30 June 2007 it employed 22 342 staff.

1.2 In a self-assessment system and with limited resources, revenue bodies need effective and efficient means for achieving very high levels of voluntary compliance. In terms of resource investments, audits have traditionally constituted the primary form of verification adopted by revenue bodies, including the Tax Office. However, there are limitations concerning the level of taxpayer and income coverage achievable from audit activities. On the other hand, large scale data matching activity in relation to the categories of income is a powerful tool for validating and disclosing much of the income reported by taxpayers in their returns.

Data matching

1.3 At its most fundamental level, data matching is the linking of datasets and comparison of corresponding data items. With respect to public administration, data matching has been defined as 'the large scale comparison of records or files of personal information, collected or held for different purposes, with a view to identifying matters of interest'.[19]

Data matching legislation and guidelines

1.4 The *Data-matching Program (Assistance and Tax) Act 1990* (Data-matching Act) regulates use of Tax File Numbers (TFN) in comparing personal

[18] Australian Taxation Office, Annual Report 2006–07, p. 283.

[19] Office of the Privacy Commissioner, *The Use of Data Matching in Commonwealth Administration–Guidelines*, February 1998.

information held by the Tax Office and by other agencies such as Centrelink and the Department of Veterans' Affairs.

1.5 For data matching activities, other than the programs specifically regulated by the Data-matching Act, the Privacy Commissioner has issued voluntary advisory guidelines for data matching in Commonwealth administration.[20] Privacy aspects of the Tax Office's data matching are discussed in *Chapter 2 – Legislative and Policy Framework.*

Data matching in the Tax Office

1.6 Data matching has been a feature of tax administration in Australia for many decades. Initially data matching in the Tax Office was a manual activity. For example, from 1936 onwards, when the Tax Office first received statements of interest from the banks, it was able to check if individual tax payers correctly reported interest earnings. During the 1970s the Tax Office moved to computer based data matching systems. In the late 1980s, after Parliament passed the TFN legislation, the Tax Office implemented the automated Income Matching System. This system has been progressively improved. The data matching system that has evolved is the Australian Taxation Office matching system (ATOms). [21]

Data acquisition and identity matching

1.7 The Tax Office's data matching capability is underpinned by the need to first obtain datasets to match, to confirm and validate the identity of entities and then to link two or more datasets to match information.

Data acquisition

1.8 Data is provided directly to the Tax Office by taxpayers (individuals and entities) and by tax agents. Almost all of the data received from taxpayers and tax agents is received via annual Income Tax Returns, and monthly and quarterly Business Activity Statements. The Tax Office also obtains significant quantities of third party data (i.e. external data) from federal and state government agencies, as well as the private sector. For example, federal

[20] The Office of the Privacy Commissioner is an independent office which has responsibilities under the *Privacy Act 1988* (Cwth). The Privacy Commissioner notes that data-matching poses a particular threat to personal privacy because it involves analysing information about large numbers of people without prior cause for suspicion.

[21] ATOms refers to a number of systems, including the Information Matching Analysis and Selection System, Case Actioning System and Compliance Online Enquiry and Amendment System.

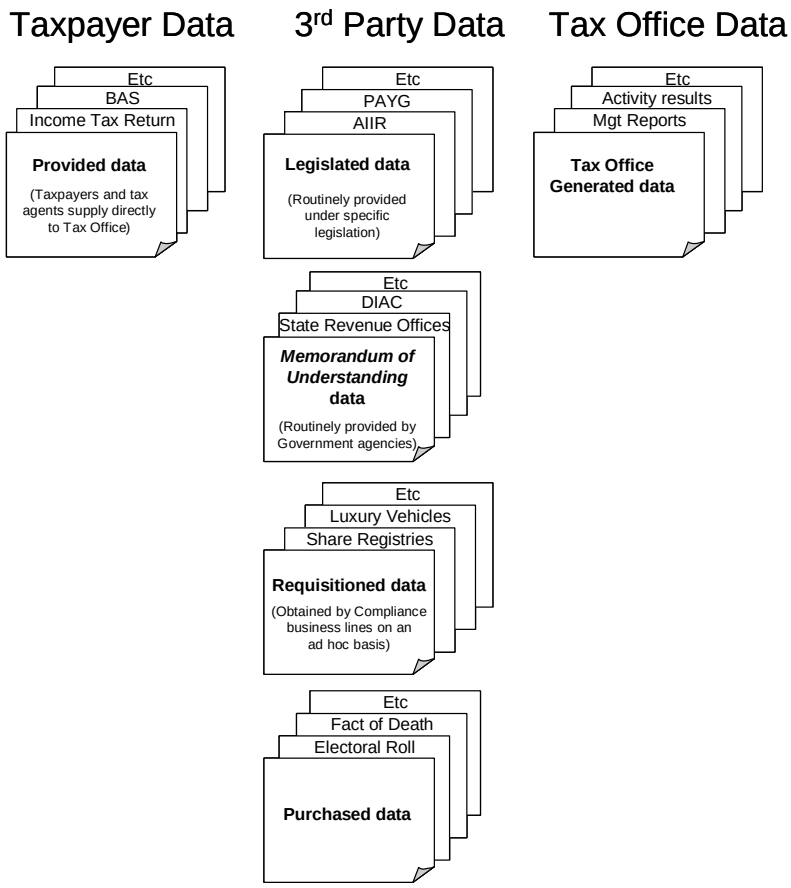
government agencies provide reports of benefits and pensions paid, while reports of wages and interest income are provided by private sector entities.

1.9 This external data may be routinely provided to the Tax Office under specific legislation, a memorandum of understanding, or requisitioned under the Tax Commissioner's information gathering authority on an ad hoc or specific project basis. Data is supplied to the Tax Office through a variety of channels, such as the Electronic Lodgement Service, Tax Agent and Business Portals, disk, or paper.

1.10 The Tax Office also uses acquired data-sets to generate some data that is used for tax administration purposes. Data acquisition is discussed in *Chapter 3 – data matching and analytics capability in the Tax Office.* Figure 1.1 depicts the categories of data received by the Tax Office.

Figure 1.1

Categories of data



Source: ANAO analysis of Tax Office information. Terms such as 'legislated data', requisitioned data' etc are defined in the Glossary at the end of this Report. Abbreviations are listed at the front.

1.11 More specific details of the data sets acquired by the Tax Office are provided in Appendix 2.

Identity matching

1.12 On receipt of third party data the Tax Office initially compares it to other Tax Office data in order to confirm and/or validate the identity of the person or entity. Data that does not contain an Australian Business Number (ABN) or TFN (which is the case for data provided by AUSTRAC and for most requisitioned data), the Tax Office attempts to attach the appropriate ABN or

TFN to the data.[22] This is undertaken by the Tax Office's identity matching facility which proceeds to match identity information such as ABN or TFN, name, date of birth, address contained in the third party data with Tax Office registration information.[23] Highest confidence matches occur with an exact match on TFN, name and date of birth. This identity matching enables the subsequent comparison of two or more datasets. Identity matching is discussed in *Chapter 3 – Data Matching and Analytics Capability in the Tax Office.*

Data matching

1.13 The Tax Office defines data matching as 'the acquisition of external data used by the Tax Office for a variety of different purposes such as detection of un-disclosed income, non-lodgement of returns and activity statements, and entities outside the tax system.'[24]

1.14 The Tax Offices undertakes two categories of data matching:

- large scale post-issue automated system processing (e.g. after the majority of tax returns have been processed, ATOms automatically matches legislated and other data against data taxpayers provide directly to the Tax Office, detects discrepancies in stated income, and produces a pool of cases for compliance activity. This occurs annually for individuals, not partnerships and trusts); and
- data matching projects (e.g. Compliance business lines[25] initiate projects that involve acquiring and matching external data to internal Tax Office databases and producing a pool of cases for risk identification or compliance activity. See Appendix 2, Table 5 for a list of the Tax Office's data matching projects since 2000).

[22] Only certain people and organisations are authorised by law to request quotation of a TFN. These include employers, some Australian Government agencies, trustees for superannuation funds, payers under the pay-as-you-go system, higher education institutions, the Child Support Agency and investment bodies such as banks. Section 202C of the *Income Tax Assessment Act 1936* authorises the Tax Office to request quotation of a TFN, however, taxpayers are not compelled to quote their TFN in any dealings with the Tax Office. In situations where payments of employment or investment income are made to taxpayers, the Tax Office requires that payers withhold tax at the highest marginal rate, plus the Medicare levy, if the taxpayer has elected not to quote their TFN and has no exemption from quoting.

[23] Since the quotation of the TFN is not mandatory in Australia, there will always be transactions and records within data bases integral to tax administration that lack this necessary numeric identifier. If the Tax Office's identity matching engine fails to establish with a sufficient level of confidence the identity of the person belonging to the transaction or record, it may be necessary for Tax Office staff to investigate these transactions or records in order to link them reliably to a person and/or entity.

[24] Tax Office, External Data Matching PS CM 2004/17.

[25] Compliance business lines include Large Business and International, Micro Enterprises and Individuals, Superannuation, Serious Non Compliance, Excise, and Goods and Services Tax.

1.15 Traditionally, data matching in the Tax Office has had a compliance checking focus, with third party data being used to verify the accuracy of information reported by taxpayers in their returns. Data matching projects within the Tax Office are coordinated through the Data Matching Steering Committee (DMSC).

Analytics

1.16 The Tax Office also undertakes a range of analytics activities (e.g. discovering relationships, patterns and trends in datasets) using data provided directly to the Tax Office, internally generated data (such as the results of compliance activities) and externally sourced datasets. The advanced analytics work based on data mining techniques primarily occurs within the analytics group, within the Office of the Chief Knowledge Officer (OCKO); whereas basic analytics work using more conventional tools is distributed broadly across most business lines. The OCKO is a relatively recently created area, preceded by the DMSC.

1.17 The Tax Office undertakes this work to improve the understanding, and support the management of compliance, especially serious non-compliance, risks, provide better services and improve the efficiency and effectiveness of tax administration.

Recent ANAO audit reports

1.18 Data matching is an essential feature of contemporary tax administration. Recent ANAO Audit Reports have evaluated data matching in various areas of tax administration. These included:

- Audit Report No.16 2006–07, *Administration of Capital Gains Tax Compliance in the Individuals Market Segment*;
- Audit Report No.12 2006–07, *Management of Family Tax Benefit Overpayments*;
- Audit Report No.35 2005–06, *The Australian Taxation Office's Administration of Activity Statement High Risk Refunds*;
- Audit Report No.33 2005–06, *Administration of Petroleum and Tobacco Excise Collections, follow-up*;
- Audit Report No.17 2005–06, *Administration of the Superannuation Lost Members Register*;

- Audit Report No.47 2004–05, *Australian Taxation Office Tax File Number Integrity* (see also Audit Report No.37, 1998–99, *Management of Tax File Numbers*);
- Audit Report No.39 2004–05, *The Australian Taxation Office's Administration of the Superannuation Contributions Surcharge*; and
- Audit Report No.48 2003–04, *The Australian Taxation Office's Management and Use of Annual Investment Income Reports*.

Recent developments

Easier, Cheaper and More Personalised Change Program

1.19 The Tax Office is currently implementing major changes to its business processes by means of an initiative called the *Easier, Cheaper and More Personalised (ECMP) Change Program*. In its 2006–07 annual report, the Tax Office stated that the ECMP represented a $453 million investment in new technology and business re-engineering. The purpose of the ECMP is to make taxpayer interactions with the Tax Office easier, cheaper and more personalised, within the parameters allowed by tax and superannuation laws.[26] The ECMP originated in the Tax Office's 2002 *Listening to the Community* program, when the Tax Office worked with the general community, small business, industry and tax agents to develop ideas to make improve the Tax Office. The Change Program business line, Business Solutions business line, and the Chief Knowledge Officer have responsibility for delivering elements of the *ECMP Change Program*.

Change Program business line

1.20 The Change Program business line is developing new standardised business systems and processes, which will be applicable across the Tax Office. It is also developing new information technology systems, which are being delivered in stages. In the first two stages, the Tax Office introduced a new Client Relationship Management system and implemented a new single Case Management system. In the third stage, the Tax Office will replace a range of computer processing systems with a single, integrated, system for all tax types. This system is currently referred to as an Integrated Core Processing (ICP) system. The ICP will comprise one registration system, one accounting system and one processing system covering all taxes administered by the Tax Office.

[26] Australian Taxation Office, 2006, *2006–07 Making it Easier to Comply, The easier, cheaper and more personalised program.*

The Tax Office plans to implement the ICP progressively throughout 2008 and 2009.

Business Solutions business line

1.21 The Business Solutions business line has the responsibility to ensure the design and implementation of all *ECMP Change Program* technology is aligned across the Tax Office. It is also responsible for ensuring that external stakeholders understand the *ECMP Change Program*, and leading whole-of-government and international initiatives, such as the standard business reporting (see paragraph 1.28).

Office of the Chief Knowledge Officer

1.22 The Office of the Chief Knowledge Office (OCKO) is responsible for developing a vision and strategy for the management of information within the Tax Office. This includes improving the Tax Office's capability in the area of information management. The OCKO is also responsible for developing the Tax Office's strategic intelligence and analytics capability.

ECMP Change Program and data matching

1.23 The ECMP will have an effect on the Tax Office's data matching and analytic activities as it involves developing new identity matching functionality, creating a single information technology system that stores whole-of-taxpayer data, and implementing automated case analysis and selection processes. Furthermore, the achievement of *ECMP Change Program* objectives may be facilitated by data matching and analytic activities. The *ECMP Change Program's* themes include: treating taxpayers as individuals; helping taxpayers manage their tax affairs in the way that suits them; saving time through online interactions; increasing certainty for taxpayers; and seamlessly integrating tax with everyday business activities. Significantly, the Tax Office aims to use technology to enable it to become a more service and customer-focused organisation.[27] The use of data originally obtained for semi-automated post assessment data matching to pre-fill tax returns before completion by a taxpayer, is one example of how the holistic management and coordination of data matching in the Tax Office, including within the *ECMP Change Program*, enables the Tax Office to realise the potential of its data matching and analytics capability.

[27] Second Commissioner Greg Farr 5 April 2006. *"Making it easier, cheaper and more personalised - dealing with the Australian Taxation Office",* Government Technology World 2006.

Pre-filling tax returns

1.24 The automatic pre-filling of fields in a tax return or record keeping tool by a revenue authority is designed to assist taxpayers meet their tax obligations. Based on experience in countries where the practice is advanced, pre-filling of tax returns can provide substantial benefits to the taxpayer and revenue authority, including a reduced compliance burden for taxpayers, a relatively easy (automated) process to advise adjustments, fewer taxpayer errors, simpler and less costly administration, and improved image of the revenue body.[28]

International experiences

1.25 Pre-filling originated in Denmark in the late 1980s and spread over the following decade to other Nordic region countries (e.g. Sweden and Finland).[29] The practice is now well established and regarded as highly successful. In more recent years, other countries (e.g. Chile, France, and Spain) have proceeded to adopt the practice.

1.26 In the Nordic region countries, most citizens (around 18 million) receive a pre-filled tax return (on paper and/or via the Internet). Given the comprehensiveness of the third party reporting systems in place in these countries, revenue bodies are able to pre-fill much of the tax return for the majority of taxpayers, thus substantially reducing their compliance burden. For many taxpayers, all that is required is that they examine their return and confirm its accuracy and completeness with the revenue body. The potential for pre-filling varies from country to country, and is influenced by tax policy design features and other factors in each jurisdiction. However, there are a number of critical success factors that have been identified from the experience of Nordic region countries that generally apply for successful pre-filling. These include the use of high integrity taxpayer identifiers in all third party reports, comprehensive and timely third party reporting obligations (e.g. covering major income, deduction, and asset-related items of information) and the wide use of technology by the reporting bodies.

---

[28] The Head of Tax Administration Division, OECD Centre for Tax Policy and Administration, Richard Highfield, 2006. *Pre-populated Income Tax Returns: The Next 'Big Thing' in Reform of the Administration of Australia's Personal Income Tax System?* Keynote address to 7th International Tax Administration Conference, ATAX, University of New South Wales, Australia, 20–21 April 2006.

[29] OECD Centre for Tax Policy and Administration, 2006. *Using Third Party Information Reports to Assist Taxpayers Meet their Return Filing Obligations – Country Experiences With the Use of Pre-populated Personal Tax Returns.*

Pre-filling of income tax returns for individuals by the Tax Office

1.27 The Tax Office began trialling the pre-filling of electronically lodged (i.e. e-tax) 2004–05 income tax returns for individuals using Centrelink payment summary information and medical expenses recorded by Medicare Australia. In the *Budget Speech 2007–08*, the then Treasurer announced the Tax Office would introduce pre-filling for the nine million taxpayers who currently lodge their tax returns electronically (either directly or through a tax agent).[30] Pre-filling would be available in 2008–09 (for the 2007–08 income tax year's returns), and pre-filled information would include income from salary and wages, interest, dividends, information on private health insurance, and any benefits paid from the Government, including the family tax benefit. The then Treasurer stated that pre-filling would dramatically simplify income tax returns, and provided the Tax Office with an additional $20 million in 2007–08 for its implementation. Pre-filling is examined in *Chapter 3 – Data matching and analytics capability in the Tax Office*.

Standard business reporting

1.28 In August 2006, the then Treasurer created a committee of Australian and state government officials to examine the case for introducing standardised business reporting. Standard business reporting aims to reduce the reporting burden for business through eliminating unnecessary or duplicated reporting, and improving the interface between business and government. On 30 August 2007, the then Treasurer announced that the government would spend $208 million over the next three years implementing a new method for standardised business reporting to government. This project is being coordinated by the Treasury and will facilitate participating agencies to review and streamline their requirements for business reporting. While the benefits of standardised business reporting are intended to primarily accrue to businesses, significant benefits are expected for government agencies, including the Tax Office, from increased data sharing opportunities and improved data quality.

Australian Law Reform Commission Review of Privacy

1.29 In January 2006, the Australian Law Reform Commission (ALRC) received Terms of Reference from the then Australian Attorney-General for an inquiry into the extent to which the *Privacy Act 1988* (Cwlth) and related laws

[30] Treasurer of the Commonwealth of Australia, *Budget Speech 2007–08*, 8 May 2007 on the second reading of the Appropriation Bill (no. 1) 2007–08, p. 6.

continue to provide an effective framework for protecting privacy in Australia. The review was initiated given:

- the rapid advances in information, communications, storage, surveillance and other relevant technologies;
- possible changing community perceptions of privacy and the extent to which it should be protected by legislation;
- the expansion of state and territory legislative activity; and
- emerging areas that may require privacy protection.

1.30 The ALRC has produced two issues papers, and a discussion paper. A final report is due to be released in early 2008. Any subsequent changes to privacy legislation may affect the Tax Office's data matching activities.

Anti-Money Laundering and Counter-Terrorism Financing Act 2006

1.31 The *Anti-Money Laundering and Counter-Terrorism Financing (AML/CTF) Act 2006* received Royal Assent on 12 December 2006. It imposed a number of obligations on reporting entities when they provide designated services. On 13 July 2007, the Attorney-General's Department released draft provisions setting out designated services that the further amendments of the *AML/CTF Act 2006* will cover. These provisions are expected to extend the designated services to include a range of professional and business services, including real estate services. This may enable the Tax Office to utilise additional external data for its data matching activities.

Council of Australian Governments

1.32 The Council of Australian Governments (COAG) is the peak intergovernmental forum in Australia. COAG's role is to initiate, develop and monitor implementation of policy reforms that are of national significance and that need cooperative Australian government action.

Personal property securities reform

1.33 In April 2007 COAG agreed, in principle, to establish a national system for registering personal property securities.[31] The Government announced in the 2007–08 Federal Budget that it would provide $113.3 million over five years (including $2.1 million in 2006–07) to harmonise Australia's personal property

[31] Personal property is property other than buildings or land (which may include, for example, boats, vehicles and other valuable consumer goods).

security laws in one Commonwealth act and develop a single national online register of personal property security interests. The Attorney-General's Department is responsible for providing legal and policy advice across Government on personal property securities and for establishing a national system for personal property securities registration.

National Identity Security Strategy

1.34 In April 2007, under COAG commitment, the Commonwealth, state and territory governments signed an agreement to implement a National Identity Security Strategy. The Strategy recognises that excess, redundant or false records in agencies' identity registers can undermine revenue collection; provision of low cost services; and obligations under the *Privacy Act* to retain accurate personal details. The Tax Office, in partnership with the Attorney-General's Department, is conducting the data matching pilot to test the integrity of a sample of 25 000 records from the Tax Office, with these records being matched with data held by Centrelink, Medicare Australia and NSW Registry of Births, Deaths and Marriages. These agencies will attempt to match Tax Office records with information held within their database and provide a rating that relates to the degree to which the record could be matched. The Tax Office will then compile the results from all three agencies. The aim of the Pilot is to identify key data matching elements that can be used to improve the accuracy of personal information held on government databases. Due to privacy concerns, the results of the data matching process cannot be used to undertake administrative or compliance action.

1.35 A Privacy Impact Assessment was developed for this program and is available from Attorney-General's Department. A Gazettal notice was published in May 2007 advising the public that this pilot was to occur.

National Information Sharing Strategy and National Address Management Framework

1.36 The Online and Communications Council, which operates within COAG protocols for Ministerial Councils, is the peak cross-government ministerial forum to consider and reach agreement on strategic approaches to information and communications issues of national importance. The Online and Communications Council agreed to the development of a National Information Sharing Strategy that would provide the processes and tools to enable jurisdictions to share information while preserving their legislative obligations and safeguarding citizens' privacy. It is also progressing development of a National Address Management Framework that the

Australian Government and State and Territory Governments will use for address verification services.[32]

1.37 These developments are designed to identify key data matching elements and will improve data sharing across government as a whole. This will affect the Tax Office's data matching activities as it may (subject to a 'fitness for purpose analyses' by the Tax Office) provide additional external data sources and provide avenues to improve the integrity and quality of its internal database.

Recent changes in superannuation legislation

1.38 Under the Government's superannuation simplification reforms, from 1 July 2007, there will be significant changes to the tax treatment of superannuation, in the areas of:

- tax-free benefits for people aged 60 and over;
- limits on concessional and non-concessional contributions to super;
- tax deductions for the self employed;
- treatment of eligible termination payments;
- tax file number reporting requirements for superannuation; and
- unclaimed superannuation reporting.

1.39 The Tax Office is increasing its efforts to reunite taxpayers with lost superannuation funds, by improving its identification and notification processes for the superannuation Lost Members Register. Data matching is a critical activity for the Lost Members work.

Australian Taxation Revenue Office Coordination Forum

1.40 The Australian Taxation and Revenue Offices (ATRO) Coordination Forum provides a means to strengthen the relationship between the Tax Office and state and territory revenue agencies,[33] and allow for joint advancement of mutually beneficial projects, such as interoperability. Commencing in July 2005, the Tax Office, and participating state and territory revenue offices signed a Service Level Agreement and a Memorandum of Understanding for

[32] Online and Communications Council, Fourteenth Communiqué, 27 June 2007.

[33] They include the ACT Revenue Office, the NSW Office of State Revenue, the NT Revenue Management, the Queensland Office of State Revenue, the Revenue SA, the Tasmanian State Revenue Office, the Victorian State Revenue Office, and the Western Australia Office of State Revenue.

the Exchange of Confidential Information and Other Aspects of Cooperation. As a result, the Tax Office has improved its processes for acquiring land titles data from the state and territory revenue offices for its capital gains tax activities, including for the Tax Office's data matching projects.[34]

Audit overview

1.41 In a self-assessment system the Tax Office seeks to achieve high levels of compliance by using the Compliance Model[35] of behaviour to guide the application of risk mitigation strategies. These strategies range from the communication of information to help taxpayers fulfil their obligations to criminal prosecution to stop and deter egregious non-compliance. In terms of resource investments, audits have traditionally constituted the primary form of verification adopted by revenue bodies, including the Tax Office, to check taxpayers' compliance with their obligations. However, there are practical limitations regarding the level of taxpayer and income coverage achievable from audit activities.

1.42 Since the 1970s the Tax Office has been developing a computer-based data matching capability which is deployed to assist the mitigation of risks associated with the management of taxpayer compliance. Traditional semi-automated data matching validates the income reported by taxpayers in their returns and discloses income that had not been reported to the Tax Office. In many cases it generates automatic correspondence to tax payers. It can also bring to notice discrepancies in the information provided by taxpayers in their tax returns and authoritative relevant third party data. As noted in *Chapter 3 – data matching and analytics capability in the Tax Office*, its potential to support taxpayer service needs (and to significantly reduce taxpayers' compliance burden) is only now starting to be addressed in a substantial way.

