KPMG

cutting through complexity

Management initiated review

Privacy breach – Data management

Department of Immigration and Border Protection

5 April 2014

Contents

The contacts at KPMG in connection with this document are:

s 22(1)(a)(ii)

Contents	**2**
Warranties and disclaimer	**3**
Executive summary	**4**
Introduction	**6**
Background	6
Scope and objectives	6
Approach	6
Overview of work performed to date	**7**
Observations	**8**
Chronology of the creation, review and publishing of the document	8
Departmental policies and guidance regarding online publishing	9
Forensic examination of the data publication	13
Policies and management practices that contributed to the data publication	15
Recommendations	**17**
Appendices	**18**
Appendix A	**19**
Appendix B	**20**
Contact us	**21**

KPMG
cutting through complexity

Warranties and disclaimer

This report has been prepared at the request of the Department of Immigration and Border Protection (the DIBP) for the purpose of assisting the DIBP identify how access to personal information was gained by unauthorized person/s, in connection with a document uploaded to its website on 10 February 2014, and providing recommendations to assist with managing the risk of a recurrence. The report covers fieldwork performed up to and including 13 March 2014.

We have prepared this report pursuant to the terms of reference set out in our Management Initiated Review, dated 24 February 2014, and the terms and conditions of the Deed of Standing Offer (DOSO) with KPMG, under which the DIBP has engaged KPMG to provide audit and Forensic services, and they are not to be used for any other purpose without our prior written consent. Accordingly, KPMG accepts no responsibility in any way whatsoever for the use of these documents for any purpose other than that for which they had been prepared.

The services provided in connection with this engagement comprise an advisory engagement, which is not subject to assurance or other standards issued by the Australian Auditing and Assurance Standards Board and, consequently no opinions or conclusions intended to convey assurance have been expressed.

This report must not be shown, copied, provided, disseminated, given to or relied on by any other person or entity without our express written consent which may be withheld in our absolute discretion.

We have considered and relied upon information, which we believe to be reliable, complete and not misleading. Nothing in these documents should be taken to imply that we have verified any information supplied to us, or have in any way carried out an audit of any information supplied to us other than as expressly stated in this report. The statements and observations included in these documents are given in good faith, and in the belief that such statements and findings are not false or misleading.

These observations are based solely on the information provided to us during the course of our fieldwork up to 13 March 2014. We reserve the right to amend any findings, if necessary, should any further information become available.

KPMG
cutting through complexity

Executive summary

Background

The 'Immigration Detention and Community Statistics Summary' (the Document), a Microsoft Word document dated 31 January 2014, which was published on the DIBP's website, allowed access to source data containing personal information of approximately 10,000 detainees. The personal information was contained in a Microsoft Excel Workbook, used to perform the underlying analysis, which was embedded into the Document. KPMG was subsequently appointed to undertake a Management Initiated Review (MIR) into the matter. The scope of the MIR is set out in the report.

The review has been undertaken through two work streams. One focussed on establishing the chronology of events leading to the publication and the second a technical examination of the data associated with the publication and the potential extent of access to that data.

Summary of work performed to date

- Planning discussions with limited Senior Executives ("SES").
- Review of policies and procedures relating to the creation, review or upload of the document subject of the review, and discussions with personnel involved in that process.
- Analysis of the final version of earlier publications of the Immigration Detention and Community Statistics Summary, published online in the twelve months prior to 31 January 2014, to determine whether any other versions exhibit vulnerabilities that may allow access to sensitive or personal information.
- Analysis of Internet Protocol (IP) addresses, external to the DIBP, identified as attempting to access the file.
- Identification of country of location for IP addresses identified as having retrieved either the entire file or enough of it to potentially contain a portion of the embedded spreadsheet data, containing the personal details of the detainees.
- Engagement with a web archiving service to remove the publication from its website

Observations

How did the data become embedded in the 'Immigration Detention and Community Statistics Summary' dated 31 January 2014?

