



Office of State Revenue recovers an extra \$33 million for the people of NSW

The Office of State Revenue (OSR) is a department of the Treasury of the State of New South Wales (NSW), Australia. NSW has a population of over 6 million, with the majority of people and businesses located in Sydney, its capital city.

“Utilizing the SSA–NAME3 algorithms and performing a thorough analysis of the relationships between people, companies and addresses the software has enabled the OSR to successfully match over 50 million records to date, despite variances in naming conventions. This has also had the benefit of reducing manual processing.”

Tom Park,
Principal Data Analyst,
Strategy Branch.
OSR

The OSR administers State taxation, collects revenue, develops policy and implements legislation relating to State taxation for and on behalf of the people of NSW. OSR's mission is to collect all revenue due and administer revenue laws for the benefit of the people of NSW, while at the same time maximising effectiveness, efficiency and client satisfaction.

Government revenue departments such as the OSR must enforce tax compliance to maximise public revenue, thereby reducing the tax burden on truly compliant taxpayers.

An important component of compliance investigation is identifying high-probability and high-worth cases to investigate. Such identification is greatly assisted by the process of matching identity data between different tax systems and by matching tax data to external data sources (e.g. the “Yellow Pages”). Apart from the inevitable error and variation that exists in identity data, data sourced from different systems suffers from an additional component of variation introduced by different input methods and users, different controls, forms and database designs and other constraints or practices.

The Strategy Branch of the Compliance Division uses Identity Systems' Data Clustering Engine (DCE) v2.3 to improve the quality of the data it holds about individuals and organisations and assists with the investigation of that data for compliance and distribution of unclaimed monies.

DCE runs on a Windows NT platform, matching and grouping files sourced from within the OSR as well as externally. DCE is being used for a variety of projects including:

1. Unregistered Clients Project –

This project is designed to identify businesses that should be registered for Payroll Tax but who currently are not on record. While the primary objective is to identify new clients, other important goals include reducing the turnaround time for finding non-compliant businesses in addition to minimising manual checks. This is achieved by improving the match rate against data from external sources, such as ASIC's Business Register. Revenue generated by payroll tax is one of the major sources of income for the Office.

2. Unclaimed Money Project –

This project aims to identify a current mailing address for recipients of Unclaimed money. Unclaimed money is money that has been held in an account that has been inactive for at least six years. Unclaimed money held by OSR has been deposited by employers, government agencies, superannuation funds, local councils and State Trustees because they have not been able to locate and return the money to the owners. OSR aims to correctly identify and notify the recipients, again by matching to external data sources.

3. ABN /ACN /ANZSIC Code Update –

This exercise aims to verify OSR client records with external data and update the OSR database where records are incorrect or incomplete, with the aim of improving the completeness of the OSR database.

In each of these projects, the project goals have been achieved. For example, in the Unregistered Clients Project, over 800 new clients and approximately \$33 million in additional revenue has been identified. In addition, the match rate has doubled and manual checking has been significantly reduced. In the ABN /ACN /ANZSIC Code Update project, the completeness of the OSR database has also risen sharply.

“Hard match algorithms miss matches because of differences in layout, user errors in writing and transcribing information and incomplete data. A soft match algorithm, such as that used by the Data Clustering Engine compensates for these errors and differences in data” said Tom Park, Principal Data Analyst, Strategy Branch.

The Data Clustering Engine is an ideal solution for this type of business requirement because it is designed as a batch, project orientated, data grouping and investigation engine for all forms of identification data. Utilizing the SSA-NAME3 algorithms and performing a thorough analysis of the relationships between people, companies and addresses the software has enabled the OSR to successfully match over 50 million records to date, despite variances in naming conventions. This has also had the benefit of reducing manual processing.

“Identity Systems' software has significantly improved our capacity to get results through data matching”, said Ermil Sipp, Deputy Director Strategy Branch.

“Identity Systems has provided technical support by way of telephone and email support and on-site training. They have assisted us in upgrading to the latest release and in converting our v1.8 SSA-NAME3 rules into a v2.3 Custom Population. This has really simplified the maintenance of this part of the software.”

According to Tom Park, “the Data Clustering Engine plays a key role in our compliance toolkit. The technology has lived up to our expectations”..

For more information about the NSW Office of State Revenue, please visit <http://www.osr.nsw.gov.au>.

“Identity Systems' software has significantly improved our capacity to get results through data matching.”

Ermil Sipp,
Deputy Director
Strategy Branch
OSR

About Identity Systems

Identity Systems is the pioneer in enabling organizations to build and maintain high-quality identity data search and matching software solutions.

The company has been in this area of specialization since 1986, and has over 500 clients worldwide who rely on its robust, enterprise-wide software. For more information please visit our website.