1.43 The spread and power of IT systems used by business and Government, in conjunction with more secure facilities on the World Wide

---

[34] The State Revenue Offices have agreed to provide property data to the Tax Office every six months. See ANAO Audit Report No.16 1006–07 *Administration of Capital Gains Tax Compliance in the Individuals Market Segment*.

[35] The Compliance Model (see Fig 3.1 of this Report) summarises a considerable body of knowledge about the reasons why people function the way they do in relation to society's institutional arrangements. The Model shows the most cost-effective compliance strategy that the Tax Office should adopt for a particular group of taxpayers. The Model provides a knowledge-based framework for determining the most appropriate strategy to take in relation to a compliance problem, given what is known about taxpayers, their situation, circumstances, and lines of business. The Model was first presented in 1998 in the Second Report of the Cash Economy Task Force, *Improving Tax Compliance in the Cash Economy*; ATO April 1998.

Web, provides significantly increased capacity for timely electronic exchanges of data and opportunities for new and improved services to tax payers. This expanding use of high quality data, especially for the provision of web-based services at the time taxpayers complete their tax returns, represents a new paradigm in tax administration.

Audit objective and scope

1.44 The objective of the audit was to evaluate the Tax Office's corporate management of data matching, including analytics.

1.45 The ANAO examined the Tax Office's strategic goals and governance arrangements for data matching and analytics, its compliance with privacy requirements and whether the Tax Office is achieving intended results, which include revenue collection, optimised compliance and provision of improved services to taxpayers.

1.46 Tax Office executives have been increasingly drawing on the interrelationships and conceptual commonalities of Tax Office data matching and analytics activity.[36] Accordingly, the audit included these relationships and conceptual commonalities within the scope of the audit. The audit was guided, therefore, by a broader definition of 'data matching': meaning 'finding relationships and patterns in large volumes of data'. This includes the more traditional idea of data matching as 'bringing together data from different sources and comparing it'.

1.47 At the conceptual level of this broader definition, traditional semi-automated Tax Office data matching is increasingly becoming an integrated element within Tax Office's more recently initiated analytics work. Accordingly, in this report, the phrase 'data matching and analytics capability' includes all the computer-based methodologies that the Tax Office applies to bring together large volumes of data from different sources, making comparisons, linking data sets, comparing corresponding items, finding

[36] For example, *"As our Integrated Core Processing matures, we will be able to develop more refined risk models with wider data warehousing, analytics, and data mining and matching capabilities. These capabilities will mean that we will be able to better differentiate between those taxpayers who are trying to do the right thing and those that are not; and by integrating data matching work with our risk profiling we are making sure that they are working hand-in-glove. This work happens without the Tax Office contacting people and wasting their time."* And *"In the future we will be developing more sophisticated multi-factor data matching, not just using, for example, TFN or name matches. We will match simultaneously across multiple data sources and more attributes, such as birth date and address. This wider range of factors will provide progressively higher degrees of reliability."* Michael D'Ascenzo, Commissioner of Taxation, *Simplifying Tax Administration in a Complex World: The Challenge of Infinite Variety*, Australian Tax Teachers' Association Conference, University of Queensland 22-24 January 2007.

relationships and patterns and constructing descriptive or hypothetical representational and/or functional relationships.

1.48 The audit focussed on the conduct of data matching and analytics projects by Tax Office business lines, the significant synergies and interrelationships between these projects and other Tax Office activities, such as identity matching, ATOms and analytics projects.

1.49 The audit did not cover whole-of-government data matching issues, such as the Tax Office's relationships with external government agencies, or implications of the then proposed Access Card.

Audit methodology

1.50 The ANAO conducted fieldwork in the Tax Office's Canberra, Melbourne, Sydney and Brisbane offices between March and June 2007. This included examination of relevant Tax Office documentation and interviews with staff from the Operations, Large Business and International, Micro Enterprises and Individuals, Superannuation, Serious Non Compliance, Excise, and Goods and Services Tax business lines.

1.51 The ANAO sought feedback from external stakeholders, including:

- the Office of the Privacy Commissioner – on the Tax Office's compliance with the privacy requirements as stated in the federal *Privacy Act 1988*, the *Data-matching Program (Assistance and Tax) Act 1990* and the advisory *Guidelines for the Use of Data-Matching in Commonwealth Administration*;
- Treasury – on the standard business reporting project; and
- Geoscience Australia – on COAG initiatives, such as interoperability and the national address management framework.

1.52 The ANAO engaged an external advisory panel to provide expert comment on the audit scope, processes, and reporting.

1.53 This audit was undertaken in conformance with ANAO auditing standards and cost $512 976.

Acknowledgements

1.54 The ANAO would like to thank Tax Office management and staff for their assistance in the conduct of the audit. In particular, we would like to

thank our liaison officers for their considerable time, effort and expertise contributed during the audit.

1.55 The ANAO would also like to express appreciation to the stakeholders consulted for their contribution to the audit.

2. Legislative and Policy Framework

This chapter identifies the key features of the legislative framework for the Tax Office's data matching activities, and reviews the Tax Office's adherence with privacy legislation.

Authority for data matching

Tax Office's legislative authority for acquiring data

2.1 Prior ANAO reports that examined the tax file number system provide details of the legislative and policy framework that regulates Tax Office data matching and related activities.[37] Tax legislation places responsibility on taxpayers to declare all of their assessable income and claim only deductions and/or offsets to which they are entitled in deriving their income. A range of economic transfer payments are also provided to taxpayers through the taxation system.[38] Individual and business taxpayers provide information directly to the Tax Office (e.g. through an income tax return or a business activity statement), and this information is stored in the Tax Office's databases. Tax returns and activity statements provide the essential core data concerning the assessment of taxpayers' taxable income. The Tax Office aims to verify this core data using, among other techniques, various data matching activities. This occurs on an annual financial year cycle.

2.2 Datasets regularly provided directly to the Tax Office, and the legislation under which the data is provided, are listed in Appendix 2, Table 1. On receipt of information from taxpayers, the Tax Office verifies the information, and identifies discrepancies, anomalies and indicators of risk. The Tax Office does this by matching new information against previously verified (and therefore high integrity) Tax Office and external data.

2.3 The external data the Tax Office uses for data matching (and other activities) is acquired via four main avenues, namely:

- Datasets an external party provides to the Tax Office, as a result of specific legislated requirements (e.g. Annual Investment Income

[37] Audit Report No.47 2004–05, *Australian Taxation Office Tax File Number Integrity* and Audit Report No.37 1998–99, *Management of Tax File Numbers*.

[38] These include grants, benefits, tax offsets and redistribution programs.

Reports (AIIR) provided by financial institutions).[39] These legislated data sets cover major income streams such as salary and wages and interest and dividends, but do not cover other income streams such as capital gains, net rents, and income from self employment. See Appendix 2, Table 2 for a list of legislated datasets the Tax Office regularly receives, and the legislation under which the data is provided.

- Datasets a Government Agency provides to the Tax Office, under a memorandum of understanding (e.g. State Revenue Offices). The datasets the Tax Office regularly receives under memoranda of understanding are listed in Appendix 2, Table 3.
- Commercially available datasets the Tax Office purchases (e.g. electoral roll data). The datasets the Tax Office purchases are listed in Appendix 2, Table 4 and;
- Datasets the Tax Office requisitions, on a non-routine basis, for data matching projects (e.g. data is requisitioned from sellers of luxury vehicles to address undeclared income and other tax risks). Further examples of specific datasets the Tax Office acquires, and the Commissioner of Taxation's legislated access and information gathering powers that provide the authority for this data acquisition are listed in Appendix 2, Table 5.

2.4 The Tax Office's legislative authority essentially derives from two areas of the law:

- specific provisions in legislation requiring annual reporting of certain incomes (e.g. wages, interest, dividends);[40] and
- the Tax Commissioner's information gathering powers.

2.5 Memoranda of Understanding covering the acquisition of data, where prepared, are made under the legal authority of the Commissioner's powers. Conversely, some external data sets are acquired on a normal commercial basis without regard to any legislative authority.

[39] See Audit Report No.48 2003–04, *The Australian Taxation Office's Management and Use of Annual Investment Income Reports*.

[40] In this Report, data provided to the Tax Office because of these provisions is referred to as legislated data.

Data matching projects – external data requisitioned using legislated access and information gathering powers

2.6 In the past, the Tax Office used different approaches to decide which of the Commissioner of Taxation's powers were to be used in procuring data. Some projects used a specific legislative authority, listed a narrow use for the data, and only obtained data fields relevant to their projects so as not to be seen to be 'fishing' for data or impose additional burden on data providers. This approach limited the ability of other areas within the Tax Office to use the data obtained. It may have necessitated a second approach to the provider if the data was also needed by other areas. Other projects deliberately sought access under all of the Taxation Commissioners powers, and listed broadly-defined criteria for use of the data, in order to maximise the ability of other areas to use the data for tax administration. Enhanced consultation between all relevant business lines and the Chief Knowledge Officer (CKO) during the project planning stage would could help make sure that data are acquired under the appropriate legislative authority, potentially minimise the compliance burden on data providers, and would enable optimal the use of the data by the Tax Office.

Compliance with privacy requirements

2.7 The Office of the Privacy Commissioner is an independent Office which has responsibilities under the federal *Privacy Act 1988* (Cth) to promote an Australian culture that respects privacy. The Privacy Commissioner notes that data-matching poses a particular threat to personal privacy because it involves analysing information about large numbers of people without prior cause for suspicion.

Using data matching in Commonwealth administration

2.8 The Office of the Privacy Commissioner published a booklet, *The Use of Data Matching in Commonwealth Administration – Guidelines* (the Guidelines), which contains a number of procedures to ensure data matching activities are designed and conducted in accordance with sound privacy principles. The Guidelines are not legally binding; however the Tax Office has voluntarily undertaken to ensure all data matching projects that procure third party data comply with the Guidelines. The Tax Office also endeavours to make sure that data matching using data obtained through a memorandum of understanding complies with the Guidelines. However, any data matching the Tax Office conducts using legislated data falls outside the scope of the Guidelines.

Protocols, gazette announcements and evaluations

2.9 The Tax Office has, as required by the Guidelines, submitted to the Privacy Commissioner *Data Matching Program Protocols* for its data matching projects, announced its intention to conduct authorised data matching projects through the Commonwealth Gazette, and submitted project evaluations to the Privacy Commissioner. Details of the Tax Office's submissions to the Privacy Commissioner, and announcements in the *Commonwealth Gazette* since inception of the DMSC in 2003–04, are listed in Table 2.1. The ANAO notes that as compliance with the *Guidelines* is voluntary, the Privacy Commission reviews the Tax Office's submissions, providing the Tax Office with comment, not approval.

Table 2.1

Tax Office submissions to the Privacy Commissioner, 2003–04 to 2006–07

Tax Office data matching projects	2002–03	2003–04	2004–05	2005–06	2006–07
Data matching privacy protocols					
submitted to the Privacy Commissioner	1	12	9	10	12
reviewed by the Privacy Commissioner	1	12	9	10	12
Notices of intention to conduct data matching published in the Commonwealth Gazette Notices	1	12	9	10	13
Requests for exemptions from data matching privacy protocols					
submitted to the Privacy Commissioner	0	0	1	0	4
reviewed by the Privacy Commissioner	0	0	1	0	4
Evaluations of data matching projects					
required to be submitted to the Privacy Commissioner	1	13	0	0	0
submitted to the Privacy Commissioner	1	12	0	0	0

Source: Tax Office

2.10 The Tax Office also announces data matching activities on its website. As the Tax Office is using data matching to improve voluntary compliance, community awareness of data matching could be beneficial to the achievement of this objective. The ANAO considered that the Tax Office should broaden its

data matching announcement by directly advising tax agents[41] of gazetted data matching projects as well as communicating these projects through the Tax Office's marketing and education initiatives.

Security

2.11 The Guidelines also require the Tax Office to take precautions to protect taxpayers' information, including the secure storage and proper disposal of information. Tax Office employees must also comply with Practice Statement CM 2004/07, *Secrecy and Privacy Obligations for Tax Office Employees*, when undertaking data matching activities. Internal Tax Office reports identified scope for improvement in the Tax Office's compliance with security requirements, including management of updating and destroying external data stored on the data warehouse. The Tax Office subsequently initiated a range of administrative reforms to improve management of the data warehouse. This included placing the data warehouse under CKO supervision; establishing improved reporting processes for validating movement of, and access to, third party data on the data warehouse; and reorganising administration of the data warehouse by the CKO. The Tax Office advised it has also commenced a Data Stewards' Program to provide improved management of the Tax Office's Data assets, which the ANAO understands includes security of external data used in data matching.

Update of the Guidelines

2.12 As the Guidelines were developed in 1998, there are a number of factors that have a bearing on their applicability to contemporary data matching practices in the Tax Office[42], namely:

- rapid advances in technology:
  - increasing use of web-based systems for taxation administration, including the Tax Agent Portal, Business Portal, and e-tax;
  - an increasing capacity for the Tax Office to adopt a more comprehensive and holistic approach in addressing compliance risks;

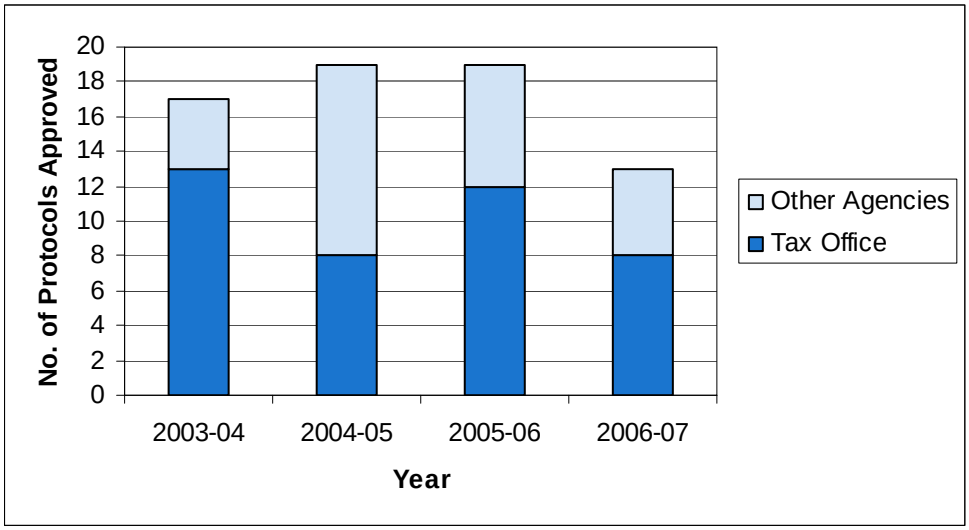
[41] Tax agents prepare about 73 per cent of individual, and over 95 per cent of business tax returns.

[42] The ANAO notes that the ALRC is currently inquiring into the extent to which the *Privacy Act 1988 (Cwth)* and related laws continue to provide an effective framework for protecting privacy in Australia. The Attorney-General initiated this review due to rapid advances in technology, changing community perceptions, expanded government activity, and emerging areas that may require privacy protection.

- changing community perceptions:
  - greater community acknowledgement that data matching is an efficient way to improve tax compliance, especially if it is addressing a complex risk or an area of non-compliance;
  - heightened community concerns about identity fraud and theft;
- expansion of data matching activity:
  - the Tax Office conducts the most data matching programs of any Commonwealth agency (see Figure 2.1);
  - an increasing need for the Tax Office to retain and use historical data for tax administration purposes and to use data for a multiplicity of data matching projects;[43] and
  - an increased role in tax administration of data mining, mapping, modelling and analytics activities.

Figure 2.1

Tax Office and federal government agency submissions to the Privacy Commissioner, 2003–04 to 2006–07



Source: ANAO analysis of Privacy Commissioner Annual Reports

[43] For example, the Tax Office has sought exemptions from the normal 90 day time limit for the destruction of data, to enable data matching project leaders to retain data for the life of projects. In particular, property data required for undertaking Capital Gains Tax compliance activities requires an extended timeframe, in order to undertake business-as-usual compliance activities, and prevent repeated requests for data from providers.

2.13 These developments indicate that the Guidelines issues by the Office of the Privacy Commissioner, to guide data matching in Australian Government agencies, would benefit from a review and subsequent update. Since 2004–05 the Tax Office has made five successful requests to the Privacy Commissioner to obtain exemptions from the Guidelines and to retain data for extended periods. The Tax Office continues to comply with the Guidelines.

2.14 In view of the foregoing considerations and developments, the ANAO raised with the Tax Office and the Office of the Privacy Commissioner the need to review the suitability of the Guidelines, with a view to drafting revised guidelines, consistent with privacy principles, which are more relevant to contemporary public sector administration, including tax administration. The Tax Office has previously raised with Privacy Commissioner the need to review the Guidelines. The proposed review could 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.

Recommendation No.1

2.15 The ANAO recommends the Tax Office, as the primary user of the Privacy Commissioner's Data Matching Guidelines:

- initiate policy-level discussion with the Office of the Privacy Commissioner with a view to updating the 1998 publication, "*The Use of Data Matching in Commonwealth Administration – Guidelines*"; and
- as appropriate, subsequently revise the Tax Office's data matching protocol, to reflect any changes.

2.16 **ATO Response**: *Agreed*. The matter will again be raised with the Privacy Commissioner as part of the established meeting program.

3. Data matching and analytics capability in the Tax Office

This chapter presents an analysis of the Tax Office's use of its capability in the areas of data matching and analytics, including key arrangements to oversight, co-ordinate, promote and evaluate the use of the capability, and assess the main ways in which it is operationally deployed. This chapter also details the pre-filling of income tax returns initiative.

Introduction

3.1 Tax Office executives have been increasingly drawing on the interrelationships and conceptual commonalities of Tax Office data matching and analytics activity.[44] Accordingly, the audit scope included these relationships and conceptual commonalities. The audit was guided, therefore, by a broader definition of 'data matching': meaning 'finding relationships and patterns in large volumes of data'. This includes the more traditional idea of data matching as 'bringing together data from different sources and comparing it'.

3.2 The ANAO noted that the Tax Office has a range of computer based activities that come within the ambit of this definition. These activities include the identity matching facility, ATOms, the data matching projects coordinated by the DMSC , and the variety of projects classified as analytics (see *Glossary*).

A Tax Office capability

3.3 In this Report the 'data matching and analytics capability' includes all the computer-based methodologies that the Tax Office applies to bring together large volumes of data from different sources, make comparisons, link data sets, compare corresponding items, find relationships and patterns and

[44] For example, *"As our Integrated Core Processing matures, we will be able to develop more refined risk models with wider data warehousing, analytics, and data mining and matching capabilities. These capabilities will mean that we will be able to better differentiate between those taxpayers who are trying to do the right thing and those that are not; and by integrating data matching work with our risk profiling we are making sure that they are working hand-in-glove. This work happens without the Tax Office contacting people and wasting their time."* And *"In the future we will be developing more sophisticated multi-factor data matching, not just using, for example, TFN or name matches. We will match simultaneously across multiple data sources and more attributes, such as birth date and address. This wider range of factors will provide progressively higher degrees of reliability."* Michael D'Ascenzo, Commissioner of Taxation, *Simplifying Tax Administration in a Complex World: The Challenge of Infinite Variety*, Australian Tax Teachers' Association Conference, University of Queensland 22–24 January 2007.

constructing descriptive or hypothetical representational and/or functional relationships.

3.4 Whilst there are considerable technical differences between the algorithms which make up the computer-based methodologies used and the particular uses to which the methodologies are put, all the methodologies have the common feature of bringing together and comparing large volumes of data from different sources. The Tax Office applies these methodologies in different ways and by different administrative procedures for a range of compliance and service improvement initiatives, and to better understand compliance risks and behaviour.

3.5 The data matching and analytics capability not only consists of these computer-based methodologies but also includes the advanced professional skills of the staff who develop the software and interpret the results generated by the software, and specialised information technology software (including proprietary and 'off-the-shelf' software) and hardware.

3.6 A paradigm shift in tax administration is noted in which high quality data and the data matching and analytics capability would be used to provide new and better web-based services at the time taxpayers complete their tax returns. This is a pioneering shift from bulk post-assessment income data matching to the pre-filling of electronic tax returns with data that would have been used in semi-automated post-assessment data matching. A feature of the paradigm shift is to use the capability to achieve compliance by the provision of services, especially web-based services, rather than by post-assessment verification. The paradigm shift is expected to make compliance easier, cheaper and more personalised. The capability will continue support risk management through the usual wide range of data matching and analytics activity that takes place after the receipt of tax returns and related information.

3.7 Although traditional semi-automated data-matching has been a feature of tax administration since the 1970s,[45] the Tax Office has only recently developed the more comprehensive data matching and analytics capability. The capability supports the Tax Office's management of compliance risk,

[45] Audit Report No.1 of 1995–96, *Income Matching System*, reported on the Tax Office's computer based system which identified discrepancies between information in tax returns and external data, noting that the system had been in operation since late 1987.

which is based on compliance behaviour summarised in the Compliance Model.[46] The previous Commissioner of Taxation stated:

> It is clear that data matching is an efficient and effective way to help ensure the integrity of our revenue system. For this reason we are continuing to explore ways to improve our data matching capability and to focus data matching exercises on genuine risks to the community's revenue base.[47]

3.8 The Tax Office's data matching and analytics capability extends beyond that traditionally considered to be data matching (e.g. identity matching, income matching using the ATOms and DMSC facilitated data matching compliance projects) and encompasses the full spectrum of Tax Office data matching and analytics activity. The Tax Office's compliance model emphasises that the provision of services to taxpayers may be the most appropriate Tax Office response to the compliance risks presented by the majority of taxpayers who seek to comply with their tax obligations.[48] The provision of educative services is generally considered as part of Tax Office risk mitigation plans.

3.9 Increasingly, as the Tax Office's *ECMP Change Program* matures with the evolving data matching and analytics capability, the provision of a service may be used more readily as a key strategic response to compliance risks, such as taxpayers misunderstanding their obligations. In this context, the ANAO considered that the administrative arrangements that operate across business lines to harness these developments, and make the most of the potential of *ECMP Change Program* and the emerging capability, could be more integrated.

3.10 The Tax Commissioner has taken a broad view of the capability:

> As our Integrated Core Processing matures, we will be able to develop more refined risk models with wider data warehousing, analytics, and data mining and matching capabilities. These capabilities will mean that we will be able to

---

[46] The Compliance Model (see Fig 3.1 of this Report) summarises a considerable body of knowledge about the reasons why people function the way they do in relation to society's institutional arrangements. The Model shows the most cost-effective compliance strategy that the Tax Office should adopt for a particular group of taxpayers. The Model provides a knowledge-based framework for determining the most appropriate strategy to take in relation to a compliance problem, given what is known about taxpayers, their situation, circumstances, and lines of business. The Model was first presented in 1998 in the Second Report of the Cash Economy Task Force, *Improving Tax Compliance in the Cash Economy*; ATO April 1998.

[47] Michael Carmody, *Data Matching Improves Compliance*, 28 September 2005, <www.ato.gov.au> [accessed 20 February 2007].

[48] "Provision of services" includes web-based services such as pre-filling, prompts, reminders and specialised calculators embedded into web-based facilities, computer generated alerts, reminder and warning letters, educational campaigns in national or specialised media and telephonic communications.

better differentiate between those taxpayers who are trying to do the right thing and those that are not.[49]

3.11 Legislation requires that quotation of the Tax File Number (TFN) is voluntary in tax administration, and the use of a unique numeric identifier is not required in most non taxation situations.[50] The Tax Office has developed its data matching and analytics capability partly to address this challenge. Nevertheless, partly because of shortcomings in the quality of third party data sets, the Tax Office is sometimes unable to associate a unique identifier to all relevant records and/or to link satisfactorily particular data items in third party data sets to entries in tax returns that should correspond. Additional forensic investigations are sometimes necessary. Relevant legislation provides that quotation of the TFN is optional for taxpayers in respect of a range of financial transactions. As a result, some income reports provided to the Tax Office do not record a TFN. A taxpayer (that is, a person, company, partnership, trust or superannuation fund) is not required to quote their TFN in any of their dealings with the Tax Office. However, since 1990 legislation governing the use of the TFN for the receipt of most Commonwealth income support payments has required that people claiming or in receipt of this assistance, have to provide a TFN as a condition of receiving such payments.[51]

3.12 Some major income streams (e.g. rents, assessable government benefits, sales of assets with possible capital gains tax implications) are not required to be reported systematically to the Tax Office by third parties. Where such data

---

[49] Michael D'Ascenzo, *The Pursuit of Simplicity – Simply Impossible,* Australasian Tax Teachers Association Conference, University of Queensland, 22–24 January 2007.