- The data extract obtained to produce the analysis for the Document contained the personal information of the detainees. Personal information was not removed prior to the analysis being performed. When the statistical information was transferred into the Document, it was done in such a way that underlying data containing personal information of detainees was embedded, which meant the information could be accessed despite there being no visible indicators of this. Reviews of the Document were primarily done on hard copy versions, rather than in electronic form, which is unlikely to have detected the embedded data.
- The process adopted in producing and publishing the document appears to have not conformed with the roles and responsibilities set out in either the web publishing and governance intranet guidance or the online style guide. Although potentially ambiguous in its reference to meta data, the online style guide sets out specific requirements for ensuring publications do not contain underlying data sets or personal information. The potential reasons contributing to the incident may include time pressures, technical ability, lack of awareness of roles and responsibilities and lack of awareness of IT security risks associated with online publishing.

Is there any indication that the privacy breach may have been malicious or intentional?

- From our work to date, we have not identified any indications that the publication of the underlying data was intentional or malicious.

Do earlier versions of the same report contain the same exposure to the underlying data?

- Our review of earlier versions, provided to us, did not identify the same issue, so it appears unique to the version dated 31 January 2014.

Is the Department likely to be exposed to the risk of inadvertent publication of data through its other on line publications?

- Although earlier versions of this document that were provided to us did not contain the same link to the underlying data, the lack of awareness of IT security issues associated with online publishing, and the process of reviewing and clearing publications in hard copy, creates a significant risk that other publications may not have been properly reviewed and therefore, would be at risk of having

KPMG
cutting through complexity

underlying data issues. The same would apply to any electronically transmitted information, not just on-line publishing.

What was the extent of potential access to the 'Immigration Detention and Community Statistics Summary', dated 31 January 2014 and therefore, the underlying personal data?

- The potential data access and distribution is widespread. There were 123 "hits" on the document from 104 unique IP addresses. Of those, 26 contacts downloaded the full document and therefore must be assumed to have obtained the underlying data. 75 contacts related to the document being browsed over the internet through a computer based browser, capable of downloading the full document and therefore the underlying data, though we could not confirm that a download had in fact occurred. The remaining contacts are considered a low risk of obtaining the document and the data, due to the nature of the technology used to look at the document.
- The predominant access country is Australia, though many other views or potential access originated from a range of other countries.

Can the personal data be tracked and recovered from those who potentially accessed it?

- No, with very limited exceptions. Some of the contacts have occurred through arrangements which anonymise the source of the contact. Our experience with Internet Service Providers is that they will not disclose the details of their customers, unless subject to a Court order or warrant. As such, the DIBP would be unable to identify who, in fact, accessed the document. The nature of some of the potential countries of origin for the contacts would make cooperation difficult. Once the data is downloaded, it could then be emailed to anyone or posted anywhere.
- The exceptions are cases where we can identify the downloader, e.g. Archive.org, in which case we have contacted them and arranged to have the data removed. Likewise, with organisations such as Guardian Australia, who we understand cooperated with destroying the data.

What steps can the Department take to manage the distribution of the data further?

- The DIBP could establish an internet search facility to monitor for publication of the document and then approach any individual or organisation identified as publishing the information, to have it removed and destroyed. The DIBP could also consider engaging a "scraping" service to monitor social media for any references to the information.

Recommendations

- Develop and implement a procedure whereby any data to be extracted for the purpose of analysis is normalised and cleansed in a secure environment, to ensure that any personal or sensitive data is removed prior to any analysis being performed;
- Update online publishing quality assurance checklists to require approvers to confirm that the document has been reviewed in its native electronic form;
- Hold online publishing workshops involving Director level representation from Information Technology (IT) Security, Web Operations and Governance and all Branches involved in the creation of material that may be published online. The objectives being to:
  - clearly agree and delineate roles and responsibilities;
  - define an appropriate point in the online publishing process to undertake an information technology security quality assurance review and assign accountability to an appropriately qualified and experienced role;
  - review the current online publishing clearance processes for both content approval and publishing to streamline the process for the purpose of creating efficiencies; and
  - consider the current prioritisation matrix for online publishing and update as appropriate to ensure that applicable criteria is in line with current business practices.
- Develop an IT security training program, to be delivered to all personnel operating in an area of the DIBP responsible for handling private or sensitive data, and include specific day-to-day scenarios covering typical risks associated with handling such data;
- Incorporate lessons learned from this review into Privacy training to be delivered in connection with the imminent operation of the Australian Privacy Principles; and
- Ensure that all policies, procedures and other guidance materials relating to roles and responsibilities of personnel involved in the creation, review and publishing of online content is updated on a timely basis and accessible to all areas of the DIBP.