[50] A taxpayer (that is, a person, company, partnership, trust or superannuation fund) is not compelled to quote their TFN in any of their dealings with the Tax Office. This makes Australia unusual amongst nations with TFN type identifiers. Since 1990 legislation governing the use of the TFN for the receipt of most Commonwealth income support payments has required that people claiming, or receiving, this assistance have to provide a TFN as a condition of receiving such payments. TFN withholding taxes are imposed (subject to several exemptions) for not quoting a TFN for specific financial transactions (such as the interest earned on investment accounts). The TFN withholding tax imposed is at the top marginal tax rate (plus the Medicare levy) and applied to the interest earned above a minimum threshold of $120 per annum for interest earned and to each $1 earned as dividends. The top marginal tax rate (plus the Medicare levy) is imposed on the actual salary and wage income being paid for not quoting the TFN on the Employment Declaration Form regardless of the actual income. For further details see Audit Report No.37 1998-99 *Management of Tax File Numbers Australian Taxation Office*.

[51] TFN withholding taxes are imposed (subject to several exemptions) for not quoting a TFN for specific financial transactions (such as the interest earned on investment accounts). The TFN withholding tax imposed is at the top marginal tax rate (plus the Medicare levy) and applied to the interest earned above a minimum threshold of $120 per annum for interest earned and to each $1 earned as dividends. The top marginal tax rate (plus the Medicare levy) is imposed on the actual salary and wage income being paid for not quoting the TFN on the Employment Declaration Form regardless of the actual income. For further details see Audit Report No.37 1998-99 *Management of Tax File Numbers Australian Taxation Office*.

is requisitioned, existing legislation would preclude the reporting body from having access to the TFN, requiring that other less effective matching criteria must be used. The ANAO noted these features do not apply in a number of sovereign states that are members of the OECD.[52]

Data matching and analytics: administrative framework

3.13 Australia's taxation system of self assessment places responsibility on taxpayers to declare all of their assessable income and claim only deductions and/or offsets to which they are entitled. In administering the taxation system, the Tax Office has to balance the underlying principles of self assessment with the need to manage risks, such as lack of compliance with legislation, failings in the integrity of the tax system, damage to the Tax Office's reputation, and erosion of the Commonwealth's revenue base, as efficiently and as effectively as possible.[53]

3.14 In a self-assessment system the Tax Office seeks to achieve high levels of compliance by using the Compliance Model of behaviour to guide the application of risk mitigation strategies. The Compliance Model, which is depicted in Fig 3.1, summarises a considerable body of knowledge about the reasons why people function the way they do in relation to society's institutional arrangements. The Model shows the most cost-effective compliance strategy that the Tax Office should adopt for a particular group of taxpayers. The Model provides a knowledge-based framework for determining the most appropriate strategy to take in relation to a compliance problem, given what is known about taxpayers, their situation, circumstances, and lines of business. The Model was first presented in 1998 in the Second Report of the Cash Economy Task Force, *Improving Tax Compliance in the Cash Economy* ATO, April 1998.

3.15 Risk mitigation strategies range from the communication of information to help taxpayers fulfil their obligations to criminal prosecution to stop and deter egregious non-compliance. In terms of resource investments, audits have traditionally constituted the primary form of verification adopted by revenue bodies, including the Tax Office, to check taxpayers' compliance

[52] OECD Centre for Tax Policy and Administration, 2006. *Using Third Party Information Reports to Assist Taxpayers Meet their Return Filing Obligations – Country Experiences With the Use of Pre-populated Personal Tax Returns.* See <http://www.oecd.org/dataoecd/42/14/36280368.pdf> especially appendix 1.

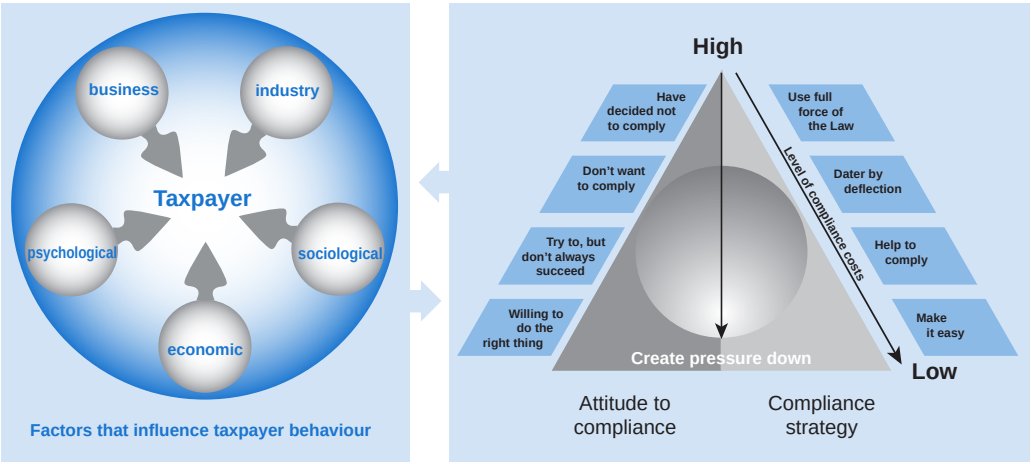
[53] There is also a range of legislative measures, including withholding taxes, designed to mitigate some compliance risks.

with their obligations. However, there are practical limitations regarding the level of taxpayer and income coverage achievable from audit activities.

3.16 The Tax Office's compliance work is guided by the *Taxpayers' Charter* and the *Compliance Model* (Figure 3.1). The Tax Office has committed to administer the tax system fairly by helping people 'do the right thing',[54] by making it as easy as possible for taxpayers to comply, and by identifying people who are not meeting their obligations and dealing with them appropriately.[55] Its business intent is to 'to optimise voluntary compliance and make payments under the law in a way that builds community confidence'.[56]

Figure 3.1

Tax Office's compliance model



Source: Tax Office Compliance Program 2006–07.

3.17 The Tax Office publishes the Compliance Program annually. Data matching and analytics have been identified in the last four compliance programs as a key post assessment compliance activity.

3.18 The *Tax Office Strategic Statement 2006–2010* stated that one of the purposes of data matching and analytics is to enable the Tax Office to make better risk management choices and to differentiate between those who want

[54] For example, in a recent speech the Commissioner of Taxation stated "We treat taxpayers fairly in accordance with the law, and at the same time we have strategies to deter, detect and address non-compliance based on risk management. Ultimately, this is to ensure that taxpayers who do the right thing are not at a competitive disadvantage relative to others". "Making tax less taxing for small business" *Speech by Michael D'Ascenzo, Commissioner of Taxation, to the Council of Small Business Organisations of Australia, Melbourne, Tuesday 27 June 2006;* p. 1.

[55] See http://www.ato.gov.au/corporate/content.asp?doc=/Content/5704.htm

[56] Australian Taxation Office, 2006, *Strategic Statement 2006–10*, p. 3.

to do the right thing and those who do not. The Statement advises that the Tax Office aims to be invisible to compliant taxpayers, other than when they require help but will be highly visible to those who choose not to comply.

3.19 In the Tax Office the identification of strategic processes to manage compliance risks takes place within a broad framework of strategic planning, risk identification and assessment. Data matching and analytics activities may form a component of these strategic responses. In addition, data matching and analytics projects have invariably informed the identification of compliance risks and the shaping of strategic processes to mitigate these risks.

3.20 The Tax Office has a formal 'Health of the System Assessment' (HOTSA) process that informs the development of the Compliance Program and a range of risk mitigation plans. The HOTSA process assesses the 'health' of the revenue system in a holistic context, considering components such as particular revenue types (e.g. GST, income tax), market segments (e.g. individuals, large businesses) or risk areas (e.g. aggressive tax planning).

3.21 The Tax Office planning processes address the major classes of risks identified by the HOTSA process. The detailed management of particular risks may occur in a number of Tax Office business lines and will often include data matching and analytic activities. The DMSC assists business lines with the coordination of proposed data matching projects. The DMSC also encourages business lines to consider the merits of data matching activities as a response to identified risks. A comparable process for the coordination of analytics projects across the Tax Office or for facilitating the use of analytics in risk responses had not been established.

Data matching and analytics: methodologies

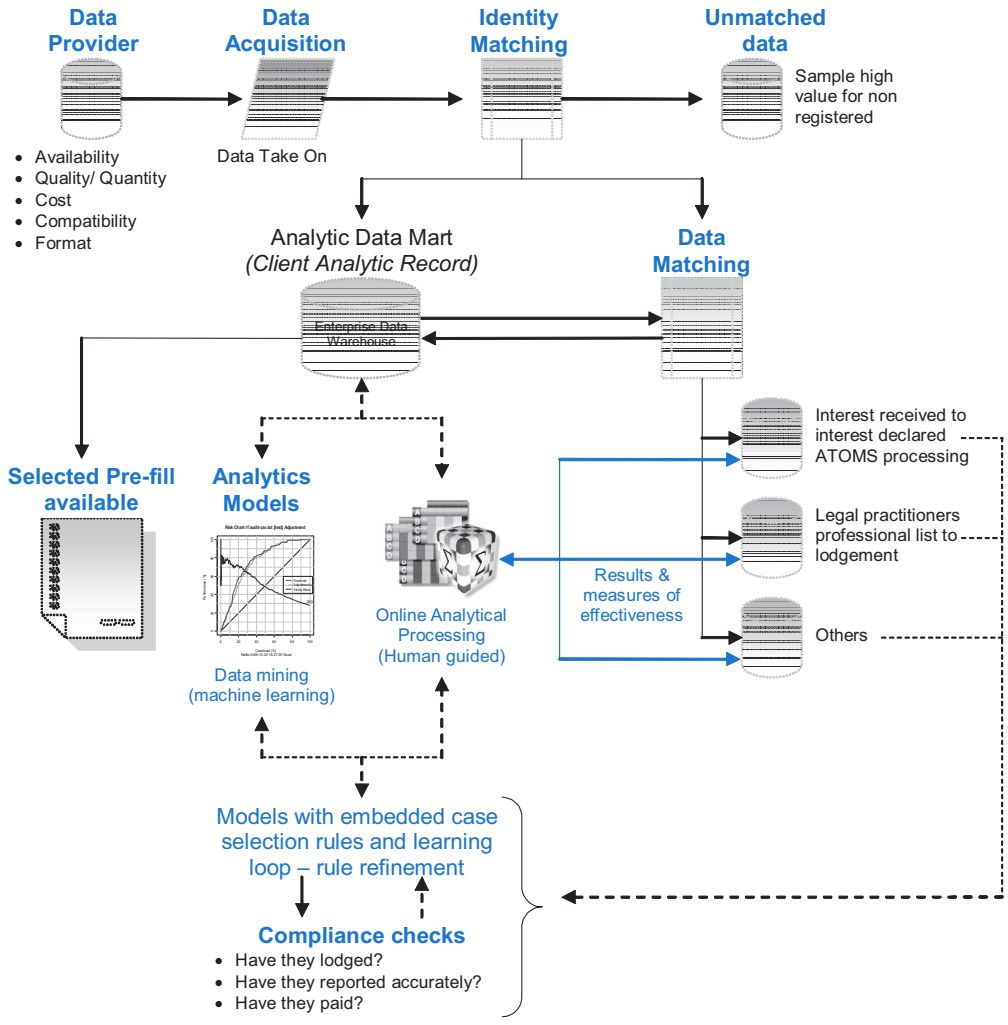
3.22 Data matching and analytics includes the identity matching systems and ATOms, the data matching projects coordinated by the DMSC and the variety of analytics projects administered by the OCKO. The Business Solutions business line also conducted an evaluation of the Tax Office's identity matching facility. It is designing a new identity matching facility and replacement components for ATOms.

3.23 The DMSC consists of representatives from all compliance business lines, the operations business line and the OCKO. The DMSC meets quarterly, and facilitates the use of data matching, the acquisition and management of data, and the effective use of data acquired by business lines.

3.24 An overview of the identity matching systems, ATOms, the data matching projects coordinated by the DMSC and the analytics projects is presented in the following paragraphs. Figure 3.2 depicts the broad relationships between data matching, identity matching, pre-filling and analytics activities.

Figure 3.2[57]

Tax Office data matching and analytics capability



Source: Tax Office

Identity matching

3.25 In Audit Report No.48 2003–04, *The Australian Taxation Office's Management and Use of Annual Investment Income Reports*, the ANAO found the

[57] Figure 3.2 shows the Tax Office's broad approach to the identification of data matching and analytics strategies as appropriate responses to risks and the consequential action of data acquisition. ANAO is satisfied that this conceptual diagram generally depicts actual practice. Furthermore, the ANAO noted that Tax Office practice in relation to the diagram has improved over recent years and that the initiatives of the DMSC have assisted in this regard.

Tax Office's identity matching system of computer programs, particularly their phonetic scoring capacity,[58] to be highly effective. There is international recognition of the Tax Office's expertise in this area (especially by national revenue authorities of some member states of the OECD) and that there has been increasing whole-of-government interest in identity matching.[59]

3.26 The identity matching system provides the Tax Office with a corporate level identity matching service and can be used to enable a broad range of matching related activities. Further, this system is essential in an environment where source data is provided by the public, or might not contain identifiers such as TFNs or ABNs. As previously noted, relevant legislation provides that quotation of the TFN is optional for taxpayers in respect of a range of financial transactions. A taxpayer (that is, a person, company, partnership, trust or superannuation fund) is not required to quote their TFN in any of their dealings with the Tax Office. As a result, there can be transactions and records within relevant data bases that lack this numeric identifier. In such cases the Tax Office has a requirement for a sophisticated identity matching system that can establish, with a sufficient level of confidence, the identity of the person or entity belonging to the transaction or record. Some transactions or records may ultimately require investigation by Tax Office staff in order to link reliably to a person or entity.

3.27 The Tax Office's matching engine has capacity to process up to 400 000 000 transactions annually. In 2007 it processed just over 220 000 000 transactions, including all of the AIIR data, as well as Centrelink statements, Pay As You Go (PAYG) data, family tax benefit, superannuation, Health Insurance Commission transactions, Baby Bonus, Higher Education Contribution Scheme/Student Financial Supplement Scheme (HECS/SFSS) transactions and a large range of ad hoc projects, including routine matching with the Australian electoral roll.

3.28 The matching routines are applied to both individual and non-individual records. The engine produces high match rates with or without identifiers, with match rates typically exceeding 90 per cent on individual data streams and 80 per cent on non-individual data streams.

---

[58] This is achieved through use of a fully customised edit list (name dictionary) capable of identifying typing and spelling errors, ethnic name anglicisation issues, orthographic and phonetic misinterpretation, truncation and abbreviation, missing words and extra words and word sequence variation.

[59] See Audit Report No.48 2003–04, *The Australian Taxation Office's Management and Use of Annual Investment Income Reports*; paragraph 2.26, p. 39.

3.29 The Tax Office developed a prototype, or proof of concept, negative search facility to support the Tax Office's registration of individuals. The Tax Office had established the need for a negative search facility in relation to the registration of individuals.[60] It was for this reason that the prototype was developed. The expectation was that it would establish with a known degree of confidence that a new individual registrant was not on the Tax Office database.

3.30 A negative search facility would enable the software to conclude that the searched-for item is not on a database. In identity matching a successful negative search concludes that the identity sought is definitely not on the database. As there is a degree of asymmetry between Tax Office and third party databases, a negative search facility would conclude (with a given level of confidence) whether a name on the Tax Office database was not on third party data bases and conversely, whether a name on a third party database was not on the Tax Office data base. Given the asymmetry, a positive conclusion - that a name on one data base is on another (with a given level of confidence) - is different from a negative one – that a name on one data base is not on the other (with a given level of confidence). One of the values to tax administration that a negative search facility provides at the time of registration, for example, is to ensure a new registrant is not an existing one, thus improving the integrity of tax data bases.

3.31 In the absence of a negative search facility, discrepant cases have to be resolved by individual manual investigation. The ANAO considered that the new environment of the *ECMP Change Program's* ICP, pre-filling and other possible web-based services requires the highest level integrity of data bases obtainable. Therefore incorporation of a negative search facility is a key aspect of achieving that integrity.

3.32 The Tax Office advised that the further development of the negative search proof of concept prototype was put on hold, pending the developments outlined in the next section in relation to superannuation administration.

Improvements to the Tax Office's identity matching

3.33 In 2006 the Tax Office sought tenders to enhance existing identity matching methodology. The Tax Office wished to examine new generation identity matching computer software that would take maximum advantage of

[60] A 'negative search' facility can match and report a 'No Match' as a success; that is, establish with a specific level of confidence, that the entity does not exist in the data being searched.

the new technology of the *ECMP Change Program*, provide a more agile solution and be sustainable in the longer term.

3.34 A feature of the *ECMP Change Program* is that all processing will take place on the one system, ICP. All third party data would be placed on the ICP, as well. In moving away from multiple data sources, the Tax Office considered that there would be overall economies of scale if identity matching could be integrated into the ICP methodology, as was planned for all other computer-based core tax administration systems. In addition, the Tax Office anticipated that the identity matching capability would require greater flexibility in future years as additional third party data sets were added to the capability and new services, based on enhanced identity matching, would be required.

3.35 At the time that the Tax Office was determining this strategy, the Government announced reforms to superannuation which required, amongst other things, that the Tax Office expand its approach to finding lost members of superannuation funds and improve the integrity of the administration of superannuation. As superannuation administration was being integrated into the ECMP, the Tax Office decided to leverage off the opportunity and develop the new approach to identity matching in the Superannuation Business Line in the first instance.

3.36 The Tax Office's existing identity matching facility required enhancement to support the identification and notification of Australians with superannuation fund accounts of which they were unaware. The Tax Office sought a general identity matching facility that could satisfy the specific needs of the Lost Members Register and then be extended to support income tax administration in the future. The Tax Office was hopeful commercial off-the-shelf identity matching software in the marketplace would meet its needs. However, the Tax Office's evaluation of the tenders submitted concluded none of the products submitted significantly or clearly offered the potential for exceeding the outcomes obtained with the current Tax Office identity matching facility. As a result, the Tax Office is proceeding internally to add to the current identity matching facility.[61] These enhancements will be used in the ICP.

3.37 According to the Tax Office, the new identity matching software will augment the existing identity matching system on which all Tax Office identity matching now takes place.

[61] The ANAO understands the Tax Office may again seek tenders, in a few years time, for identity matching facilities to enhance or replace the software now being developed.

3.38 The new identity matching software the Tax Office is now developing will not, however, include a negative search facility. Even though the Tax Office identified this as a highly desirable feature of the required identity matching functionality, the Tax Office decided that in the absence of a suitable off-the-shelf negative search facility, it would only proceed, at this stage, with the internal design of mandatory, not highly desirable, features.

Scope for further improvement to the Tax Office's identity matching

3.39 The ANAO considered there is scope to improve the efficiency and effectiveness of the current general purpose identity matching facilities. The ANAO considered non-individual identity matching, especially in relation to trusts, companies and superannuation funds could be improved considerably. There is also scope to improve individual identity matching through including additional cascading routines and improved use of archived addresses. In addition, there may be scope to make important improvements to the identity matching facility by closer working relationships between the analytics group in the CKO's business line and the Client Identification Compliance Directorate in the ME&I business line. Implementation of these improvements could, among other things, ensure the better use of available data, especially legislated data.

3.40 The provision of on-line, web-based services, such as pre-filling, requires very high quality data. A negative search facility can improve the quality of data by giving assurance that an individual does not exits on a data base. In addition, the need for a negative search facility in the Tax Office's general purpose identity matching software had been established. The further consideration of the benefits of bringing the negative search prototype into production, given the present hiatus, is warranted. Further consideration of adapting the prototype to satisfy the requirements of the Superannuation Business Line is also warranted. The ANAO noted that the more wide spread use of a negative search facility in tax administration requires more research and development.

Recommendation No.2

3.41 The ANAO recommends that in order to achieve additional efficiencies to tax administration and improve the integrity of key tax data bases, the Tax Office improve non-individual identity matching, give further consideration to the negative search proof-of-concept facility and continue to review options to improve its key identity matching facility.

3.42 ATO Response: *Agreed*. The Tax Office will continue to look at opportunities to improve its identity matching capability in respect of both individual and non-individual taxpayers. Cost benefit considerations have precluded the development of a negative search facility. The Tax Office will continue to review the negative search proof of concept as new software and technology improvements and resources become available.

The Australian Taxation Office's matching system

3.43 The Australian Tax Office's matching system (ATOms) automatically matches data from legislated data sources against data taxpayers provide directly to the Tax Office.[62] ATOms detects discrepancies in stated income, and produces a pool of cases that may be selected for compliance activity. With implementation of the *ECMP Change Program*, the Tax Office will progressively replace a number of the systems that comprise ATOms, and case analysis and selection will be made using operational analytics.

3.44 Processing statistics and amounts of additional revenue collected from automated data matching are collated and in some years have been published in the Tax Office's annual compliance program report. Table 3.1 provides details of the Tax Office's automated data matching using ATOms for the last three available financial years (details for 2006–07 are not yet available).

[62] ATOms refers to a number of systems, including the Information Matching Analysis and Selection, Case Actioning System and Compliance Online Enquiry and Amendment System.

Table 3.1

Automated data matching using legislated data sources

Automated data matching using the income matching system	2003–04	2004–05	2005–06
Annual Investment Income Report and Dividend and Interest Schedule data			
- Number of individuals identity matched	10,364,485	10,674,374	10,609,227
- Number of individuals identity matched at a high level of confidence	9,168,898	9,498,095	9,426,953
- Number of individuals subject to income matching as a tax return had been lodged	6,834,851	7,079,577	6,977,535
- Number of potentially discrepant cases	129,747	144,779	56,065
- Number of discrepant cases selected for actioning	112,928	143,783	30,335
- Liabilities raised ($ millions)	$45.6	$66.5	$29.3 (to 7 Feb 08)
Pay-As-You-Go (PAYG) data (e.g. salary, lump sums, fringe benefits)			
- Number of individuals identity matched	9,956,905	10,115,280	10,459,616
- Number of individuals identity matched at a high level of confidence	9,473,933	9,656,167	10,002,133
- Number of individuals subject to income matching as a tax return had been lodged	8,471,855	8,597,048	8,775,822
- Number of potentially discrepant cases	54,477	59,614	36,151
- Number of discrepant cases selected for actioning	45,949	59,263	22,123
- Liabilities raised ($ millions)	$32.1	$61.9	$27.7 (to 7 Feb 08)
Centrelink data (e.g. various benefits and allowances, pensions)			
- Number of individuals identity matched	5,628,326	5,552,119	5,462,710
- Number of individuals identity matched at a high level of confidence	4,653,355	4,573,995	4,479,026
- Number of individuals subject to income matching as a tax return had been lodged	1,946,149	1,907,019	1,828,926
- Number of potentially discrepant cases	82,617	81,503	40,027
- Number of discrepant cases selected for actioning	66,679	77,312	27,839
- Liabilities raised ($ millions)	$39.6	$44.2	$23.3 (to 7 Feb 08)
Medicare Australia (e.g. private health insurance premiums, Medicare levy surcharge data)			
- Number of individuals identity matched	7,391,452	7,430,279	7,114,119
- Number of individuals identity matched at a high level of confidence	4,926,381	4,912,739	4,695,381
- Number of individuals subject to income matching as a tax return had been lodged	53,370	59,378	Not yet available
- Number of potentially discrepant cases	53,370	59378	
- Number of discrepant cases selected for actioning	50,463	52,380	Not yet
- Liabilities raised ($ millions)	$24.6	$30.7	available

Notes:

(a) The 'Number of individuals identity matched' is the number of individuals identity matched to high, medium or low levels of confidence, not the total number of data records for those individuals. To obtain this figure, data records received were sorted to produce a count of unique TFNs. This produced a count of individuals that can be aligned with the statistics for cases actioned.

(b) The 'Number of individuals identity matched at a high level of confidence' is the number of matched individuals whose data records passed established criteria that allow them to be used with confidence in automated processes that identify potential discrepancies between income reported to the Tax Office by third parties and that declared within the individual's tax return.

(c) The 'Number of individuals subject to income matching as a tax return had been lodged' does not include those individuals whose income details have been reported, but have no obligation to lodge a tax return.

(d) The 'Number of potentially discrepant cases' are those where a potential income discrepancy has been detected.

(e) The 'Number of discrepant cases selected for actioning' is the number of cases that were actioned from the pool of potential cases for actioning, because the amount of the discrepancy was above nominated thresholds.

(f) AIIR data statistics also include statistics for Dividend & Interest Reports under Subsection 161A (1) ITAA36 and Sections 388-50 TAA53.

(g) All of the figures within the table are derived from data for discrete financial years, so that the figures for identity matching outcomes, discrepant cases actioned and liabilities raised all relate to that same year's data.

(h) Statistics provided represent a snapshot in time and some may change slightly as new data is received or further tax returns are lodged. The number of potentially discrepant cases may change as some data from third-party reporters is identified as incorrect or of poor quality and is rejected then replaced, as poor data may have been the cause of some of these potential cases.

Source: Tax Office.

3.45 Another income stream for which the Tax Office conducts data matching is foreign sourced income. Table 3.2 provides details of the Tax Office's foreign sourced income data matching for the last four operational years.

Table 3.2

Data matching using data on foreign sourced income

Data from Tax Treaty Partners	2003–04	2004–05	2005–06	2006–07
FIR data (e.g. foreign income)				
- No. of records matched	95,395	91,137	164,779	121,273
- No. of discrepant cases actioned	1,171	1,568	1,212	1,800
- Liabilities raised ($)	3,002,988	1,452,596	2,053,364	1,222,540

Notes:

(a) These statistics are based on operational years and represent work effort, as data is often received from tax treaty partners in multi-year lots that refer to various Australian operational years.

(b) These identity matches may be to a high, medium or low confidence level

(c) Statistics are for ME&I business line only

Source: Tax Office

3.46 Data from legislated sources is also available to authorised Tax Officers from all business lines, for use in various tax administration activities such as audit, data mining and analytics. Statistics and results from this use of legislated data are not consolidated or reported corporately.

Australian Transaction Reports and Analysis Centre data

3.47 Australian Transaction Reports and Analysis Centre (AUSTRAC) data represents a further source of data to augment the Tax Office's data holdings. Authorised Tax Officers have online access to the AUSTRAC database. This enables selected searching of the database to obtain information that is relevant to a case. The Tax Office (and the Privacy Commissioner) does not consider this access to be data matching, as it is not '*the computerised comparison of two or more databases or extracts*' (Privacy Commissioner *Guidelines* paragraph. 15.1), rather it is the identification of relevant information via a one-to-one comparison of records with common identifiers. Results of these searches are published in the Tax Office's Annual Report.

3.48 The memorandum of understanding between AUSTRAC and the Tax Office allows for AUSTRAC to provide the Tax Office with bulk datasets for data matching.[63] To implement Recommendation No.6 of Audit report No.47 2004–05, *Australian Taxation Office Tax File Number Integrity*, the Tax Office's Serious Non Compliance business line has initiated an 'AUSTRAC data matching project'. The purpose of this project is to determine the extent to which it is possible to reliably attach TFNs to AUSTRAC reports by matching AUSTRAC data with Tax Office client data. This project has received two bulk datasets from AUSTRAC, and is exploring options to improve identity matching. Among other options, the project plans to use data from Tax Office datasets (such as bank account records from AIIR, or relevant personal identification and financial institution details in other databases) to attempt to match a TFN to AUSTRAC records.

3.49 Tax Office data mining of AUSTRAC data occurs within the analytics area of the OCKO. Following a successful trial, an AUSTRAC data miner has been permanently seconded to the Tax Office in the analytics area. Specific data mining projects that used AUSTRAC data include the Operation Wickenby, High Risk Refunds, and Tax Havens projects. These data mining projects have also experienced difficulties in obtaining high quality TFN matches. As a result, additional techniques to improve identity matching are being investigated.

[63] Memorandum of Understanding between the Director, Australian Transaction Reports and Analysis Centre and the Commissioner of Taxation on access to and use of Australian Transaction Reports and Analysis Centre Data, 2003.

3.50 While the AUSTRAC identity matching project and the team in the OCKO have interacted, the ANAO considers that additional synergies and efficiencies may be achieved through closer integration of the two teams.

3.51 The Tax Office advised that a review of the Tax Office's use of AUSTRAC data has been established and that the purpose of the review is to identify how the Tax Office can make best use of AUSTRAC in the future environment. The Tax Office further advised that the scope of the review includes the following:

- Engage other business lines in this review by:
  - consulting with the OCKO;
  - consulting with identified stakeholders to determine their business requirements; and
  - seeking business lines resources and commitment to the review.
- Identify how the Tax Office can best optimise the use of current AUSTRAC data as well as exploring (if possible) how to use the new AUSTRAC data resulting from the AML/CTF Act reforms by:
  - gaining an understanding of what AUSTRAC data contains and will contain under the reforms;
  - determining how AUSTRAC data can add value to Tax Office intelligence;
  - discovering the different categories of products available from AUSTRAC now and into the future; and
  - confirming business line requirements for AUSTRAC data.
- Identify the required direction for optimising the use of AUSTRAC data by the Tax Office in the future, including the development of guiding principles and/or parameters for future use of AUSTRAC data and integration of the intelligence gained from the AUSTRAC data into appropriate business line activities.
- Influence the AUSTRAC review by:
  - participating in the AUSTRAC review to gain knowledge of what AUSTRAC systems can or should provide (this will include identifying any current positives and/or shortcomings);

– Based on business line requirements, advising AUSTRAC of Tax Office future needs (acknowledging that this will probably depend on the resources, time and expertise available to AUSTRAC).

3.52 The desired outcomes from the review will be:

- a strengthened relationships between AUSTRAC and the Tax Office as business intent is shared and understood;
- an optimum use of AUSTRAC data across the Tax Office;
- a more strategic and holistic use of AUSTRAC data;
- an embedded practice of using AUSTRAC data in 'business as usual' processes, including bulk data matching activities; and
- a review on existing measures and a development for new ones (if necessary).

Memoranda of understanding data

3.53 The Tax Office routinely receives datasets from Australian Securities and Investments Commission (ASIC) and State Revenue Offices under memoranda of understanding.

3.54 ASIC data provided to the Tax Office relates to licensed investment advisers and security dealers, and their representatives, including any relationships or roles they have with Australian companies. It is used to help the Tax Office review tax compliance of investment advisers and securities dealers, and their representatives, licensed by ASIC as well as identify any non compliance issues.

3.55 A memorandum of understanding provides the formal framework for exchanging confidential information and other aspects of cooperation between the Tax Office and participating State and Territory revenue offices. In 2004, the Tax Office initiated a capital gains tax project focusing on acquiring real property data to help administer Australian taxation law. More information on this project is available in Audit Report No.16 2006–07, *Administration of Capital Gains Tax Compliance in the Individuals Market Segment*.

3.56 Audit Report No.16 2006–07 noted a range of difficulties associated with the Tax Office's use of real property data provided by state government

authorities.[64] The ANAO notes the Tax Office is taking action on these matters and now receives real property data from all relevant state authorities. The Tax Office advised that all state authorities now provide property data of high quality ready for use which is updated on a regular basis. These datasets are crucial in achieving improved levels of capital gains tax compliance, as well as GST and other taxes.

Purchased data

3.57 The Tax Office's business lines may purchase datasets from commercial providers for use in audit and other activities. If the purchased data has been obtained for data matching purposes, the purchased datasets are recorded on the data matching register. If data has been purchased for other purposes, it is not subject to the privacy guidelines and is not required to be on the register.

3.58 The Tax Office purchases two main data sets, fact of death and electoral office data sets. The Tax Office uses both to update and maintain the integrity of client registers. These updates occur on a monthly basis. Table 3.3 provides the numbers of updates to the Tax Office's client registers as a result of administrative matching.

Table 3.3

Administrative matching of purchased data to update client registers

Data set	2005–06			2006–07		
	No. of records processed through identity matching engine	Records matched to a TFN	No. of updates made to client register	No. of records processed through identity matching engine	Records matched to a TFN	No. of updates made to client register
Electoral Roll	26,340,732	98.1%	Not applicable	53,060,530	98.2%	Not applicable
Fact of Death	125,211	91.5%	108,198	141,257	92.0%	108,962

Notes:

(a) The identity matches to a TFN may be achieved at a high, medium, or low confidence level.

(b) The updates made to the register from Fact of Death data occurred as a result of high quality matches only.

(c) The Tax Office investigates reasons for low and medium confidence level matches.

Source: Tax Office.

[64] See Audit Report No.16 2006–2007 *Administration of Capital Gains Tax Compliance in the Individuals Market Segment* paragraph 3.29 p. 57.

Requisitioned data

3.59 Table 3.4 shows the number of the Tax Office's ad hoc data matching projects (using requisitioned data) since establishment of the DMSC in 2003–04. Results from these data matching projects are not consolidated and reported corporately.

Table 3.4

Data matching projects using requisitioned data

Ad hoc data matching projects	2003–04	2004–05	2005–06	2006–07
New Projects Registered with the Gatekeeper	70	49	45	18
Suspended or Cancelled Projects	18	17	13	-
Completed Projects	52	20	16	7
DMSC Evaluations of Projects	20	14	7	-
Work-in-Progress Evaluations	1			
Project Closure Evaluations	*13			

Notes:

(a) The Data Matching Register was first introduced in early 2005. The 2003–04 figures represent, in some respects, historical information about projects which had already been completed. As such, information about projects which may have been commenced during 2002–03 is also included. The 2004–05 figures are somewhat more accurate but, again, some of these projects were already finalised before they were registered.

(b) Project Closure Evaluations include final evaluations provided to the Privacy Commissioner in respect of Para. 76 of the Data Matching Guidelines.

(*) 13 of the 14 evaluations required in accordance with Para. 76 of the Data Matching Guidelines have been conducted. At this stage there is no requirement to conduct evaluation for any subsequent year projects.

Source: Tax Office

3.60 The Tax Office advised that it does not prescribe benchmark measures for data matching processes. Each situation is considered based on the circumstances appropriate in that situation. These decisions are taken by the risk owner and will take into consideration a range of factors relevant to decisions to proceed, moderate or cease further work. These decisions can change as information available changes during the year, for example a new policy imperative. The ANAO raised as a matter for consideration by the Tax Office whether it may not be beneficial to provide generic guidance to staff about specific compliance improvements such as an increase in voluntary compliance in a nominated area; a specific measure of risk amelioration; a nominated proportion of additional revenue to be raised; or improved strike rates required, to be achieved from the proposed use of the requisition data to be acquired.

Analytics

3.61 In general terms, the purpose of analytics is to discover and model relationships and patterns relevant within and between variables in large volumes of data relevant to tax administration. Analytics is a broadly defined family of sophisticated methodologies which include data mining, data linking, data modelling, and expert business rules. The methodologies may use advanced mathematics and statistics in the algorithms of which they are composed. The methodologies may require considerable computer power and substantial volumes of data to generate results useful to tax administration. They may be based on a detailed understanding acquired by other means of patterns of activity in the data, including that which will be subject to the methodologies of analytics.

3.62 An analytics project may be **descriptive** in that it may describe relationships that might not otherwise have been detected. It may be **inferential** in that it generates hypotheses about relationships relevant to tax administration, including indications of new risks. It may be **predictive** in that it predicts likely tax-relevant behaviour based on characteristics in the data. It may be **decision-oriented** in that it can be used to make administrative decisions such as case selection, related to, for example, refunds, debt, lodgement or assessment.

3.63 Analytics projects may use any or all of the six categories of data shown in Figure 1.1. However, analytics projects may use requisitioned data (containing personal information about individuals) only if the protocol approved by the Privacy Commissioner permits the use of the data in this way.

3.64 The Tax Office's CKO is responsible for the administration of analytics projects. The CKO is responsible for developing a coherent vision and strategy for managing information in the Tax Office, including building organisational capability in the areas of strategic intelligence and analytics. The Corporate Intelligence and Risk Unit has been established within the CKO area to focus on enhancing the Tax Office's organisational intelligence and risk capabilities as well as developing data analytics to inform risk profiles. A key purpose of data analytics is to enable the Tax Office to provide more personalised products and services to clients through better analysis of the Tax Office data and information.

3.65 Appendix 3 summarises the Tax Office's analytics projects and shows the wide range of this activity.

Data acquisition and quality

Acquisition

3.66 An agency's organisational arrangements are necessarily a compromise between many different and sometimes competing priorities and have a range of strengths and weaknesses. ANAO considered that there are several risks associated with current arrangements that have a bearing on data acquisition and may require closer attention. One risk is that a business line focus on business line priorities could mean that whole-of-Tax-Office interests in the availability and use of data might not be optimally achieved. Another is that the better understanding of the detail of the risks that initially warranted the acquisition of the data by a business line might not be fully shared with other business lines. The ANAO considered that this risk could be managed satisfactorily if the risk owner who initiated the acquisition of the data (and therefore the data matching and analytics strategic response) maintains a strong corporate perspective throughout the project.

3.67 The risk associated with the corporate use of data acquired by business lines may require a different approach as business line initiation of the acquisition of requisitioned data may not be the optimum way for the Tax Office to acquire corporate data.

3.68 In relation to requisitioned data, a range of data management issues might not be addressed until after receipt of data. These issues include the timeliness, quality and usability of the data, its management as a whole-of-Tax Office resource, and identification of the most appropriate channel by which to receive it. In addition, the Tax Office may find there are relatively long lead times between identifying a business line compliance issue and availability of the data in a form that can be used for data matching.

3.69 In relation to legislated data, the ANAO noted that the Tax Office has arrangements to address the issues of timeliness, quality and usability of the data, and its management as a whole-of-Tax Office resource. In relation to one set of legislated data, the AIIR, the ANAO examined these issues, amongst others, in a previous report.[65] The ANAO has not examined the Tax Office's management and use of any other legislated data.

[65] Audit Report No.48 2003–04, *The Australian Taxation Office's Management and Use of Annual Investment Income Reports*.

3.70 The entities that have an obligation to supply the tax Office with legislated data range in size and technical capability from high volume, highly technical reporters such as major financial institutions to low volume, technically unsophisticated enterprises, such as small family businesses. The Tax Office advised that it is not practical or appropriate to identify one reporting channel for all reporters. Nevertheless, the Tax Office advised of a clear emphasis on encouraging reporters to move to electronic reporting. For most, this means using one or more of the reporting modules within the Tax Office's Electronic Commerce Interface (ECI) application. With regard to high volume reporters, the Tax Office has initiated a project to review the Tax Office's bulk take-on procedures with the intention of developing an automated capability for exchanging data securely, on a frequency that provides the best level of efficiency for the Tax Office and its clients.

3.71 Reporting channels for legislated data are detailed within various reporting specifications issued by the Tax Office. They currently include the following:

- paper based reporting;
- magnetic media (computer floppy disk or magnetic cartridge tape). The Tax Office withdrew the option of reporting on magnetic cartridge tape from June 2006 as a result of the industry-wide phasing out of this medium; ANAO understands that the Tax Office has extended its use on a 'case by case' basis to assist some larger reporters in their transition to other reporting channels;
- optical media (CD or DVD); and
- electronic transmission, specifically through the Tax Office's ECI application.

3.72 In 2003 the Tax Office adopted a strategic approach to the selection and use of channels for communication to and from the Tax Office's clients. The Tax Office revised the strategy in 2005. The Tax Office explained the channel strategy in this way:[66]

> The aim of the channel strategy is to create an effective method for investing in the development, enhancement, cross enterprise management, and use of customer contact channels for the delivery of Tax Office products and service, whilst minimising the burden and cost of compliance. The strategy is a means

[66] Tax Office Channel Strategy and Framework at a Glance, 23 February 2005.

> of encouraging voluntary compliance (by making compliance easier) through consideration of the way clients are treated, client preferences and inherent channel characteristics. In this way it is consistent with the Taxpayer's Charter and aligned with the principles of brand management and the compliance model.

3.73 Whilst the focus of the strategy has been on communications with the Tax Office's clients in relation to tax returns and responsibilities directly related to these, the ANAO accepted that it is an appropriate framework for guidance about the identification of the best reporting channels for legislated and requisitioned data. The reporting channels for legislated data are detailed within various reporting specifications issued by the Tax Office for legislated data. Full implementation of the channel strategy may require changes to those reporting specifications.

3.74 Once the *ECMP Change Program* is operational, there will be greater scope, and greater need, for a whole-of-Tax Office approach.[67] This is because the *ECMP Change Program* will help the Tax Office obtain a whole-of-taxpayer view of the risks associated with taxpayers whether they be individuals, companies, trusts, partnerships, superannuation entities or government agencies. Increasingly, compliance management and provision of better services requires a whole-of-Tax Office outlook.

3.75 Requisitioned data is a whole of Tax Office resource. However, corporate priorities and business line priorities could diverge not only in relation to the acquisition of a particular set of data, but also in relation to matters such as the timing, quality, priority of use of the data and the channel by which the data is provided to the Tax Office. It may be preferable for the Tax Office to adopt a more corporate and strategic approach to the acquisition of data, having regard to the data being primarily a corporate rather than business line resource.

3.76 In 2003 the Cash Economy Task Force made a number of recommendations to improve the Tax Office's business-to-consumer dealings. Aspects of Recommendation 5.10 of the Report of the Task Force are relevant to data acquisition, as the task force recommended the Tax Office:

> [I]dentify a broader range of third party information sources and negotiate with their owners to provide appropriate information on a regular and timely

[67] See *Chapter 1 Background and Context*, for a description of the *Easier, Cheaper and More Personalised (ECMP) Change Program*, and the Tax Offices introduction of a new ICP system.

basis and develop and implement more systematic and regular approaches to third party data matching.[68]

3.77 The Tax Office agreed to this recommendation. The Tax Office advised that the recommendation has been fully implemented. The ANAO noted that the initiatives of the DMSC have assisted with the implementation of this recommendation. Nevertheless, the ANAO observed that the Tax Office does not have a corporate, strategic data acquisition plan.

Quality

3.78 The quality of internal Tax Office data, the Tax Office identity matching engines, and external data all affect the effectiveness of data matching activities.

3.79 Having quality internal data is critical for the Tax Office to maximise the potential of its data matching activities. ANAO Report No.47 2004–05 found:

> [T]he Tax Office has cleansed the tax database substantially, removed extraneous records and upgraded data quality; tightened up identity requirements; brought TFN registration arrangements up to a higher standard; and is undertaking more effective data matching, especially with the use of more extensive third party datasets.[69]

3.80 The ANAO notes that significant work is underway to ensure only high integrity (i.e. fit for purpose) data is placed on the new ICP system. New arrangements are being established to deal with the quality of the Tax Office's internal data holdings, the quality of legislated data and incoming third-party datasets in the new ICP environment. The new arrangements are being implemented, but will not be fully effective until the ICP becomes operational, which is likely to be during 2008–09.

Pre-filling income tax returns

3.81 The ANAO's Performance Audit Report No.37 of 1998–99, *The Management of Tax File Numbers*, brought to the Tax Office's attention the scope for improved services the Tax Office could provide using its high quality TFN and related data. The ANAO suggested the Tax Office would be able, in principle, to provide a unique service to many taxpayers, by using the

[68] 2003, Cash Economy Task Force, *The cash economy under the New Tax System*.

[69] ANAO Report No.47 2004–05, *Australian Taxation Office Tax File Number Integrity,* p. 13.

information it had available to it for use in its data matching activities, to determine the income tax payable for certain taxpayers without the need for those taxpayers to complete and lodge an income tax form. The Tax Office examined this suggestion, concentrating in the first instance on two categories of taxpayers; those in only receipt of income from Centrelink and those in only receipt of wages or salary income i.e. apart from any income from Centrelink. Addressing the issue of work-related expenses and other deductions (claimed by over 80 per cent of employee taxpayers) was an added complication.

3.82 In 1998 the Commonwealth Government announced the Tax Office would proceed with a pilot project to pre-populate the tax returns of some people. The Government stated that the:

> Tax Office will explore the option of replacing the existing taxpayer prepared annual return with a Tax Office generated income statement. The statement would contain the income details that have been reported through the new withholding and other systems. Taxpayers could simply confirm the information by telephone to receive their refund, or add details of any other income and claims for rebates or deductions as appropriate. Conceptually a statement approach would apply to 3½ million taxpayers whose income is derived from wages and salaries, dividends and interest investments and who have the more straight-forward rebates and deductions.[70]

3.83 The Government said the Tax Office would pilot these new statements for the income year 2000–01.

3.84 The Tax Office defines pre-filling as provision of information that the Tax Office currently uses for data matching purposes, directly into an individual electronic tax preparation or record keeping tool.[71] It is intended to make the process of preparing and lodging tax returns easier and quicker for individual taxpayers. For taxpayers, it also reduces the likelihood of later Tax Office review or audit. The ANAO noted that pre-filling entails a system of data matching. The ability to pre-fill individuals' income tax return information is available to individuals through e-tax and tax agents through the Tax Agent Portal.

---

[70] A New Tax System - The Howard Government's Plan for a New Tax System, August 1998.

[71] Tax Office, *Pre-filling Intent Statement*.

3.85 The Tax Office began piloting the pre-filling concept using e-tax[72] for 2004–05. This pilot was limited to pre-filling data from Medicare Australia and Centrelink. In 2005–06 the Tax Office expanded its pre-filling pilot to include the 30 per cent childcare rebate, and bank interest and managed fund information from selected financial institutions.[73] The ANAO noted the limited range of fields and data available for pre-filling in 2005–06 restricted the relevance of the pre-filling pilot facility to a relatively small proportion of e-tax users. In 2006–07, approximately 1.9 million individuals used e-tax to lodge their return, and of these, approximately 1.1 million used pre-filling. A further 1.9 million pre-filled reports for individuals were downloaded by tax agents through the Tax Agent Portal.[74]

3.86 The Tax Office received an additional $20 million funding in the 2007–08 Federal Budget to significantly accelerate the pre-filling initiative for the 2007–08 income tax return year. The Tax Office advised that in response to the Budget announcement and funding, pre-filling will be a fully operational facility within the ICP environment in 2008–09. It will be in full deployment for around nine million taxpayers, many of whom would use the facility through their tax agents.

3.87 The Tax Office anticipates that for 2007–08, more people will use the pre-filling facility (for 2006–07 income tax year) because:

- pre-filling service will be more visible in e-tax;
- more information will be available at an earlier date making it a useful service to a wider audience; and
- the 2007–08 Budget announcement on pre-filling initiatives and media interest raised public awareness.

3.88 Pre-filling requires timely, high quality data from employers and financial institutions, preferably online.

---

[72] The Tax Office has been developing e-tax since 1999 with the aim of making it easier and simpler for taxpayers with less complex tax affairs to complete their tax return. E-tax is a free service the Tax Office provides that allows taxpayers to lodge their individual tax returns online.

[73] When e-tax 2005–06 was decommissioned, the proportion of the 1.6 million e-tax users who used pre-filling information is as follow:
- 84,257 (5%) downloaded bank interest details;
- 65,909 (4%) downloaded information to claim the 30% Child Care Tax Rebate (CCTR);
- 104.009 (6%) downloaded out-of-pocket medical expenses from Medicare Australia; and
- 7,954 (<1%) downloaded managed fund information.

[74] These figures are as at February 2008.

3.89 The ANAO noted the pre-filling initiative requires the Tax Office to have relevant third party datasets available as soon as possible at the end of the financial year. In the Nordic countries, where the practice of pre-filling is well established and reportedly operating with a high degree of success, third parties face strict and timely end-of-year reporting obligations under the law with final information reports typically required within 3–4 weeks after the end of the relevant fiscal year.

3.90 The legislation that regulates provision of third party data necessary for the pre-filling initiative (e.g. the employment and financial data) requires the data be provided to the Tax Office no later than 14 August and 31 October respectively.[75] However, a number of data providers have voluntarily provided the data required by the Tax Office earlier than the legislated requirements, in order to enable pre-filling activities to occur. This is beneficial to a significant number of e-tax users who lodge their tax returns early in order to obtain a tax refund. Because of this, the longer term success of pre-filling under current arrangements will depend on the continued voluntary efforts of the suppliers of legislated data.

3.91 However, given the Government's intention to simplify income tax returns by providing pre-filling services for around 9 million individual taxpayers with effect from 2007–08 income year, the ANAO considered the Tax Office's reliance on the goodwill of data providers to provide data earlier might not be sufficient to optimise the implementation and efficiency of pre-filling initiative.

3.92 The Tax Office advised during the year 2008, it will encourage reporters to submit their reports earlier. For the longer term, the ANAO suggested the Tax Office consider the need to recommend legislative change to bring these dates forward and mandate electronic transmission in a range of areas.

3.93 The inclusion of the TFN as a key identifier in legislated data has been a factor in the higher integrity of that data and the improved administration that comes from its use.[76] An issue for consideration is the attachment of the TFN to some third party data in a similar manner. The ANAO noted that any options for legislative amendment would require discussion with the Treasury in the first instance.

---

[75] This is because historically, this data was supplied to the Tax Office for post-issue data matching against income details provided by taxpayers.

[76] Audit Report No.47 2004–05, *Australian Taxation Office Tax File Number Integrity* (see also Audit Report No.37 1998–99, *Management of Tax File Numbers*).

Recommendation No.3

3.94 The ANAO recommends that to simplify income tax returns by providing pre-filling for increased numbers of taxpayers the Tax Office discuss with the Treasury options in relation to the pre-filling initiative, namely:

- bringing the dates forward for the provision of requisite third party data;
- mandating electronic transmission in a range of areas; and
- including the TFN on some additional data sets, having regard to the need to balance privacy concerns and improving the efficiency and effectiveness of public administration.

3.95 **ATO Response**: *Agreed.* The Tax Office is committed to improve services to taxpayers through the pre-filling initiative, and recognises the need to bring forward the availability of data to support this. For 2007/2008 the focus has been on encouraging large providers to voluntarily lodge information early, and wherever possible, electronically. Early indications are that these providers recognise the benefits for their customers and employees and will co-operate. The Tax Office will discuss changes to enhance the pre-filling process with Treasury. Where the inclusion of a tax file number on additional data sets is considered warranted to support the treatment of compliance risks, the matter will be discussed with Treasury.

Optimising the potential of data matching and analytics

Building on existing strengths

3.96 The ANAO found that the Tax Office makes good use of the substantial data received from third parties as a legislative requirement.[77] Each set of data is used for data matching. Nevertheless, the ANAO considers that the Tax Office could more fully use this data; there are some subsets of it that are not used. The Tax Office's data matching activity is largely limited to individual taxpayers; there is scope for using legislated data in data matching and analytics projects to better manage compliance risks about non-individual taxpayers (companies, trusts, partnerships, superannuation funds and government agencies). The Tax Office could also make better use of the data third parties provide under legislative requirements in respect of certain classes of investment income received by individual taxpayers, such as

[77] Appendix 1 contains a comprehensive list of the data available to the Tax Office.

dividends and distributions from partnerships, trusts (including unit trusts) and private companies. There may be an implied obligation on the Tax Office to use completely all data that is provided to it by other parties in fulfilment of the legislative obligation on them. The ANAO understood that the Tax Office would address any implied obligation having regard to a range of factors that require the prioritisation of activity, including an assessment of risks relevant to the data and the allocation of resources required to address the risk. The ANAO notes that, in response to Recommendation No.7 of Audit Report No.48 2003–04, *The Australian Taxation Office's Administration and Use of Annual Investment Income Reports*, the Tax Office intends to include the capacity to do this as part of Release 3 of the *ECMP Change Program* scheduled for deployment in 2008. There may also be scope to make greater use of PAYG Payment Summaries.

3.97 The ANAO considered there may also be scope for additional data matching and analytics projects in several other areas, such as assessable government payments[78], debt, fringe benefits tax, GST and the cash economy. Most of these areas would benefit from better use of existing data, not necessarily acquisition of more data. Each of these areas has been examined in previous ANAO audits, which also canvassed the opportunities for improved tax administration that could arise from the better use of existing data.

3.98 The Tax Office has historically used data matching and analytics for compliance activities. It aims to identify non-compliance with respect to:

- registration – by identifying people and entities outside the tax system;
- lodgement – by detecting late lodgement or failure to lodge; and
- reporting – by detecting undisclosed or under-reported income.

[78] Assessable government payments include Australian and state government payments in relation to bounties; diesel fuel rebates; fuel tax credits; fuel grants; fuel sales grants; cleaner fuel grants; product stewardship (oil) benefit; Department of Agriculture, Fisheries and Forestry grants; drought relief; purchase of water tanks; school bus subsidies; heritage and conservation (including water conservation) incentives; employee subsidies; export incentive grants; industry restructuring and adjustment payments; industry research and development subsidies and grants; and Medicare payments for medical practitioners. The Tax Office's administration of assessable government payments has been the subject of previous audits and remains an area of compliance risk, as reported in previous audits (Audit Report No 3 of 1998–99, *Assessable Government Industry Assistance; Follow-up Audit*, tabled on 12 August 1998) most recently Audit Report No. 47, 2004–05, *Australian Taxation Office Tax File Number Integrity*, especially Chapter 6 and recommendation No. 7). The Tax Office recently estimated that aggregate annual government outlays on assessable government payments would be around $10 billion. There are continuing problems, ranging from incomplete reporting requirements on tax returns, inadequate or incorrect advice being given by federal and state authorities; poor controls over data quality by the Tax Office and other relevant Commonwealth authorities, lack of compliance with GST reporting and payment requirements; to inadequate data matching by the Tax Office. In the absence of any recent detailed risk assessment, ANAO considered that assessable government payments are potentially an area of serious compliance risk.

3.99 It also aims to detect instances of identity fraud, identity theft and the misuse of identity information.

3.100 There may be scope to add 'meeting established obligations' to this compliance agenda, specifically relating to tax debt. According to the Tax Office's Compliance Program, making correct and timely payment is part of the taxpayers' obligation. The ANAO noted that the Tax Office has several analytics projects that build models for debt case selection and prioritisation of debt collection cases. In addition, the ANAO noted that data matching projects within compliance business lines have an opportunity to provide some inputs for the Debt business line. The ANAO understood that during 2007 the Debt business line received additional funding for a range of data matching and analytics projects.

3.101 The ANAO also considered there may be scope for acquiring a number of additional external datasets (such as elements of the SENSIS database for serious non compliance projects).[79]

3.102 Most of the 'situational' or 'project based' data matching projects are quite unlike the large-scale post-issue automated systems data matching conducted through the ATOms. This is because the projects involve a much larger element of risk exploration, analysis and identification. This leads to a decision on compliance activities after the data has been acquired and analysed. Even though these projects contain a significant research and development component, they invariably result in Tax Office administrative action in relation to tax payers.

3.103 The ANAO considered there may be scope for increased exploratory use of data matching and analytics to identify and improve the Tax Office's understanding of a range of compliance risks. Such risks include those to the integrity of the tax system, the Tax Office's reputation and to the revenue base.

3.104 The ANAO noted that there are synergies amongst that variety of data matching and analytics methodologies that can substantially improve the value of data to the Tax Office. Specific data matching methodologies can improve the quality of various analytics methodologies, for example by improving the quality and value of data inputs. Conversely, analytics methodologies can also provide technology to assist data matching activities, by improving, for example, the efficiency and effectiveness of the selection of

[79] By 'SENSIS' the ANAO means the most up-to-date electronic edition of the white and yellow paged phone books.

'at risk' cases as a result of degrees of mismatch. In addition, the value of data to the Tax Office can be improved by matching and analysing it in relation to a variety of methodologies and additional data bases and using it in a multiplicity of ways.

3.105 The Tax Office obtains greatest benefit from third party data when the data matching and analytics capability can be applied to it fully, and the data set can be used with a range of other datasets to improve compliance and/or provide better services.

3.106 The ANAO considered that there would be merit in the analytics group consulting more closely with the DMSC when it next undertakes strategic planning, outreach and prioritisation of projects. This would provide an opportunity for the closer integration of the data matching and analytics capability with whole-of-Tax Office and whole-of-client operations.

Improving taxpayer experiences

Making compliance easier

3.107 The ANAO considered that there is scope for the Tax Office to better manage corporately and strategically the data matching and analytics capability. Amongst other things, this could enable a closer alignment of compliance responses to the compliance model. As the compliance model shows, the appropriate Tax Office response for the majority of tax payer compliance issues is an educative, enabling response, not a punitive one. However, for it to be fully effective, this response has to directly address the problem that gave rise to the compliance concern. In addition, the response would be more effective if it made compliance easier and cheaper for the tax payer, as the *ECMP Change Program* acknowledges.

3.108 These two factors (addressing the specific tax payer compliance matter and making compliance easier and cheaper for the tax payer) ultimately require a response tailored to the situation of the specific tax payer. The provision of web-based electronic services and tax payer specific communications that the emergent data matching and analytics capability may enable the Tax Office to function in this way.

3.109 The same capability can be deployed to improve the efficiency and effectiveness with which the Tax Office manages compliance risks at the apex of the compliance pyramid. In this case the emergent data matching and analytics capability could enable the Tax Office to detect serious compliance

risks at the earliest possible stage and institute effective mitigation strategies tailored to specific behaviour.

3.110 An example of how the Tax Office has used the data matching and analytics capability to make compliance easier can be seen in the CGT Compliances Challenges project. Here, the Tax Office uses state revenue office property sales data to identify sales of properties for which rental income has previously been declared; the Tax Office alerts property sellers that have not yet lodged an income tax return, that they may need to consider the capital gains tax implications of the sale.

3.111 Pre-filling has the potential to improve taxpayer experiences by making compliance easier and cheaper, as relevant fields of the tax return would be already filled in.

Provision of services in relation to economic transfers

3.112 Given the Tax Office's increasing role in delivering social policy through administering economic transfers, such as grants, benefits, tax offsets and redistribution programs, there may be significant scope for the Tax Office to provide services to its clients.[80] The provision of improved services in relation to economic services is but one feature of improved tax administration that could be enabled by the emergent data matching and analytics capability.

3.113 In addition, given that the provision of a service can be the most appropriate Tax Office response to typical compliance risks, the data matching and analytics capability can enable the provision of this service. This may be a more efficient and effective use of the capability than using it merely to address incomplete or inaccurate reporting by tax payers of their tax liabilities.

3.114 One example of the Tax Office providing a service through the data matching and analytics capability is the management of superannuation co-contributions. In this instance, the Tax Office uses income tax return data, and superannuation contribution information provided by superannuation funds to calculate eligibility for the co-contribution payment and automatically make the payment.

[80] The most significant economic transfers are the various fuel grant schemes that provide assistance to users of various types of fuels, mainly for business purposes. They include the energy grants credits scheme, the fuel sales grants scheme, the product stewardship for oil program and the cleaner fuels grants scheme. Other transfers that provide assistance for families and individual taxpayers include the family tax benefit, the baby bonus and the private health insurance benefit. Superannuation co-contribution payments and superannuation guarantee payments provide the mechanism through which certain superannuation transfers are effected, while the research and development tax offset and the large scale film production expense tax offset provide selected assistance to industry.

3.115 The ANAO considered that there is scope for the Tax Office to identify proactively and manage strategically, opportunities for making compliance easier and improving service provision through the better management of the data matching and analytics capability. Although the prime responsibility for initiating work about the provision of new services in this way rests with risk owners, there is scope for the DMSC to contribute via its facilitation and coordination role.

A paradigm shift

3.116 Introduction of pre-filling represents a paradigm shift for the Tax Office's use of data matching and analytics in tax administration, as it involves a shift from bulk post-assessment income data matching to pre-filling of electronic returns.[81] In the longer term, as the pre-filling initiative matures and more taxpayers use the service, there may be a reduced need for the Tax Office's post-assessment automated data matching program.

3.117 Provision of services (enabled by data matching and analytics) is another feature of this shift in tax administration. While there is a challenging range of new opportunities for the Tax Office to address as the understanding of this service paradigm improves, the legislation under which relevant data is acquired is firmly grounded in the compliance-centred paradigm of tax administration.

3.118 The ANAO considered that the DMSC has successfully facilitated the more corporate and strategic use of data matching as an integral component of tax administration. This has resulted in improved compliance, improved services and the more efficient and effective use of resources. It has also enabled the Tax Office to better understand risks. The procedural arrangements established by the DMSC have enabled the Tax Office to improve the use of data matching and compliance with Privacy Commissioner Guidelines. Whilst the ANAO has not examined data matching in other national revenue authorities, the ANAO acknowledges that the Tax Office is highly regarded by the leading national revenue authorities of the OECD as being a leader in this field.

3.119 The ANAO considered that the Analytics team in the OCKO has established its value to tax administration during the three years since its establishment. It has pioneered many new sophisticated mathematical and

[81] Commissioner's online update, May 2005.

statistical methodologies resulting in improved compliance, improved services and the more efficient and effective use of resources.

3.120 The ANAO suggested that it may now be opportune to build on these achievements given the implementation of the *ECMP Change Program* during the next few years. The ANAO considers that the recognition by the Tax Office of the emergent data matching and analytics capability and the improved management of it, both corporately and strategically, would be a useful step to take in this direction. The ANAO acknowledged that necessarily this capability will be distributed throughout the Tax Office as it is deployed on a daily basis to support the primary tasks of improved compliance, provision for services and the collection of revenue. The deployment of the capability in this way may heighten the need for stronger corporate and strategic guidance. In addition, the ANAO considered that the Tax Office needs to develop a strategic plan for acquiring and using external data over a three to five year period, and that this plan have regard to:

- compliance risks;
- a whole-of-Tax Office perspective;
- including as goals for the use of external data compliance verification, improving service provision and making compliance easier and cheaper;
- forming part of a rolling data acquisition and use program;
- identifying and prioritising additional external datasets for acquisition (e.g. SENSIS data) as well as additional data matching and analytics projects across the whole Tax Office (e.g. trusts and partnerships; superannuation; assessable government payments; debt; fringe benefits tax; GST and the cash economy);
- including preferences for channels and methods of data acquisition; and
- guiding future decision making with respect to:
  - investment in enhancing the quality of external data sources; and
  - payment for external data.

Recommendation No.4

3.121 The ANAO recommends that the Tax Office develop a three to five year corporate and strategic plan for the acquisition and use of external data.

3.122 ATO Response: *Agreed*

4. Governance and Coordination Arrangements

This chapter reviews the Tax Office's data matching governance arrangements, including coordination of activities, performance monitoring and reporting, and guidance for staff.

Introduction

4.1 The ANAO's examination of the Tax Office's governance and coordination focused on:

- management of the data matching and analytics capability;
- performance monitoring and reporting; and
- guidance for Tax Office staff.

Management of the data matching and analytics capability

4.2 The administration of the Tax Office's data matching and analytics capability is distributed amongst the business lines. Identity matching and ATOms are administered by the ME&I business line; the data matching projects coordinated and facilitated by the DMSC are the responsibility of the business line that initiated them; and analytics is the responsibility of the OCKO. The Business Solutions business line is responsible for the pre-filling project, and has also undertaken an evaluation of the identity matching facility, with a view to designing a new one that will initially be used by the Superannuation business line, and may have wider application throughout the Tax Office.

4.3 The administration of the data matching and analytics capability is necessarily distributed as the capability is integral to tax administration. The Tax Office has established the DMSC to coordinate and facilitate the best corporate results from the capability, and other coordination and consultative arrangements, to encourage best practice and intra-business line consultation.[82]

[82] Examples of the coordination arrangements initiated by the DMSC include the provision of guidance for staff, in the form of a data matching policy *External Data Matching Practice Statement CM 2004/17*, templates to use when preparing a *Data Matching Concept Brief* and a *Data Matching Program Protocol,* and guidelines and FAQ's on the Tax Office intranet. The DMSC has also established a Data Matching Gatekeeper, a data matching register, and an annual certificate of assurance process that aims to ensure compliance with the Privacy Commissioner's *Data Matching Guidelines.*

The ANAO considered these arrangements have generally worked well and the DMSC has been pivotal in the good results the Tax Office has achieved from the deployment of the capability.

4.4 The ANAO considered that the CKO's analytics group has established its value to tax administration but could proceed more strategically in the future. For example, it would be appropriate for the CKO to take the initiative with business lines to promote further use of data matching and analytics capability in tax administration. To do this, it would be necessary for the group to develop a strategic plan and adopt a greater outreach approach about deploying its capability. The ANAO considered this approach requires increased coordination with the DMSC to ensure the two areas proceed jointly and efficiently with respect to strategic planning, outreach and prioritisation of projects.

4.5 The ANAO observed that the DMSC and the CKO's committees (e.g. the Information Management Advisory Committee, the Intelligence and Risk Leadership Group and the Data Stewards Reference Group) charters were updated in March 2007 to align with the Tax Office's Committee Management practice statement. However despite the close interrelationships of their goals and activities,[83] they were developed in isolation from one another, and do not reference each other in their 'governance and relationships' sections.

4.6 The ANAO noted that the Tax Office recognised the importance of applying Tax Office-wide agreed terms and definitions and is currently developing a common business language repository of key business concepts. During the audit the Tax Office took steps to ensure that there would be a consistent understanding and interpretation of key concepts such as 'data matching', 'identity matching', 'analytics', 'data mining', 'data modelling', and 'data mapping' across the Tax Office. The ANAO noted that while the abovementioned terms have been defined, the process of gaining consistent usage across the organisation may need further development.

[83] For example, the DMSC states that its role is to 'promote a corporate approach to identifying data that meets business needs, acquisition of that data … and the use of that data within the organisation' while the Data Stewards Reference Group states that its propose is 'improving and maintaining the quality of data', and the Information Management Advisory Committee's scope includes all relevant principles, policy, strategy, assurance arrangements and capabilities for the full range of activities and programs designed to improve the management of information as a core Tax Office resource.

Recommendation No.5

4.7 To minimise duplication and close gaps in responsibilities and streamline co-ordination arrangements, the ANAO recommends the Tax Office clarify the roles of the various committees relevant to the management of the data matching and analytics capability, so their individual responsibilities are integrated more closely.

4.8 **ATO Response:** *Agreed*

Performance monitoring and reporting

4.9 Performance monitoring and reporting are important aspects of governance. Timely collection and analysis of performance information about the value that data matching and analytics can add to tax administration is important. It can assist the Tax Office make informed decisions about investing in the development of the capability and the extent to which it is used to achieve Tax Office outcomes. In some instances the role of data matching and analytics may be so intermingled with other actions, that identifying the valued added attributable to it may prove difficult. In other cases, several years may elapse before the intended changes are achieved. In spite of the difficulties, managers have available a range of evaluation methodologies that have been designed to address these and other challenges of measurement and attribution. The methodologies cover the full range of evaluation from qualitative to quantitative.

4.10 Major Tax Office initiatives in which data matching and analytics has been instrumental in achieving improved compliance jointly with other actions may take place over several years before significant improvements can be noted. Identification of the precise quantum contribution made by data matching and analytics may not be possible, but usually a qualitative evaluation will be achievable.

Case study: data matching and improved compliance in the legal profession

4.11 Data matching and analytics are applied jointly with other components of tax administration. These include litigation, communication, marketing, education and outreach campaigns, and the provision of improved, and perhaps more targeted, services. Furthermore, more effective Tax Office action may be achieved sometimes by using data matching and analytics in

conjunction with related matters such as debt identification and collection, lodgement enforcement and/or rulings and legislative change.

4.12 The ANAO considered the Tax Office's actions since 2000 in relation to the legal profession are a good illustration of a successful Tax Office campaign which aimed to rectify serious problems of non-compliance and achieve improved industry wide compliance. Data matching was integral to the results the Tax Office has achieved, but the overall success depended upon a wide range of Tax Office actions operating synergistically.

4.13 As noted in previous Auditor-General and other reports, data matching can assist the Tax Office achieve significant improvements in the levels of compliance amongst taxpayers. This is revealed in the improved levels of compliance now occurring in the legal profession as a result of the legal profession project established by the Tax Office in response to the findings and recommendations of Audit Report No. 23 1999–00, *Management of Tax Debt Collection.* Since the establishment of that project, there has been a significant improvement in the levels of compliance throughout the legal profession. Data matching has been integral to this result.

4.14 During 1998–99 at the time of the tax debt collection audit, over 20 per cent of taxpayers in the legal profession were tax debtors when the average level of debt per occupation was only 2.6 per cent. There were also serious problems of non-lodgement, late lodgement and persistent tax debtors.

4.15 The Tax Office advised that there are no new instances of debtors being bankrupted repeatedly in the legal profession like those described in Audit Report No.23 1999–00. Lodgement and debtor behaviour has also significantly improved as shown in the summary in Table 4.1 below.

Table 4.1

Results of the legal profession project

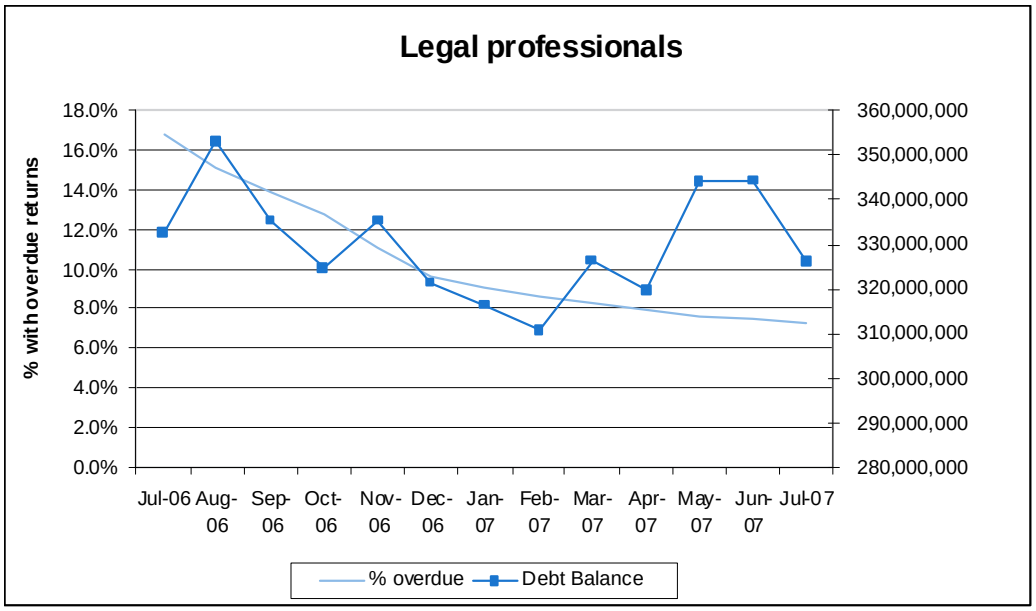
Data Matching and Tax Compliance Improvement
Data Matching has proven useful to the Australian Taxation Office to assess the tax compliance of members of the Legal Profession and to help the profession improve its performance in meeting its tax obligations.
The Tax Office observed significant numbers of legal professionals failing to lodge income tax returns and not paying tax debts. It was felt that a systemic problem may exist within the profession with respect to taxation obligations.
The Tax Office first obtained details of all Barristers and Solicitors with practising certificates and all active members of the Judiciary. This information was used to ensure that all lawyers and members of the Judiciary were registered with the tax system by having Tax File Numbers, and then to determine how well the profession as a whole was performing with respect to lodgement of documents and payment of tax.
The acquisition of the data was valuable in gaining a more complete picture of the profession as it provided an up to date snapshot, unhampered by information gaps in the information available in the tax system which occur when taxpayers fail to meet their obligations.
The Tax Office identified a small number of lawyers who were not properly registered and substantial numbers who had failed to meet other tax obligations, in particular lodgement of income tax returns.
In addition to pursuing lodgement for returns outstanding, the Tax Office was able to use the total industry picture to demonstrate to the regulators of the profession that problems existed and through a series of meetings strategies were developed to address the situation co-operatively.
By taking this holistic approach the Tax Office and the profession have been able to encourage non-compliant legal professionals to get their tax affairs in order. The Tax Office has collaborated on educational programs that fit into the existing professional programs to build strong understanding of responsibilities for the future. The Tax Office has provided information that can be used in articles in professional and association magazines. Where necessary the Tax Office has pursued cases individually. The approach has seen the Tax Office developing global responses to the issues identified.
The Tax Office has also looked at the links between legal firms, self-managed super funds and service trusts and the individuals connected to them to determine the compliance of entities associated with the profession.
The results of these activities have been a decline in the number of overdue income tax returns for legal professionals and a reduction in the rate of growth of the debt for the profession as a whole.

Source: Tax Office

4.16 The following chart shows the improvement during the 2006–07 year in the lodgement of lawyers' income tax returns in respect to the 1998–2005 years of income.

Figure 4.1

Lodgement of lawyers income tax returns

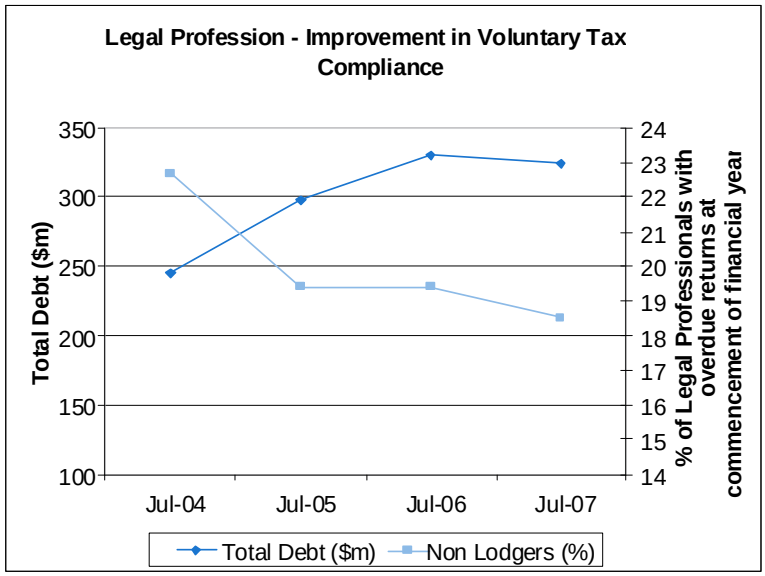


Source: Tax Office

4.17 The chart shown in Figure 4.2 compares the movement over time of voluntary compliance in both debt and lodgement of income tax returns. The percentage of overdue returns reflects the position at the outset of each new financial year showing a trend decrease in legal professionals who are expected to lodge and do not lodge on time. This does not reflect the additional direct impact of compliance activities to pursue lodgement, which substantially reduces the number of overdue returns over the course of each year.

Figure 4.2

Legal Profession voluntary compliance in debt and lodgement of income tax returns.



Source: Tax Office

Performance monitoring

Health of the System Assessments

4.18 The Tax Office conducts an annual Health of the System Assessments (HOTSA) process, to assess the 'health' of the revenue system with respect to the entire organisation, and its components such as particular revenue types (e.g. GST, income tax), market segments (e.g. individuals, large businesses) or risk areas (e.g. aggressive tax planning).

4.19 The DMSC has been coordinating and facilitating data matching projects using requisitioned data since 2003–04. In 2005 the DMSC initiated and facilitated an informal cross business line data matching risk workshop, the output of which was a draft *Data Matching HOTSA*.[84] The scope of the *Data Matching HOTSA* also included data acquisition and identity matching. It did not include analytics as that function was being established at the time. The Tax Office advised the DMSC considered that analytics was outside the scope of the informal HOTSA. The ANAO considers the DMSC's development of the *Data Matching HOTSA* was a sound initiative, as the process by which it was

[84] This HOTSA remains a draft as it was not corporately initiated or endorsed.

developed focused attention on data acquisition, identity matching and data matching within the Tax Office, and strengthened cross business line understanding and information sharing. It facilitated business lines addressing many of the key concerns raised by the process and helped the Tax Office make better and more corporate and strategic use of data matching. The resultant document contained a number of draft recommendations identifying opportunities for improvement. The *Data Matching HOTSA* was not repeated in 2006, however the Tax Office advised that in July 2007 it conducted a second cross business line data matching risk workshop.

4.20 The ANAO found the Tax Office's data matching practices are well established. The ANAO considered that, given the increased incidence of data matching and analytics work over the last five years, the planned introduction of the *ECMP Change Program*, and the demonstrable synergies between data matching and analytics methodologies, there is merit in the Tax Office consolidating corporate understanding of data matching and analytics, the achievements made and the lessons learnt from data matching and analytics. This could be done by conducting a high level stocktake of data matching and analytics activities relevant to improving the corporate and strategic management of the capability. ANAO noted that the informal *Data Matching HOTSA* also considered that a stocktake of the Tax Office's data matching capabilities was warranted.

Recommendation No.6

4.21 To improve the efficiency and effectiveness of the corporate and strategic management of the data matching and analytics capability and inform the further development of the capability, the ANAO recommends that the Tax Office complete a high level 'stock take' of achievements and lessons learned from its deployment.

4.22 ATO Response: *Agreed*.

Compliance Effectiveness Measures *Project*

4.23 The ANAO noted that the Tax Office has initiated a *Compliance Effectiveness Measures* project. The aim of this project is to develop a corporate framework that would enable the Tax Office to measure the change in taxpayers' compliance resulting from its activities. While this project does not specifically address the effectiveness of the Tax Office's management of the data matching and analytics capability, the ANAO considered it may be

advantageous to adapt the forthcoming framework to the measurement of the efficiency and effectiveness of the management of the capability.

Evaluation

4.24 As part of the DMSC's continuous improvement approach to its duties, a DMSC subcommittee conducts work-in-progress and project closure evaluations of the data matching projects that use requisitioned data. These reviews have proved beneficial to the Tax Office.

4.25 The Tax Office does not task the DMSC to evaluate the Tax Offices' legislated data matching activities. The Tax Office explained that this role was one of the normal evaluative tasks of business lines.

Cost effectiveness

4.26 The *Data Matching Concept Brief* and *Data Matching Program Protocol* prepared by the projects that match requisitioned data include an estimate of the costs and benefits (qualitative and quantitative) for each project. On conclusion of a project, a *Data Matching Protocol Evaluation Report* is also prepared for the Privacy Commissioner, which includes a compilation of the costs and benefits, and an attempt to assess the overall effectiveness of the project in relation to original objectives.

4.27 The ANAO conducted an analysis of a sample of concept briefs, program protocols and evaluation reports, and found:

- in some cases, the cost and benefit estimates identified in a project's *Data Matching Program Protocol* followed a generic pro-forma, and did not include project-specific estimates;
- projects used differing approaches and methods when estimating a project's costs and benefits in the *Data Matching Concept Briefs*;
- as data acquisition and use comprise one element of a project, the specific costs and benefits relating to these activities were not always separately identified and estimated in the *Data Matching Concept Briefs* and *Data Matching Program Protocols*;
- projects used differing approaches and methods when reporting the project's costs and benefits in the *Data Matching Protocol Evaluation Reports*; and
- estimates and evaluations were undertaken on a project-by-project basis, and were not integrated to give a whole-of-Tax Office picture of

the costs and benefits associated with data acquisition and use in, say, given timeframes.

4.28 The ANAO saw merit in placing the Tax Office's cost benefit analysis on a more disciplined footing and extending it to take a whole-of-Tax Office perspective. There may be scope for the Tax Office to apply some of the rigour of cost benefit analysis or cost effectiveness analysis to the periodic evaluation of data matching and analytics projects.[85] Introducing a more rigorous approach to the cost benefit/cost effectiveness evaluation of the Tax Office's management of the data matching and analytics capability could improve the strategic value of this capability to Commonwealth administration.

4.29 As previously noted, addressing the measurement and attribution difficulties in this area may present challenges to the adoption of standard financial, economic and/or social science evaluation methodologies to unique aspects of tax administration. In addition, there are significant differences amongst the wide range of data matching and analytics methodologies. At this operational level, there may not be scope to apply a standard evaluation methodology. However, there would be scope to apply a more rigorous, disciplined approach, perhaps using a range of data matching and analytics methodology evaluation techniques. There will, therefore, necessarily be differences in the evaluation techniques relevant to each.

4.30 It may seem that the many differences mean that a more disciplined approach to the evaluation of data matching and analytics methodologies could not be achieved. The ANAO considered this unlikely and suggested that the likely advantages to tax administration from the adoption of a more disciplined and rigorous evaluative framework would outweigh any disadvantages.

4.31 If evaluation is focussed more strategically on the relationship between the objectives to which the use of the data matching and analytics capability has been applied rather than micro aspects of methodologies or techniques, information could be generated that would be useful to the assessment of value for money achieved by the data matching and analytics capability. Among other things, there is scope for the Tax Office to have improved procedures for measuring the efficiency and effectiveness of the role of data matching and analytics capability in tax administration. The use of the

[85] See the Department of Finance and Administration's Financial Management Reference Series, *No.5 Introduction to Cost–Benefit Analysis and Alternative Evaluation Methodologies* (2006) and *No.6 Handbook of Cost–Benefit Analysis* (2006).

improved evaluation could produce better high level information useful for resource management in relation to, for example, the allocation of resources to data matching and analytics rather than other strategies to achieve the outcomes of tax administration. This could prove useful to risk owners, relevant sub-plan managers and senior business line staff who have the responsibility to apply resources most cost-effectively to achieve Tax Office outcomes.

4.32 The ANAO notes the draft *Data Matching HOTSA* also considered there was a need to understand the cost of data matching from a whole-of-Tax Office perspective. The draft *HOTSA* recommended that, '*The [Tax Office] also needs to better understand the investment in data matching activities across the [Tax Office]. A pilot model needs to be trialled to cost activities (p.14)*'.

4.33 A further consideration is the need for the Tax Office to estimate the benefits, if any, (such as reduced costs or other savings) that taxpayers may achieve as a result of initiatives enabled by data matching and analytics to make compliance easier and cheaper for them. The need to do this may become more compelling if the Tax Office wishes to evaluate the extent to which the provision of new and better services, some enabled by data matching and analytics, make compliance with tax obligations easier, cheaper and more personalised, as is intended by the *ECMP Change Program*.

Reporting

4.34 The DMSC receives an annual performance report from the Data Matching Gatekeeper, who reports on the progress of the data matching projects using requisitioned data listed in the Compliance Plan and various ad hoc reports from the Identity Matching team and business lines on an as-needs basis. The DMSC is phasing out the receipt of progress reports about implementation of the recommendations from Audit Report No.48 of 2003–04 about managing AIIR data, and does not receive reports about data matching using other legislated data sets.

4.35 Prior to March 2007, the DMSC was the peak body in the area of data acquisition, identity matching and data matching. In March 2007, the Tax Office Executive determined that the DMSC report to the Compliance Executive Meeting; however, future reporting requirements were not determined.

4.36 The ANAO considered that routine high level trend analysis reports may help the relevant sub-plan managers, business line risk owners and senior

business line management and the Tax Office Executive monitor the success of data matching and analytics activities and identify areas for improvement. The ANAO suggested that there would be merit if these reports covered appropriate components of the data matching and analytics capability, including identity matching, data matching (using legislated, memorandum of understanding and requisitioned data) and analytics. This could be achieved if, for example, business lines, the DMSC and the CKO all used common evaluation methodologies, at least at the strategic end of high level trend analysis, and coordinated their reporting arrangements so that the Tax Office Executive, relevant sub-plan managers, business line risk owners and senior business line management received a more integrated and comprehensive report.

Guidance to Tax Office staff

4.37 The DMSC provided data matching guidance to Tax Office staff through publication of a data matching practice statement, publishing data matching information on the Tax Office's intranet and in internal newsletters, and creating a data matching gatekeeper position to advise staff, where required. This guidance has enabled the Tax Office to achieve compliance with the *Data Matching Guidelines*. The DMSC provided data matching project leaders with templates to use in preparing *Data Matching Concept Briefs* and *Data Matching Program Protocols*. The Tax Office introduced tighter controls on the development and consideration of data matching opportunities, and brought data matching strategies to business line managers' attention, facilitating an improved whole-of-Tax Office and whole-of-taxpayer approach.

4.38 Nevertheless, a range of issues emerged in recent years for which staff need further guidance. The DMSC and the Gatekeeper provide advice to staff on an ad hoc basis about key data acquisition matters, including quality, format, and timeliness. However, it may be more efficient for the Tax Office to issue generic guidance about these matters. It may be appropriate, therefore, for the Tax Office to issue data acquisition guidance to staff. DMSC, in consultation with the OCKO and business lines, could prepare this guidance.

4.39 Tax agents have a critical role in the improvement of compliance amongst taxpayers.[86] Given the reliance on tax agents by the community, the ANAO noted that data matching in relation to taxpayers has a dual role in

[86] See Audit Report No.30 2006–07 *The Australian Taxation Office's Management of its Relationship with Tax Practitioners: Follow-up Audit.*

relation to tax agents. The dual role arises because compliance issues, especially those of a more serous kind, arising from the activities of taxpayers, may indicate compliance concerns in relation to tax agents. Furthermore, information about compliance shortcomings in the returns of taxpayers can be used by the Tax Office to help tax agents manage their businesses more efficiently and effectively. Because of this dual role, the ANAO considered the Tax Office should use any data matching activity that focuses on taxpayers to produce information relevant to addressing these two concerns about tax agents. This should help the Tax Office better identify tax agent specific compliance risks, especially those of the more serious kind. It will also help the Tax Office assist tax agents to better understand compliance problems of taxpayers and some tax agents specifically, especially those who lodged the returns of the taxpayers that data matching has revealed as presenting compliance risks.

4.40 The ANAO noted the benefits of the templates issued by the DMSC and used to prepare *Data Matching Concept Brief* and *Data Matching Program Protocol*. However, the ANAO considered that there was scope to improve the *Data Matching Concept Brief* by adding to the templates so as to:

- require project managers to take a whole-of-Tax Office (and/or whole-of-client) perspective;
- include a rigorous process for obtaining consultation comments from all business lines and relevant data stewards against a range of criteria;
- integrate data matching and analytics; and
- outline how the project might be used to provide better, including more personalised, services and make tax compliance easier (both being goals of the *ECMP Change Program* initiative).

Emerging issues

Online access to legislated and other external data

4.41 The Tax Office currently receives data, including legislated data, via a multiplicity of channels. This can be a source of administrative inefficiency. The ANAO considers there is a need for the Tax Office to streamline data acquisition, thus reducing to the minimum the number of channels through which data is received. The ANAO considers the Tax Office should aim to obtain legislated data online as a priority. There is merit in the Tax Office having a strategy to move to online data receipt to replace receipt of data as

tapes, discs or paper, especially for legislated data, but potentially for many third party datasets. The Tax Office's Channel Strategy[87] could provide guidance. It was, however, developed with a focus on communications between the Tax Office, taxpayers and tax agents in relation to tax returns and matters directly related to these.

4.42 Coupled with improved timeliness of access to the data, which online access could enable, the Tax Office could have greater scope to take immediate action on suspected serious non compliance. Under current arrangements, apart from problems of data quality, there can be substantial delays between the Tax Office's realisation of a suspected non compliance issue and the availability of the third party data needed to take action.

Intergovernmental cooperation: strategies for improving the Tax Office's ability to use external data

4.43 In Audit Report No.16 2006–07, the ANAO noted difficulties the Tax Office experienced in making optimum use of real property data because of difficulties associated with identity matching of real property transactions.[88] This is a component of COAG's *National Identity Security Strategy*. As noted in Audit Report No.47 2004–05,[89] the international inter-governmental body, the Financial Action Task Force, identified the real property data maintained by government authorities as relevant to managing the risks associated with money laundering and terrorist financing.

4.44 On 13 July 2007, the Attorney-General's Department released draft provisions setting out designated services which will be covered by the second tranche of the *Anti-Money Laundering and Counter-Terrorism Financing Act 2006*. This second tranche will require, among other things, the recording of some real property and conveyancing transactions and advisory services that meet specified criteria, such as thresholds or suspicious transactions. The record, to be provided to the AUSTRAC, and thereby forming a subset of the AUSTRAC database, will identify fully the provider of the service (e.g. a real estate agent or a solicitor), the vendor and the purchaser.

---

87 See paragraph 3.67 for an explanation about Tax Office's Channel Strategy.

88 See Audit Report No.16 2006–2007 *Administration of Capital Gains Tax Compliance in the Individuals Market Segment* paragraph 3.35 p. 58.

89 See Audit Report No.47 2004–2005 *Australian Taxation Office Tax File Number Integrity* paragraph 6.52 p. 116.

4.45 The ANAO notes there are two broader issues concerning the Tax Office's efficient access to data from federal and state government agencies. One is that of interoperability between information technology systems at the federal and state levels. The other issue is that of online receipt of data.

4.46 The Tax Office is examining these issues at the state level through the Australian Taxation and Revenue Offices Coordination Forum. As well, the issues are being examined at the federal and state levels in a project coordinated by Geosciences Australia. Aspects of both are also being examined as part of COAG's *National Identity Security Strategy*. Given that these issues (interoperability between information technology systems at the federal and state levels and the online receipt of data) are being examined by a range of agencies and in a number of fora, there may be scope for improved examination of these issues in ways that would be advantageous to tax administration.

4.47 The standard business reporting project, which aims to reduce the reporting burden for business through eliminating unnecessary or duplicated reporting, may assist in the development of improvements with respect to the quality of real property data, interoperability between information technology systems at the federal and state levels and the Tax Office's receipt of online data.

Ian McPhee
Auditor-General

Canberra ACT
24 April 2008

Appendices

Appendix 1: Agency Response

Australian Government
Australian Taxation Office

SECOND COMMISSIONER OF TAXATION

Mr David Crossley
ANAO Performance Audit Executive Director
Australian National Audit Office
GPO Box 707
CANBERRA ACT 2601

Contact Officer: Paul McKenzie
Telephone: (02) 621 61634
Facsimile (02) 621 66776
Issue Date:

Dear Mr Crossley

Re: AUSTRALIAN NATIONAL AUDIT OFFICE PEFORMANCE AUDIT OF THE AUSTRALIAN TAXATION OFFICE'S USE OF DATA MATCHING AND ANALYTICS IN TAX ADMNISTRATION

Thank you for your letter dated 4 March 2008 and for the opportunity to provide comments on the proposed report on the Australian Taxation Office's use of data matching and analytics in tax administration.

The Tax Office accepts the six recommendations as presented in the section 19 report.

I would like to thank the Australian National Audit Office audit team for the cooperative and professional manner they have adopted in working with us on this matter. I look forward to continuing the good working relationship developed in this performance audit.

Attached is the Tax Office response to recommendations (Annexure 1), and summary of our comments to be used in the report brochure (Annexure 2).

If you require further information on this matter, please contact Paul McKenzie, Acting Assistant Commissioner, Micro Enterprises and Individuals business line, on (02) 621 61634.

Yours Sincerely

Bruce Quigley
Second Commissioner
Law
Australian Taxation Office
4 April 2008

Cc: Richard Mackey Australian National Audit Office
Kirsty Stuart Australian National Audit Office

PO BOX 900 CIVIC SQUARE ACT 2608 AUSTRALIA (ADDRESS)
+61 (0)2 6216 1111 (TELEPHONE)
+61 (0)2 6216 2743 (FACSIMILE)

ANNEXURE 1

	Recommendation	ATO Response
1 2.15	The Australian National Audit Office recommends the Tax Office as the primary user of the Privacy Commissioner's Data Matching guidelines: ▪ Initiate policy-level discussion with the Office of the Privacy Commissioner with a view to updating the 1998 publication, "The Use of Data Matching in Commonwealth Administration – Guidelines"; and ▪ As appropriate, subsequently revise the Tax Office data matching protocol, to reflect any changes	**Agreed** The matter will again be raised with the Privacy Commissioner as part of the established meeting program.
2 3.41	The Australian National Audit Office recommends that in order to achieve additional efficiencies to tax administration and improve the integrity of key tax data bases, the Tax Office: ▪ improve non-individual identity matching; ▪ give further consideration to the negative search proof-of-concept facility; and ▪ continue to review options to improve its key identity matching facility.	**Agreed** The Tax Office will continue to look at opportunities to improve its identity matching capability in respect of both individual and non-individual taxpayers. Cost benefit considerations have precluded the development of a negative search facility. The Tax Office will continue to review the negative search proof of concept as new software and technology improvements and resources become available.
3 3.94	The Australian National Audit Office recommends that to simplify income tax returns by providing pre-filling for increased numbers of taxpayers, the Tax Office discuss with the Treasury options in relation to the pre-filling initiative, namely: ▪ Bringing the dates forward for the provision of requisite third party data; ▪ Mandating electronic transmission in a range of areas; and ▪ Including the tax file number on some additional data sets, having regard to the need to balance privacy concerns and improving the efficiency and effectiveness of public administration.	**Agreed** The Tax Office is committed to improve services to taxpayers through the pre-filling initiative, and recognises the need to bring forward the availability of data to support this. For 2007/2008 the focus has been on encouraging large providers to voluntarily lodge information early, and wherever possible, electronically. Early indications are that these providers recognise the benefits for their customers and employees and will co-operate. The Tax Office will discuss changes to enhance the pre-filling process with Treasury. Where the inclusion of a tax file number on additional data sets is considered warranted to support the treatment of compliance risks, the matter will be discussed with Treasury.

4 3.120	The Australian National Audit Office recommends that the Tax Office develop a three to five year corporate and strategic plan for the acquisition and use of external data.	**Agreed**
5 4.7	To minimise duplication and close gaps in responsibilities and streamline co-ordination arrangements, the Australian National Audit Office recommends the Tax Office clarify the roles of the various committees relevant to the management of the data matching and analytics capability, so their individual responsibilities are integrated more closely.	**Agreed**
6 4.20	To improve the efficiency and effectiveness of the corporate and strategic management of the data matching and analytics capability and inform the further development of the capability, the Australian National Audit Office recommends that the Tax Office complete a high level 'stock take' of achievements and lessons learned from its deployment.	**Agreed**

Appendix 2: Authority for acquisition of data

Provided data

Table A 1 (a)

Data provided to the Tax Office by taxpayers or their tax agents

Data Type / Description	Legislation
Income Tax Return	Part IV - *Income Tax Assessment Act 1936* (ITAA 1936) in particular - sec.161, 161A, 162
Schedules (e.g. CGT, Rental)	Part IV - ITAA 1936 in particular - sec.161, 161A, 162
Business Activity Statement	With respect to PAYG activities - Section 16-150 - Part 2-5 PAYG withholding, Schedule 1 *Taxation Administration Act 1953*. With respect to GST - GST Returns - *A New Tax System (Goods and Services Tax) Act 1999* - Div 31- see sec.31-15 re: approved form. With respect to Fringe Benefits Tax - Notification of instalment in approved form - see sec. 104 *Fringe Benefits Tax Assessment Act 1986* (i.e. a separate return is required for the annual FBT return - instalments required to be lodged with BAS).
Excise forms	The following are legislative requirements: • Section 39 of the Excise Act (Application for Licences) • Section 58 of the Excise Act (Entry for Home Consumption) • Regulation 52 of the Excise Regs (Application for Remission, Rebate or Refund) • Regulation 78A of the Excise Regulations (Conditions Relating to Payment of Drawback of Duty) The following are inferred in the legislation as the Collector must make a decision and it is based on information supplied by the taxpayer: • Section 61A of the Excise Act (Permission to Remove Goods that are Subject to the CEO's Control) • Section 61C of the Excise Act (Permission to Deliver Certain Goods for Home Consumption Without Entry) The following are administrative requirements that certain information is supplied to the Tax Office: • Excise Registration Form. This allows the Tax Office to create an Excise CAC on the client's AIS account. • Consent to Obtain Information Form. This provides permission for the Tax Office to obtain information on a client or company to complete the fit & proper persons checks under the Licensing part of the Excise Act (Sec39A (2)).
Self Managed Superannuation Funds forms	The *Superannuation Industry (Supervision) Act 1993*. Sec.254(1) Information is to be given to the regulator in an approved form. Sec.299U(1) provides that the Approved Form may require the TFN of the fund 'Regulator', which is defined under sec.10 to mean: (c) the Commissioner of Taxation if the provision in which it occurs is, or is being applied for the purposes of, a provision that is administered by the Commissioner of Taxation. (Note (a) and (b) not relevant to this definition.

Data Type / Description	Legislation
Application form for TFN, ABN	TFN - Sec.202B of the ITAA 1936. Where a TFN or ABN is quoted to an investment body, a TFN report must be lodged by the investment body under reg.55 of the *Income Tax Regulations 1936*. ABN - Sec.9 of *A NEW TAX SYSTEM (AUSTRALIAN BUSINESS NUMBER) ACT 1999* requires an applicant to apply for registration in the Australian Business Register. They must use a form approved by the Registrar of the Australian Business Register.

Source: Tax Office

Table A 1 (b)

Supplementary data provided to the Tax Office by taxpayers or their tax agents

Data Type / Description	Associated Schedules / Attachments
Individual tax return H/Individual Income tax return (NAT H/W 1371) *Produced for use by tax agents. Contains tax return for individuals, supplementary section, business and professional items.*	Baby bonus claim (NAT 6581) Capital gains tax (CGT) schedule (NAT 3423) Family tax benefit tax claim (NAT 4117) Family tax benefit SHORT tax claim (NAT 4153) Individual PAYG payment summary schedule (NAT 3647) Request for a determination of the deductible amount of UPP of a foreign pension or annuity (NAT 16543) Request for a determination of the deductible amount of UPP of an Australian pension or annuity (NAT 16544)
Tax return for individuals *Produced for self-preparers.*	Baby bonus claim (NAT 6581) Business and professional items schedule for individuals (NAT 2816) Capital gains tax (CGT) schedule (NAT 3423) Family tax benefit tax claim (NAT 4117) Family tax benefit SHORT tax claim (NAT 4153) Individual PAYG payment summary schedule (NAT 3647) Request for a determination of the deductible amount of UPP of a foreign pension or annuity (NAT 16543) Request for a determination of the deductible amount of UPP of an Australian pension or annuity (NAT 16544) Schedule of additional information – private health insurance policy details Schedule of additional information – item 10, 11, L1, T1 part C, M1, M2, M2 spouse's taxable income Tax return for individuals (supplementary section) (NAT 2679)
Non-lodgment advice *Produced for self-preparers. Available in TaxPack, Retirees TaxPack, Short tax return instructions and loose form.*	

Data Type / Description	Associated Schedules / Attachments
Tax return for retirees *Produced for self-preparers. This is a shorter version of the tax return and has its own RDCC transaction.*	Capital gains tax (CGT) schedule (NAT 3423) Request for a determination of the deductible amount of UPP of a foreign pension or annuity (NAT 16543) Request for a determination of the deductible amount of UPP of an Australian pension or annuity (NAT 16544) Schedule of additional information – private health insurance policy details, item 26 spouse's taxable income, item 27 spouse's taxable income
Tax return for individuals (supplementary section) *Produced for self-preparers. Once completed, is attached to the Tax return for individuals.*	Business and professional items schedule for individuals (NAT 2816) Request for a determination of the deductible amount of UPP of an Australian pension or annuity (NAT 16544)
Business and professional items schedule for individuals *Produced for self-preparers. Once completed, is attached to the Tax return for individuals. The tax return (supplementary section) must also be completed.*	Individual PAYG payment summary schedule (NAT 3647) Schedule of additional information – item 14
Individual PAYG payment summary schedule *Produced for self-preparers. Once completed, is attached to the Tax return for individuals.*	
Application for refund of franking credits for individuals *Produced for self-preparers.*	
Family tax benefit tax claim *Produced for self-preparers. Once completed, it can be lodged with the Tax return for individuals or Short tax return.*	Schedule of additional information - maintenance income details - relevant periods Schedule of additional information - dates of periods overseas Schedule of additional information - Spouse number 1 - maintenance income details - relevant periods Schedule of additional information – maintenance income details - relevant periods - in cases where there had been additional spouses Schedule of additional information - dependant number 1 - dates of periods overseas Schedule of additional information – dates of periods overseas – in cases where there had been additional dependents Schedule of additional information – dependants' circumstances Schedule of additional information – family profile (all families) Schedule of additional information – other employment-related income Schedule of additional information – return to work (relevant periods)
Family tax benefit short tax claim *Produced for self-preparers. Once completed, it can be lodged with the Tax return for individuals or Short tax return.*	

Data Type / Description	Associated Schedules / Attachments
Baby bonus claim *Produced for self-preparers. Once completed it can be lodged with the Tax return for individuals or Application for refund of franking credits, or on its own.*	
Trust tax return (NAT 0660) H/Trust income tax return (NAT HW 06060) *Produced for self-preparers and tax agents.*	Capital gains tax (CGT) schedule (NAT 3423) Schedule 25A (NAT 1125) Losses schedule (NAT 3425) Personal services income schedule (NAT 3421)
Partnership tax return (NAT 0659) H/Partnership tax return (NAT H/W 0659) *Produced for self-preparers and tax agents.*	Schedule 25A (NAT 1125) Personal services income schedule (NAT 3421)
Fund income tax and regulatory return (NAT 0658) *Produced for self-preparers and tax agents.*	Capital gains tax (CGT) schedule (NAT 3423) Losses schedule (NAT 3425)
Company tax return (NAT 0656) *Produced for self-preparers and tax agents.*	Research and development tax concession schedule (NAT 6708) Losses schedule (NAT 3425) Capital gains tax (CGT) schedule (NAT 3423) Capital allowances schedule (NAT 3424) Consolidated groups losses schedule (NAT 7888) Personal services income schedule (NAT 3421) Dividend and interest schedule (NAT 8030) Life Insurance companies taxation schedule - consolidated groups (NAT 7334CON) Life insurance companies taxation schedule (NAT 7334)
Franking account tax return (NAT 1382) *Produced for self-preparers and tax agents.*	NA. No associated Scehule/Attachment
Thin capitalisation schedule (NAT 6458) *Produced for self-preparers and tax agents.*	NA is a schedule attached to company return 2007
Strata title body corporate tax return (NAT 4125) *Produced for self-preparers and tax agents.*	Company tax return (NAT 0656)
Non-individual PAYG payment summary schedule (NAT 3422) *Produced for self-preparers and tax agents.*	

Data Type / Description	Associated Schedules / Attachments
Interposed entity election (NAT 2788) *Produced for self-preparers and tax agents.*	
Family trust election and/or family trust revocation (NAT 2787) *Produced for self-preparers and tax agents.*	
Venture Capital Deficit tax return (NAT 3309) *Produced for self-preparers and tax agents.*	

Source: Tax Office.

Note: The above tables (Table A1 (a) and A1 (b)) were prepared by the Tax Office at the request of the ANAO. Whilst the data is a comprehensive list as at February 2008, it can not be relied upon as an exhaustive list of all categories of data provided to the Tax Office by taxpayers or their tax agents pursuant to all legislative, regulatory or Tax Office administrative requirements.

Legislated data

Table A 2

Legislated datasets the Tax Office receives regularly, and the legislation under which the data is provided

Data type/description	Legislation
Annual Investment Income Reports (AIIR)	Financial institutions, share registries and other payers of investment income report AIIR data to the Tax Office in accordance with reporting obligations under Sections 202D and 202G of the *Income Tax Assessment Act 1936*, and Regulation 56 of the Income Tax Regulations 1936
AUSTRAC reportable transactions	The *Anti-Money Laundering and Counter-Terrorism Financing Act 2006* gives the Tax Office the legal right to access AUSTRAC data. There is a memorandum of understanding between both agencies to administer access to data and related data transfers.
Dividend and Interest Statements	Dividend and Interest Statements are reported under obligations created by Subsection 161A(1) of the *Income Tax Assessment Act 1936*, and Sections 388–50 of the *Tax Administration Act 1953*
Family Tax Benefit	Family Tax Benefit is governed by *A New Tax System (Family Assistance) Act 1999*, and *A New Tax System (Family Assistance) (Administration) Act 1999* There are cross references (and definitions) in the *Social Security Act 1991* and the *Social Security (Administration) Act 1999*

Data type/description	Legislation
Foreign Resident Withholdings	Under Section 16–153 of Schedule 1 to the *Tax Administration Act 1953*, payers are required to submit an annual report to the Tax Office with details of certain payments made to foreign residents
Foreign Sourced Income	The tax authorities in various countries supply this data in accordance with reporting obligations under existing tax treaties with those countries
Inbound Mail Management System	Section 202 of the *Income Tax Assessment Act 1936* (TFN legislation) Income Tax Ruling 36 (updating client records)
PAYG Payment Summaries	Under Section 16–153 of Schedule 1 to the *Tax Administration Act 1953* payers and other organisations are required to report to the Tax Office details of payment summaries issued to their payees during the financial year
Private health insurance data via Medicare Australia	This data is received from Medicare Australia and supplied to the Tax Office in accordance with a reporting obligation contained within Medicare Australia's governing legislation, the *Private Health Insurance Incentives Act 1998*
Quarterly TFN Reports (of TFN withholding tax withheld by financial institutions)	Regulation 55(1)
Reasonable Benefit Limit	This reporting provides complete and accurate information to the Commissioner when eligible termination payments are paid on or after 1/7/94 (for employers, if eligible termination payment is less than $5,000, it is not required to be reported). Information is specified in the regulations and must be sufficient to enable a final determination to be made. *Section 140M ITAA 1936.* Reasonable Benefit Limit notices are to be lodged by the end of the 14th day of the month after the month the payer makes an eligible termination payment (where the value is greater than $5,000) or commences to make payments of a superannuation pension or an immediate annuity on or after 1/7/94. *Section 140M (3) ITAA 1936.* Payers Of Benefits – Specified Information (Act S 140m) Regulation 53K For each of the following provisions of the Act, the payer mentioned in the provision must give to the Commissioner a notice containing the information set out in Schedule 2B: (a) sub-section 140M(1); (b) paragraph 140M(1A)(f); (c) paragraph 140M(1C)(f).

Data type/description	Legislation
Student Financial Supplement Scheme (SFSS data) + HECS	Section 154–55 of *Higher Education Support Act 2003* provides that Higher Education Providers are required to give the Commissioner any information in relation to students that the Commissioner needs for administration of their debt. Section 1061ZZEV of the *Social Security Act 1991* and section 12ZH of the *Student Assistance Act 1973* have similar requirements in relation to SFSS debts.
Super Fund Lost Members Register	The *Superannuation (Unclaimed Money and Lost Members) Act 1999* and its Regulations outline responsibilities for data providers and the Tax Office with respect to the Lost Members Register.
Super Fund Member Contribution Statements	Reporting of Member Contribution Statements is required by the following legislation: *Superannuation Contributions Tax (Assessment and Collection) Act 1997* Section 13, from 1 July 1996 to 30 June 2005 *Superannuation (Government Co-contributions for Low Income Earners) Act 2003* Sub-section 26(1), from 1 July 2006 to 30 June 2008 *Superannuation Guarantee (Administration) Act 1992* Section 78, from 1 July 2005 to 30 June 2008 NOTE: From 1 July 2008 the Superannuation Fund Members Contribution Statements will be covered under the *Tax Administration Act 1953*
Super Member Exit Statements	*Super (Constitutionally Protected Funds) A & C Act 1997* Pt 3, sec 12 (2) & (3) – Superannuation providers to give statements
TFN Declarations (received daily)	The legislation covering TFN declaration reporting is Section 202C–202CF of the *Income Tax Assessment Act 1936* and for electronic reporting the *Electronic Transactions Act 1999*, Section 11.
Welfare payments and Centrelink debtors	Centrelink provides the following data to the Tax Office under cover of its own legislation: Details of benefits income paid to clients, which the Tax Office uses to identify any undeclared government income. Centrelink client outstanding debts. The Tax Office takes garnishee action against refund cheques as a means of recovering amounts outstanding to Centrelink. Data the Tax Office returns to Centrelink in relation to these exchanges is provided under the *Data Matching Program (Assistance and Tax) Act 1990*, Section 16(4)(e) of the *Income Tax Assessment Act 1936* and *A New Tax System (Family Assistance Administration) Act 1999*.

Source: Tax Office

Memorandum of understanding data

Table A 3

Datasets acquired through a memorandum of understanding

Data type/description	Memorandum of understanding
Australian Securities and Investments Commission	The memorandum of understanding between the Tax Office and ASIC is to enhance liaison and relationships between the agencies. The memorandum of understanding set the information sharing arrangement between ASIC and the Tax Office to help administer or enforce the particular laws for which each agency is responsible.
State Revenue Offices	A 2005 memorandum of understanding governs exchange of confidential information and other aspects of cooperation between the Tax Office and participating state and territory revenue offices.

Source: Tax Office

Purchased data

Table A 4

Datasets the Tax Office purchases

Data type	Description
Australian Electoral Commission	The Tax Office purchases electoral roll data on a periodic basis.
Fact of Death data	The Tax Office purchases Fact Of Death data from the State and Territory government Offices of Registrar-General, on a periodic basis.
ASIC	Purchased by Large Business and International to risk profile taxpayers.
Reed Construction Data	The purchased data represents building approval information, relative to non-individuals, acquired by GST and Small and Medium Enterprises Business Lines to support their building and construction audit program. The data was initially purchased in 2005-06 and the Tax Office continues to receive updates on a regular basis. GST has since determined that the data no longer adequately supports their needs but Small and Medium Enterprises Business Line continue to find value in the data.

Source: Tax Office

Datasets the Tax Office requisitioned on an ad hoc basis for data matching projects

Between January 2000 and June 2007, the Tax Office gazetted 47 data matching projects.

The data used in these projects may have been voluntarily provided to the Tax Office; or one or more of the following legal authorities were used to compulsorily requisition the data:

- section 264 of the *Income Tax Assessment Act 1936*;
- section 353–10 of the *Taxation Administration Act 1953*;
- section 128 of the *Fringe Benefits Tax Assessment Act 1986*; or
- section 77 of the *Superannuation Guarantee (Administration) Act 1992*.

Table A 5

Tax Office's gazetted data matching projects 2000–07

No	Gazette number and date	Project name	Data sources	Project description
1	GN 3, 27 Jan 2000	Autosearching	AUSTRAC	The Tax Office electronically matched names and addresses from certain sections of Tax Office data holdings to reveal financial transaction information that related to the Tax Office names and addresses.
2	GN 39, 2 Oct 2002	Medical Profession	State Medical Practitioners Register	The Tax Office electronically matched names and addresses with certain sections of Tax Office data holdings to reveal medical practitioners' compliance with lodgement and payment obligations under taxation law.
3	S 145, 8 May 2003	Profession/ industry	State Boards for Architects	The Tax Office electronically matched names and addresses with certain sections of Tax Office data holdings to identify non compliance with lodgement and payment obligations under taxation law.
4	GN 37, 17 Sep 2003	Prison population	All State Department of Correctional Services	The Tax Office electronically matched names of people incarcerated in all states of Australia with certain sections of Tax Office data holdings to identify non compliance with lodgement and payment obligations under taxation law.

No	Gazette number and date	Project name	Data sources	Project description
5	GN 42, 22 Oct 2003	Horse racing	Stakes Payment System of the NSW Thoroughbred Racing Board	The Tax Office cross-referenced with its internal data to identify entities that were registered for GST, received income and/or prize money and accounted erroneously in relation to their GST or income tax liability; entities that were not registered for GST but should be and identified associated potential GST and income tax liabilities; and to produce relevant case selection procedures and implement appropriate educational campaigns.
6	GN 42, 22 Oct 2003	Victorian Work Cover	Victorian Work Cover Authority	The Tax Office electronically matched business names and addresses with certain sections of Tax Office data holdings to identify non compliance with registration, lodgement and payment obligations under taxation law.
7	GN 43, 29 Oct 2003	Pubs and clubs	State and Territory Gaming and Liquor Licences Register	The Tax Office electronically matched with taxpayers' income tax return information and business activity statement database to identify entities that are not registered for GST but should be and identify associated potential GST liabilities.
8	GN 44, 5 Nov 2003	Dairy adjustment	Dairy Adjustment Authority	The Tax Office electronically matched taxpayers' income tax return information held on the Tax Office's tax return database with data regarding payments made under the Dairy Structural Adjustment Program.
9	GN 48, 3 Dec 2003	Disposal of asset	Motor vehicle registrars throughout Australia; state and territory agencies responsible for surveying commercial vessels; state and territory agencies responsible for administering land titles, sale of real property and sale of businesses; Intellectual Property Australia; Civil Aviation Safety Authority; and Esanda Finance	The Tax Office cross-referenced with its internal data to identify entities that were registered for GST, disposed of capital assets and accounted erroneously in relation to their GST liability; entities that were not registered for GST but should be and identified associated potential GST liabilities; and to produce relevant case selection procedures and implement appropriate educational campaigns.

No	Gazette number and date	Project name	Data sources	Project description
10	S 452, 4 Dec 2003	Accounting profession	CPA Australia; National Institute of Accountants; National Tax & Accountants Association; Institute of Chartered Accountants Australia; Association of Taxation and Management Accountants; Taxation Institute of Australia	The Tax Office electronically matched membership information with certain sections of Tax Office data holdings to identify non compliance with registration, lodgement and payment obligations under taxation law.
11	GN 2, 14 Jan 2004	Barter Member Industry Group	Barter Exchanges	The Tax Office electronically matched names and addresses of barter members with certain sections of Tax Office data holdings to identify non compliance with lodgement and payment obligations under taxation law.
12	GN 3, 21 Jan 2004	Gold bullion, antiques and art dealings	PJ Williams & Associates; Twin Plaza Metals; Universal Coin Company; Dascorp; Johnson Matthey; E & R Gould Unit Trust; Anthony Filippini Family Trust; Graham Cornall Antiques; Sotheby's Australia; Christie's Australia; The Antiques Garden Unit Trust; Scott Livesey; and Editions Australia	The Tax Office requested and collected purchase and sale transaction details, including names and addresses, for the period 1 July 2000 to 30 June 2003. The data was electronically matched with certain sections of Tax Office data holdings to identify non compliance in terms of incorrect or non lodgement and payment of taxation obligations; and/or registration obligations under taxation law.
13	GN 12, 24 Mar 2004	DEST Apprenticeship Grants	Department of Education, Science and Training	The Tax Office electronically matched taxpayers' information held on the Tax Office's databases with data regarding payments made under the Commonwealth New Apprenticeships Incentive Scheme.
14	GN 17, 28 Apr 2004	Victorian State Revenue Office	Victorian State Revenue Office	The Tax Office matched business names, addresses, and details pertaining to property title transfers with certain sections of Tax Office data holdings to identify non compliance with registration, lodgement and payment obligations under taxation law.
15	S 217, 22 Jun 2004	Financial planner compliance matching	Australian Securities and Investment Commission	The Tax Office electronically matched details with certain sections of Tax Office data holdings to review licensed securities dealers and investment advisers, and their representatives, registered with ASIC to determine their level of tax compliance; and to identify areas of non compliance and develop strategies to address these areas and enhance tax system integrity.

No	Gazette number and date	Project name	Data sources	Project description
16	GN 31, 4 Aug 2004	Extension of disposal of asset	Motor vehicle registrars throughout Australia; state and territory agencies responsible for surveying commercial vessels; state and territory agencies responsible for administering land titles, sale of real property and sale of businesses; financial institutions; Intellectual Property Australia; and Civil Aviation Safety Authority	The Tax Office cross-referenced with its internal data to identify entities that were registered for GST, disposed of capital assets and accounted erroneously in relation to their GST liability; entities that were not registered for GST but should be and identified associated potential GST liabilities; and to produce relevant case selection procedures and implement appropriate educational campaigns
17	GN 46, 17 Nov 2004	Motor vehicle auction house	Pickles Auctions; Fowles Auction Group; Auto Auction Group	The Tax Office electronically cross matched and matched names, addresses and telephone numbers with certain sections of Tax Office data holdings to identify possible non compliance with lodgement and payment obligations under taxation law.
18	GN 12, 30 Mar 2005	BSA	Building Services Authority (Queensland)	The Tax Office electronically matched licence registration formation with certain sections of Tax Office data holdings to identify clients with any outstanding tax returns or activity statements; clients that may have incorrectly declared income on a tax return and that may have incorrectly declared total supplies on an activity statement; and clients that were not meeting other tax obligations (i.e. GST registration)
19	GN 13, 6 Apr 2005	Horse racing	Queensland Racing; Racing Victoria Limited; Tasmanian Thoroughbred Racing Council; Darwin Turf Club; Thoroughbred Racing SA Limited; Racing and Wagering Western Australia	The Tax Office cross-referenced with its internal data to identify entities that were registered for GST, received income and/or prize money and accounted erroneously in relation to their GST or income tax liability; entities that were not registered for GST but should be and identified associated potential GST and income tax liabilities; and to produce relevant case selection procedures and implement appropriate educational campaigns.
20	GN 14, 13 Apr 2005	RPDATA	RPDATA Pty Ltd	The Tax Office electronically matched names and addresses of entities within the real property market with certain sections of Tax Office data holdings to identify non compliance with lodgement and payment obligations under taxation law.

No	Gazette number and date	Project name	Data sources	Project description
21	GN 16, 27 Apr 2005	Removalist industry	Desk Top Marketing System; Australia On Disk; Infot Inc.; Australia Business Database; and Australian state and territory motor vehicle (truck) registration data	The Tax Office electronically matched names and addresses of removalist industry members with certain sections of Tax Office data holdings to identify non compliance with reporting obligations under taxation law.
22	GN 18, 11 May 2005	Motor vehicle retailing and wholesaling	All state and territory licensing bodies	The Tax Office electronically cross matched information provided in applications for motor vehicle retailer and wholesaler licenses with certain sections of Tax Office data to identify possible non compliance with lodgement and payment obligations under taxation law.
23	GN 20, 25 May 2005	Adult services	ACT Dep. of Fair Trading; NSW Dep. of Fair Trading; All local government councils in NSW; Queensland Crime and Misconduct Commission; Queensland Prostitution Enforcement Taskforce; Queensland Dep. of Tourism, Racing and Fair Trading; Queensland Office of Fair Trading; SA Office of the Liquor and Gaming Commissioner; Tasmanian Dep. of Justice; Tasmanian Dep. of Treasury & Finance; Victorian Dep. of Justice; Victorian Office of Consumer Affairs; WA Dep. of Racing, Gaming and Liquor; SENSIS (Yellow Pages); and Telstra	The Tax Office electronically matched names and addresses of entities registered as providers of adult services with Tax Office data holdings to identify possible non compliance with registration, lodgement and payment obligations under taxation law.

No	Gazette number and date	Project name	Data sources	Project description
24	S 90, 31 May 2005	Capital Gains Tax	NSW Office of State Revenue; NSW Dep. of Land Property Information; Victorian State Revenue Office; ACT Planning Authority; ACT Land Titles Office; ACT Revenue Office; NT Treasury; Queensland Office of State Revenue; Queensland Dep. of Natural Resources and Mines; Tasmanian Dep. of Primary Industries, Water and Environment; Tasmanian Dep. of Treasury and Finance; SA Revenue Office; SA Land Services Group; and WA Office of State Revenue	This notice replaced the previously GN 20, 25 May 2005 and S 87, 30 May 2005 (not mentioned in this appendix) regarding the same data matching project. The Tax Office electronically matched business names, addresses, and details pertaining to property title transfers with certain sections of Tax Office data holdings to identify non compliance with CGT, income tax and GST obligations under taxation law.
25	GN 29, 27 Jul 2005	Workcover SA	Workcover Corporation South Australia	The Tax Office electronically matched business names and addresses with certain sections of Tax Office data holdings to identify non compliance with registration, lodgement and payment obligations under taxation law.
26	GN 42, 26 Oct 2005	Low Doc Loans	PMI Mortgage Insurance; GE Mortgage Insurance; St George Bank; Australia and New Zealand Banking Group; Westpac Banking; and Suncorp–Metway	The Tax Office electronically matched names and addresses of entities who have taken out insurance on low doc loans with certain sections of Tax Office data holdings to identify non compliance with lodgement and payment obligations under taxation law.
27	GN 43, 2 Nov 2005	Workcover NSW	Workcover New South Wales	The Tax Office electronically matched business names and addresses with certain sections of Tax Office data holdings to identify non compliance with registration, lodgement and payment obligations under taxation law.
28	GN 47, 30 Nov 2005	Workcover Qld	Workcover Queensland	The Tax Office electronically matched business names and addresses with certain sections of Tax Office data holdings to identify non compliance with registration, lodgement and payment obligations under taxation law.

No	Gazette number and date	Project name	Data sources	Project description
29	GN 48, 7 Dec 2005	Trades Compliance	Office of Consumer and Business Affairs – South Australia	The Tax Office electronically matched details of licensed builders, electricians, plumbers and gas fitters with certain sections of Tax Office data holdings to identify non compliance with registration, lodgement and payment obligations under taxation law.
30	GN 49, 14 Dec 2005	Motor Vehicle Auction House	Pickles Auctions; Fowles Auction Group; and Auto Auction Group	The Tax Office electronically matched names, addresses, and telephone numbers and other information with certain sections of Tax Office data holdings to identify possible non compliance with lodgement and payment obligations under taxation law.
31	S 232, 16 Dec 2005	RPDATA	RPDATA Pty Ltd	The Tax Office electronically matched names and addresses of entities within the real property market with certain sections of Tax Office data holdings to identify non compliance with lodgement and payment obligations under taxation law.
32	GN 1, 11 Jan 2006	Legal profession (including judiciary)	Attorneys General and state and territory Legal Practitioners Registers	The Tax Office electronically matched names and addresses with certain sections of Tax Office data holdings to reveal non compliance with lodgement and payment obligations under taxation law.
33	GN 15, 19 Apr 2006	Luxury vehicle	NSW Roads and Traffic Authority; Queensland Transport; Vic Roads; Tasmanian Dep. of Infrastructure, Energy and Resources; Transport SA; WA Dep. for Planning and Infrastructure; NT Dep. of Planning and Infrastructure (Transport Division); and ACT Road Transport Authority	The Tax Office electronically matched details of individuals or entities that had purchased or acquired a motor vehicle valued at $70,000 or higher with certain sections of Tax Office data holdings to identify non compliance with lodgement and payment obligations under taxation law.
34	GN 25, 28 Jun 2006	Marine vessel	Australian Maritime Safety Authority; NSW Maritime Authority; Maritime Safety Queensland; Marine Safety Victoria; Marine and Safety Tasmania; SA Dep. for Transport, Energy and Infrastructure; WA Dep. for Planning and Infrastructure; and NT Dep. of Planning and Infrastructure	The Tax Office electronically matched details of individuals or entities who registered a marine vessel with certain sections of Tax Office data holdings to identify non compliance with lodgement and payment obligations under taxation laws.

No	Gazette number and date	Project name	Data sources	Project description
35	GN 30, 2 Aug 2006	Fishing industry	Australian Fisheries and Management Authority; Australian Maritime Safety Authority; WA Dep. of Fisheries; WA Office of State Revenue; Transport WA; Queensland Dep. of Fisheries; Queensland Office of State Revenue; Queensland Dep. of Transport; Tasmanian Dep. of Primary Industries, Water and Environment; Tasmanian State Revenue Office; Tasmanian Marine and Safety; NSW Fisheries; NSW Office of State Revenue; NSW Maritime Authority; Victorian Dep. of Primary Industries; Marine Safety Victoria; SA Dep. of Primary Industries and Resources; Revenue SA; Transport SA; NT Dep. of Primary Industry and Fisheries; NT Territory Revenue Management; and NT Dep. of Transport (Marine Branch)	The Tax Office electronically matched information about commercial fishing licence holders, commercial boat licence holders and registrations, and transfers of fishing licences and unit entitlements with Tax Office data to identify non compliance with lodgement and payment obligations under taxation law.
36	GN 33, 23 Aug 2006	Aircraft	Civil Aviation Safety Authority; Recreational Aviation Australia; and Australian Sports Rotorcraft Association	The Tax Office electronically matched details of individuals or entities that had purchased or acquired an aircraft with certain sections of Tax Office data holdings to identify non compliance with lodgement and payment obligations under taxation law.

No	Gazette number and date	Project name	Data sources	Project description
37	GN 37, 20 Sep 2006	Real property	NSW Office of State Revenue; NSW Dep. of Lands; Victorian State Revenue Office; ACT Planning and Land Authority; ACT Land Titles Office; NT Treasury; NT Dep. of Planning and Infrastructure; NT Land Titles Office; Queensland Office of State Revenue; Queensland Dep. of Natural Resources, Mines and Water; Tasmania State Revenue Office; Tasmania Dep. of Primary Industries and Water; SA Revenue Office; SA Land Services Group; and WA Office of State Revenue	This notice replaced the S 90, 31 May 2005 regarding the capital gains tax data matching project. The Tax Office electronically matched identity and transaction details pertaining to property title transfers with certain Tax Office data holdings to identify non compliance with capital gains tax, income tax and goods and services tax obligations under taxation law.
38	GN 37, 20 Sep 2006	Residential tenancies authorities	Queensland Residential Tenancies Authority; NSW Office of Fair Trading; and Victorian Residential Tenancies Bond Authority.	The Tax Office electronically matched names and address details pertaining to residential tenancy agreements with certain Tax Office data holdings to identify non compliance with rental income, capital gains tax and other income tax obligations under taxation law.
39	GN 37, 20 Sep 2006	Foreign resident	Foreign Investment Review Board	The Tax Office electronically matched details pertaining to property title transfers by non-residents, their names, and addresses with certain Tax Office data holdings to identify non compliance with lodgement and payment obligations under taxation law.
40	GN 37, 20 Sep 2006	Shares data	Link Market Services Limited; Computershare Limited; Australian Stock Exchange Limited; and Registries Limited	The Tax Office electronically matched identity and transaction details pertaining to securities with certain Tax Office data holdings to identify non compliance with capital gains tax, income tax and goods and services tax obligations under taxation law.

No	Gazette number and date	Project name	Data sources	Project description
41	GN 37, 20 Sep 2006	Horse industry	Stakes Payment System of the New South Wales Thoroughbred Racing Board	This notice was supplementary to GN 42, 22 Oct 2003 regarding the horse racing data matching project. The Tax Office electronically matched details pertaining to horse trainers, owners, jockeys, stable-hands, and track-riders, including records relating to individuals, partnerships, companies, and trusts; with certain sections of Tax Office data holdings to identify non compliance with lodgement and payment obligations under taxation law.
42	GN 44, 8 Nov 2006	Workcover Tasmania	Workcover Tasmania	The Tax Office electronically matched business names and addresses with certain sections of Tax Office data holdings to identify non compliance with registration, lodgement and payment obligations under taxation law.
43	GN 44, 8 Nov 2006	Workcover West Australia	Workcover Western Australia	The Tax Office electronically matched business names and addresses with certain sections of Tax Office data holdings to identify non compliance with registration, lodgement and payment obligations under taxation law.
44	GN 47, 29 Nov 2006	Taxi industry	Victorian Taxi Directorate and Queensland Transport	The Tax Office electronically matched names and addresses of entities within the taxi industry with certain sections of Tax Office data holdings to identify non compliance with lodgement and payment obligations under taxation law.
45	GN 14, 13 Dec 2006	Shopping centre retailers	Westfield Management; Colonial First State Property Management; Lend Lease Property Management (Aust); Stockland Property Management; AMP Capital Investors; Centro Properties Group; and General Property Trust	The Tax Office electronically matched names and addresses of entities that operate as shop owners within major shopping centres with certain sections of Tax Office data holdings to identify non compliance with lodgement and payment obligations under taxation law.
46	GN 3, 24 Jan 2007	NT WorkSafe	NT WorkSafe	The Tax Office electronically matched business names and addresses with certain sections of Tax Office data holdings to identify non compliance with registration, lodgement and payment obligations under taxation law.

No	Gazette number and date	Project name	Data sources	Project description
47	GN 18, 9 May 2007	Integrity of Identity Data Pilot	Dep. of Immigration and Citizenship; Centrelink; Medicare Australia; and NSW Registrar of Births, Deaths and Marriages	This notice referred to development of the National Identity Security Strategy project announced by the Australian Government in April 2005 and was later endorsed by COAG in September 2005. This pilot is being managed by the Attorney-General's Department in partnership with the Tax Office. It will test the integrity of a sample of 25,000 records of the Tax Office's tax file number database. It includes cross-agency data matching against the identity registers of a range of government agencies.

Source: ANAO analysis of information in the Commonwealth Gazette Notices

Appendix 3: Tax Office's data mining, modelling, mapping and analytics projects

1. The Tax Office expects its analytics projects to transform data and information into knowledge to support decision-making activities. The Tax Office expects that the projects will enable application of consistent and rigorous methods to produce profiles, reports and models that will help the Tax Office better understand taxpayer behaviours, and predict future behaviours. The Tax Office expects these projects will lead to a greater understanding of taxpayer groups and individuals so they can be treated in a manner that suits their compliance behaviour. If these objectives are achieved the analytics projects should enable the Tax Office to interact with the community and taxpayers in ways that are less intrusive and more appropriate, efficient, and effective.

2. The Tax Office has three data analytics and mining groups in the Chief Knowledge Office business line working on a wide range of projects. These include risk analyses and predictive modelling of tax compliance behaviour in relation to compliance risks in the areas of tax agents, high risk refunds, debt and lodgement, operation Wickenby, aggressive tax planning, tax havens, under reporting of foreign sourced income, and non-compliance with employer obligations.[90] The three groups liaise with business lines to develop taxpayer profiles and predictive models.

3. This work has enabled the Tax Office to be more targeted and more efficient in relation to lodgement enforcement; to have automated action about very low risk refunds; to use GST and cash economy models to be more targeted and more efficient in relation to audit and enforcement; and to better manage superannuation guarantee obligation case work.

[90] The Australian Government established Project Wickenby after information became available that international tax scams were operating in Australia and some Australians had engaged in them in a deliberate effort to take income outside the operation of Australia's tax laws. In 2006 the Australian Government provided $305 million for the project, which brings together the Australian Tax Office as the lead agency, the Australian Crime Commission, the Australian Federal Police, the Australian Securities and Investments Commission and the Commonwealth Director of Public Prosecutions. Statement by the Treasurer of the Commonwealth of Australia Thursday, 20 July 2006. <http://www.treasurer.gov.au/tsr/content/transcripts/2006/108.asp> .

Table A 6

Analytics projects 2005 to 2007

No.	Project/Model	Objective
1	All Obligations – Whole-of-client	To bring a consistent modelling approach based on the framework of predicting the change in tax that would result from a review, and generating a client risk score for specific areas of the Client Risk Score matrix.
2	Registration – Identity crime	To help automate the processes used to identify people who have either created or taken over another person's identity.
3	Lodgement – Identification of client risk	Four phases are to be delivered over 3½ years (from 2005 to 2008), relating to outstanding lodgement obligations. Phase 1 – Predict risk to revenue. Phase 2 – Predict risk to revenue of all outstanding lodgements. Phase 3 – Stage 1. Predict propensity to lodge and risk to revenue. Stage 2. Predict risk to reputation. Stage 3. Predict risk of missing information. Phase 4 – Risk score accumulated at associated entities/groups of clients.
4	Lodgement – Strategy evaluation and improvement; and propensity to lodge	Determine the effectiveness of existing treatment strategies across all clients in the known non lodging and late lodging population. Determine appropriate treatment strategy for all clients. Estimate the probability that a client will lodge within a certain time period of the due date.
5	Lodgement – Failure to lodge	Measure effectiveness of late and non lodged penalty impositions, and failure to lodge warnings for non lodgement and late lodgement. Determine imposition gaps. Identify categories of clients where penalty imposition for non lodgement is likely to result in lodgement and improvement to future lodgement compliance behaviour. Profile clients according to their propensity to lodge (i.e. identify common characteristics of those likely/not likely to respond to failure to lodge penalty or warnings).
6	Lodgement – Shift to instalments	Ability to differentiate clients based on their risk of non compliance if they take up an instalment option.
7	Reporting – GST field	Improve strike rate and return-on-investment for compliance cases generated for GST field.

No.	Project/Model	Objective
8	Reporting – Tax agent micro clients	Identify entities effecting gross tax in micro clients other than taxable income and related with a view to predicting gross tax.
9	Reporting – Low risk gross interest	To provide a list of clients with low or no 'gross interest' risk for earlier closure.
10	Reporting – Low risk 'salary and wages'	To provide a list of clients with low or no risk in salary or wages for earlier closure.
11	Reporting – Risk scores	To provide a risk score for each individual client.
12	Reporting – Promoters	To identify promoters of tax avoidance and tax evasion schemes.
13	Reporting – Super guarantee	Build an analytic model to risk-rate employers in regard to their superannuation guarantee obligations and assist with case selection.
14	Reporting – Super guarantee employee notification	The developed analytic model identifies high and low risk employee notification cases more accurately than the existing activity rating rules and increases detection accuracy.
15	Reporting – High risk refunds (income tax)	Review current processing of pre-issue high risk refund cases to improve strike rate and increase the return-on-investment by reducing the overall number of non-productive cases being reviewed.
16	Reporting – Small and medium enterprise tax agents and clients	To identify potential high-risk tax agents with small and medium enterprise clients.
17	Reporting – Micro tax agents and clients	To identify potential high-risk tax agents who have substantial micro business client components.
18	Reporting – Small and medium enterprise clients	To identify potential high-risk small and medium enterprise clients.
19	Reporting – Operation Wickenby	To detect previously unknown onshore (Australian) entities linked to known promoters operating out of a specified tax haven or other tax haven countries. To analyse the tax behaviour of the promoters and identify relationships, patterns and trends. To analyse Operation Wickenby data to identify patterns and trends indicative of emerging risks with promoter schemes. Develop classification and prediction models that help detect promoters and associated entities.
20	Reporting – Micro clients	To identify potential high-risk micro clients.
21	Reporting – Employer obligation	Case selection for non-compliant clients who under-withhold income tax withholding.

No.	Project/Model	Objective
22	Reporting – Conspicuous consumption clients	To identify individuals who have high spending patterns that are inconsistent with their incomes and who are using tax havens, schemes and/or the cash economy to disguise their income.
23	Reporting – Cash economy pubs and clubs; restaurants and cafes; and construction	To improve the strike rate of cases reviewed as high risk in unreported income.
24	Reporting – General cash economy	To improve the strike rate of cases reviewed as high risk in unreported income.
25	Reporting – High risk refunds/incorrect credits (GST)	To review current processing of identifying pre-issue high risk refund cases. To improve targeting of high risk cases to avoid unnecessary delays in refunds and to increase the return-on-investment by reducing the overall number of non-productive cases being reviewed.
26	Accounting – Legal and complex cases	Model the payment risk of legal and complex case management debt cases for case selection prioritisation.
27	Accounting – Propensity to pay	A review of the effectiveness of the legal and complex risk model.
28	Accounting – Income tax withholding	To demonstrate the goodness of fit of the risk model to the large income tax withholding client population. To provide client segments based on payday periodicity or aperiodicity.
30	Accounting – Capacity to pay	To identify taxpayers with low capacity to pay due to financial distress or insolvency.
31	Accounting – Propensity to pay (income tax and GST)	Improve efficiency of revenue collection and debt case finalisation by choosing treatments based on client propensity to pay.
32	Accounting – C2P & P2P (Version 2)	To provide a high level description of the purpose and intent for the design of the Capacity to Pay (C2P) and Propensity to Pay (P2P) risk models.
33	Account – Referrer project	Model the payment risk of debt cases for selection for referrer project. (This is a pilot project to understand how third parties could collect debt for the Tax Office.)

Source: ANAO analysis of Tax Office's Analytics Project Fact Sheets.

Glossary

Advanced Analytics	The application of the analytics discipline using sophisticated technology to deliver **analytics models** from very large collections of data. These models use data mining techniques to discover relationships from the data and infer future behaviour or events. The models can be categorised as descriptive, predictive or decision-oriented. Such models are considerably more complex and more dynamic than **expert business rules** engines and generally identify a finer level of detail and understanding of the populations in the data. Advanced analytics typically requires considerably more computational processing of the data due to the sophisticated algorithms they utilise. – *Draft Common Business Language term, Authorised by the CKO, September 2007*
Analytics	A discipline that identifies patterns, relationships and trends from data, using a variety of mathematically based technologies principally drawn from statistics and **data mining**. Most broadly, analytics covers what might be called **basic analytics**, including data exploration and aggregation, and **advanced analytics**, which uses data mining technology for discovery and model building purposes. Using statistical and data mining technologies, significantly more complex relationships within and between entities (e.g. taxpayers) can be discovered and modelled, based on analyses over very large populations of collected data. Analytics is assisted by the use of good **data matching** and **data linking** techniques which improve the quality and value of data inputs available to a data miner. Conversely, analytics can also provide technology to assist data matching and data linking activities. – *Draft Common Business Language term, Authorised by the CKO, September 2007*

Basic Analytics	The application of the **analytics** discipline using fundamental statistical methods and online analytical processing technology. This includes data exploration and aggregation, including cube analysis, as well as statistical profiling and analysis. Analytics (statistical and data mining technology) is used in many areas, including **data matching** (matching data relating to the same entity) and **data linking** (linking entities through common characteristics). – *Draft Common Business Language term, Authorised by the CKO, September 2007*
Channels	Reporting channels for legislated data are detailed within various reporting specifications issued by the Tax Office. The channels available to third-party data providers are: • Paper based reporting; • Magnetic media (Computer floppy disk or magnetic cartridge tape). The Tax Office's Magnetic Information Processing Services section decided to withdraw the option of reporting on magnetic cartridge tape from June 2006 as a result of the industry wide phasing out of this medium, but has extended its use on a 'case by case' basis to assist some larger reporters in their transition to other reporting channels; • Optical media (CD or DVD); and • Electronic transmission, specifically through the Tax Office's Electronic Commerce Interface (ECI) application. – *Tax Office*
Confidence levels	A measure of match that indicates the quality of the matching outcome and of the likelihood that mismatch has occurred. – *Tax Office – Analysis of Identity Matching of AUSTRAC Data*

Data linking	The linking of data entities through common characteristics, e.g. linking a corporate group based on a common owner or linking transactions and events over time to a common intent. *- Draft Common Business Language term, Authorised by the CKO, September 2007*
Data mapping	Identifying the links and associations between entities *- Derived from training material in the Office of the Chief Knowledge Officer*
Data matching	The process of conducting checks of information by matching data held by the Tax Office with data obtained from other sources. *- Tax Office – Common Business Language* Data Matching is the acquisition of external data used by the Tax Office for a variety of different purposes such as detection of un-disclosed income, non-lodgement of returns and activity statements, and entities outside the tax system. *- Tax Office – External Data Matching PS CM 2004/17* The large scale comparison of records or files of personal information collected or held for different purposes, with a view to identifying matters of interest. *- Privacy Commissioner, The use of data matching in Commonwealth administration – Guidelines* Data-matching involves bringing together and comparing data (including personal information) from different sources. *- AGIMO – Future Challenges for E-government* At its most fundamental level, data matching is the linking of datasets and comparison of corresponding data. *- ANAO This Report*

Data mining

Data mining transforms raw data into higher-level constructs, such as predictive models, explanatory models, filters or summaries, by using algorithms from fields such as artificial intelligence and statistics. Techniques used can range from very simple models, such as arithmetic averages, those of intermediate complexity, for example, linear regression, clustering, decision trees, case-based reasoning and k-nearest neighbour, to very complicated models including neural networks and Bayesian networks.

– *Tax Office – Final Review of Analytics Project1.pdf*

Data modelling

Data models are used primarily to represent and document structured data, i.e. data that may be stored in a database. Data modelling uses 'Entity Relationship' modelling techniques as a widely used diagramming convention to document data requirements and relationships so as to meet business requirements.

Depending on the level of abstraction, data models are categorised as contextual, conceptual, logical or physical models.

– *Data Modelling Guide, Tax Office, Version 0.1, September 2005*

Data quality	Quality assesses whether information is 'fit for purpose'. Purpose is a statement of intent and should be thought of as inclusive of all intended business purposes throughout the Information Lifecycle. In all quality systems, purpose and requirements are determined largely by the customer/stakeholder. This view therefore needs to be built into any assessment of data quality. The aspects of quality of interest to the Tax Office are: • accuracy; • completeness; and • timeliness. Other quality aspects may be considered as part of particular project initiatives, including: • relevance; • accessibility; • interpretability; and • coherence. – *Data and Information Standards, Version 1.4, Tax Office, June 2006*
e-Tax	The Tax Office's free tax return preparation software. e-Tax lets individuals prepare and lodge online individual tax returns, baby bonus claims and 30 per cent child care tax rebate transfer advice.
Expert Rules Engines	Rules engines utilise expert business rules to filter and identify taxpayer entities which exhibit known characteristics of non-compliant behaviour. These rules are based on relationships, events and attributes that have been identified through various sources including statistical analysis, data mining and historical experience. – *Draft Common Business Language term, Authorised by the CKO, September 2007*

Identity matching	Identity matching attempts to attach ABN/TFN details by matching name, address and other information contained in the external data to Tax Office registration information. - *Tax Office – External Data Matching PS CM 2004/17*
Legislated Data	Datasets third parties provide to the Tax Office, as a result of specific legislated requirements (e.g. Annual Investment Income Reports provided by financial institutions). Legislated data sets cover major income streams such as salary and wages and interest and dividends, but do not cover other income streams such as capital gains, net rents, and income from self employment.
Negative search capability	The ability to match and report a 'No Match' as success (that is, the entity does not exist in the data being searched). - *Tax Office – Request for Tender Number 06.389 Identity Matching Solution*
Pre-filling	Provision of information the Tax Office currently uses for matching purposes, directly into an individual's electronic tax preparation or record keeping tool. It is intended to assist those doing the right thing to comply with their tax obligations. - *Tax Office – Pre-filling Intent Statement*
Provided Data	Datasets a Government Agency provides to the Tax Office, under a memorandum of understanding (e.g. State Revenue Offices).
Requisitioned Data	Datasets the Tax Office requisitions, on a non-routine basis, for data matching projects (e.g. data is requisitioned from sellers of luxury vehicles to address undeclared income and other tax risks). The Commissioner of Taxation's legislated access and information gathering powers provide the authority for this data requisition.
SENSIS	The most up-to-date electronic edition of the white and yellow paged phone books.

Index

Series Titles

Audit Report No.1 2007–08
Acquisition of the ABRAMS Main Battle Tank
Department of Defence
Defence Materiel Organisation

Audit Report No.2 2007–08
Electronic Travel Authority Follow-up Audit
Department of Immigration and Citizenship

Audit Report No.3 2007–08
Australian Technical Colleges Programme
Department of Education, Science and Training

Audit Report No.4 2007–08
Container Examination Facilities Follow-up
Australian Customs Service

Audit Report No.5 2007–08
National Cervical Screening Program Follow-up
Department of Health and Ageing

Audit Report No.6 2007–08
Australia's Preparedness for a Human Influenza Pandemic
Department of Health and Ageing
Department of Agriculture, Fisheries and Forestry

Audit Report No.7 2007–08
The Senate Order for Departmental and Agency Contracts (Calendar Year 2006 Compliance)

Audit Report No.8 2007–08
Proof of Identity for Accessing Centrelink Payments
Centrelink
Department of Human Services

Audit Report No.9 2007–08
Australian Apprenticeships
Department of Education, Science Training

Audit Report No.10 2007–08
Whole of Government Indigenous Service Delivery Arrangements

Audit Report No.11 2007–08
Management of the FFG Capability Upgrade
Department of Defence
Defence Materiel Organisation

Audit Report No.12 2007–08
Administration of High Risk Income Tax Refunds in the Individuals and Micro Enterprises Market Segments
Australian Taxation Office

Audit Report No.13 2007–08
The Australian Taxation Office's Approach to Managing Self Managed Superannuation Fund Compliance Risks
Australian Taxation Office

Audit Report No.14 2007–08
Performance Audit of the Regional Partnerships Programme:
Volume 1–Summary and Recommendations
Volume 2–Main Report
Volume 3–Project Case Studies
Department of Transport and Regional Services

Audit Report No.15 2007–08
Administration of Australian Business Number Registrations: Follow-up Audit
Australian Taxation Office

Audit Report No.16 2007–08
Data Integrity in the Child Support Agency
Child Support Agency
Department of Human Services

Audit Report No.17 2007–08
Management of the IT Refresh Programme
Centrelink

Audit Report No.18 2007-08
Audits of the Financial Statements of Australian Government Entities for the Period Ended 30 June 2007

Audit Report No.19 2007–08
Administration of the Automotive Competitiveness and Investment Scheme
Department of Innovation, Industry, Science and Research
Australian Customs Service

Audit Report No.20 2007–08
Accuracy of Medicare Claims Processing
Medicare Australia

Audit Report No.21 2007–08
Regional Delivery Model for the Natural Heritage Trust and the National Action Plan for Salinity and Water Quality
Department of the Environment, Water, Heritage and the Arts
Department of Agriculture, Fisheries and Forestry

Audit Report No.22 2007–08
Administration of Grants to the Australian Rail Track Corporation
Department of Infrastructure, Transport, Regional Development and Local Government

Audit Report No.23 2007–08
The Management of Cost Recovery by Selected Regulators

Audit Report No.24 2007–08
DIAC's Management of the Introduction of Biometric Technologies
Department of Immigration and Citizenship

Audit Report No.25 2007–08
Administering Round the Clock Medicare Grants
Department of Health and Ageing

Audit Report No.26 2007–08
Tasmanian Forest Industry Development and Assistance Programs
Department of Agriculture Fisheries and Forestry

Audit Report No.27 2007–08
Emergency Management Australia
Attorney-General's Department

Audit Report No.28 2007–08
Defence's Compliance with the Public Works Committee Approval Processes
Department of Defence

Audit Report No.29 2007–08
Parent School Partnerships Initiative
Department of Education, Employment and Workplace Relations

Current Better Practice Guides

The following Better Practice Guides are available on the Australian National Audit Office Website.

Public Sector Internal Audit	
An Investment in Assurance and Business Improvement	Sep 2007
Fairness and Transparency in Purchasing Decisions	
Probity in Australian Government Procurement	Aug 2007
Administering Regulation	Mar 2007
Developing and Managing Contracts	
Getting the Right Outcome, Paying the Right Price	Feb 2007
Implementation of Programme and Policy Initiatives:	
Making implementation matter	Oct 2006
Legal Services Arrangements in Australian Government Agencies	Aug 2006
Preparation of Financial Statements by Public Sector Entities	Apr 2006
Administration of Fringe Benefits Tax	Feb 2006
User–Friendly Forms Key Principles and Practices to Effectively Design and Communicate Australian Government Forms	Jan 2006
Public Sector Audit Committees	Feb 2005
Fraud Control in Australian Government Agencies	Aug 2004
Security and Control Update for SAP R/3	June 2004
Better Practice in Annual Performance Reporting	Apr 2004
Management of Scientific Research and Development Projects in Commonwealth Agencies	Dec 2003
Public Sector Governance	July 2003
Goods and Services Tax (GST) Administration	May 2003
Managing Parliamentary Workflow	Apr 2003
Building Capability—A framework for managing learning and development in the APS	Apr 2003
Internal Budgeting	Feb 2003
Administration of Grants	May 2002
Performance Information in Portfolio Budget Statements	May 2002

Some Better Practice Principles for Developing Policy Advice	Nov 2001
Rehabilitation: Managing Return to Work	June 2001
Business Continuity Management	Jan 2000
Building a Better Financial Management Framework	Nov 1999
Building Better Financial Management Support	Nov 1999
Commonwealth Agency Energy Management	June 1999
Security and Control for SAP R/3	Oct 1998
Controlling Performance and Outcomes	Dec 1997
Protective Security Principles (in Audit Report No.21 1997–98)	Dec 1997