KPMG
cutting through complexity

Introduction

- Background
- Scope and objectives
- Limitations
- Approach

Background

The Department of Immigration and Border Protection (DIBP) has inadvertently allowed access through its website to the personal information (including names, dates of birth, nationality) of a large number of detainees. The access was gained by a person/s unknown and passed to journalist/s at The Guardian.

The DIBP loaded a Microsoft Word document titled 'Immigration Detention and Community Statistics Summary' dated 31 January 2014, onto its website on 10 February 2014. On page 10, Figure 5 of this document, a graph inadvertently allowed access to source data that contained personal information as outlined above. The DIBP removed the document from its website immediately after the matter came to its attention. The DIBP is uncertain as to the extent of who downloaded the source data and consequently continues to have access to the data.

The Privacy Act 1988 requires the DIBP to adhere to a number of Information Privacy Principles (IPPs) (soon to be replaced by the Australian Privacy Principles) governing the collection, use and disclosure of personal information. "Personal information" is relevantly defined to mean information about an individual whose identity is apparent, or can reasonably be ascertained, from the information.

Scope and objectives

The objectives of this Management Initiated Review (MIR), of which a copy is attached at Appendix A, were to identify how access to personal information was allowed by unauthorized person/s and any recommendations to prevent this occurring again.

This MIR included the following scope of work:

1) The chronology of how access to personal information was gained.
2) Who accessed the information and who may retain access to the personal information?
3) Report on remaining departmental vulnerabilities that may allow access to sensitive or personal information.
4) Policy or management practices that contributed to this significant breach of privacy.
5) Options to prevent recurrence.

Limitations

Our scope of work was undertaken to the extent practicable within the timeframe of this review and subject to the following limitations:

- In respect of item 2, details recorded in applicable system logs and the extent to which that enabled distinction between access, as opposed to retention of the relevant data;
- In respect of item 3, vulnerabilities relating to departmentally approved and initiated release of information into the public domain; and
- We did not inquire into whether a breach of the Privacy Act in fact occurred, nor did our review extend to a detailed analysis of root causes or potential disciplinary issues.

Approach

The approach taken to this MIR was as follows:

- Reviewing policies and procedures relevant to understanding the DIBP's management of personal information and protocols applied when loading material on the internet;
- Meetings with appropriately informed persons to collect information about policy, protocols and practices within the scope of this investigation;
- Providing daily a short oral briefings to Michael Manthorpe, Deputy Secretary, Portfolio Innovation and Coordination Group or, in his absence, Stephen Wood, Assistant Secretary Professional Integrity and Assurance; and
- Preparation of draft and final reports.

Overview of work performed to date

- Review of relevant policies and procedures
- Discussions with personnel involved in document's creation, review and publishing of the document
- Forensic examination

An overview of the work completed to date as part of the review is set out below:

- Planning discussions with the following personnel:
  - Deputy Secretary – Portfolio Coordination & Innovation
  - First Assistant Secretary – Detention Infrastructure and Services
  - First Assistant Secretary – Strategic Business Solutions Division
- Fieldwork discussions with personnel in the following roles involved in the creation, review or upload of the document subject of the review, or otherwise in a First Assistant Secretary role in the relevant areas of the DIBP that prepared and published the document:
  - Chief Information Officer
  - First Assistant Secretary – Detention Infrastructure and Services
  - First Assistant Secretary – Strategic Business Solutions Division
  - Assistant Secretary – Detention Services Management
  - Assistant Secretary – Community Programmes and Children Division
  - Director – ISRG reporting
  - Director – Web Operations and Governance
  - Assistant Director – ISRG Public information team
  - EL1 – Manager ISRG reporting
  - EL1 – Manager ISRG reporting
  - EL1 – Manager Data mining and analysis
  - APS6 – Business Analyst
  - APS 4 – Business Analyst
  - Area Web Coordinator – ISRG Area Web Coordinator
  - Web Publisher– Web Operations and Governance
- Forensic examination of data associated with the incident, including the following tasks:
  - Analysis of the final version of files, published online in the twelve months prior to 31 January 2014, to determine whether any other versions exhibit vulnerabilities that may allow access to sensitive or personal information, in the form of either:
    - Embedded personal information of detainees; or
    - Links (broken or otherwise) to personal information of detainees.
  - Analysis of Internet Protocol (IP) addresses, external to the DIBP, identified as attempting to access the Document, including consideration of:
    - whether the file was entirely or partially received; and
    - the identity of the IP addresses which have accessed the file, including in particular web crawlers associated with search engines, internet archiving sites, media organisations, certain networks which allows users to anonymously access the internet without releasing information relating to their location and State and Federal Government Departments.
  - Identification of country of location for IP addresses identified as having viewed the document, or retrieved either the entire file or enough of it to potentially contain a portion of the embedded spreadsheet data.

Observations

- Timeline of events
- Procedural guidance
- Forensic examination of the data publication
- Policies and management practices that contributed to the data publication

Chronology of the creation, review and publishing of the document

Set out below is a high level chronology summarising the events that transpired in connection with the publication of the document subject of this review. This is not intended to provide a full and detailed account of events that transpired, but rather provide an overview of the general process adopted in publishing the document and the key personnel involved, so as to enable the DIBP to form an assessment as to whether there are any factors present to suggest the incident may be malicious or intentional in nature, as well as to gain some insight into what contributed to the incident occurring.

Key steps in the process of publishing the Document are summarised as follows:

- A data set, including personal information of detainees was extracted from a data warehouse by a Manager within the ISRG Reporting team, for the purpose of preparing the Monthly Detention and Community Statistics Summary (the publication), to be published on the DIBP's website. In this instance, the data set was extracted manually using data analytics queries over the weekend of 1 /2 February 2014. Ordinarily this would be automated, however, it was manually expedited to assist in meeting the target publication date of 10 February 2014;
- A Business Analyst within the ISRG Reporting team, imported the data set into a MS Excel template used to prepare the analysis for inclusion in the statistical report, and also undertook quality assurance of and updated formulae in the template;
- Another Business Analyst within the ISRG Reporting team, performed quality assurance on the data set contained in the MS Excel template, for the purpose of rectifying data integrity issues. This involves clarification of potential anomalies such as blank fields and age/classification discrepancies et al. The inclusion of names within the data set facilitates this process, which typically takes in the order of a day.
- In accordance with standard procedures, a Business Analyst copied data in the form of charts and tables from the MS Excel template into the MS Word version of the publication. We understand the preceding quality assurance process and this preparation of the publication, was delayed to some extent in connection with the reconciliation of differences arising from the ISRG reporting team taking on responsibility for BVE statistics. Subsequent to this, the content clearance process commenced;
- Clearance for the content of the publication was escalated through the following persons, for review and approval[1], in accordance with the review process for Category Four documents, as set out in the ISRG reporting team's content clearance matrix:
  - EL1
  - EL2
  - Assistant Secretary
  - First Assistant Secretary
  - Deputy Secretary
- Throughout the content clearance process, various aspects of the publication were amended, based on reviewer comments. In some instances, this involved alterations to the underlying MS Excel template and copying of data to the MS Word publication.
- A Business Analyst submitted the publication to the ISRG Website team's functional email box for publishing to the internet. The publication was receipted by the Area Web Coordinator for ISRG.
- The publication was then escalated through the clearance process of the ISRG website team. This involved review and approval from the following personnel:
  - Area web coordinator
  - Assistant Director
  - Assistant Secretary, in place of the Director

[1] In some instances we sighted documentary evidence of completion of the clearance process, in other instances only verbal clearance was provided. Documentation has not been provided to evidence whether First Assistant Secretary level clearance was obtained.

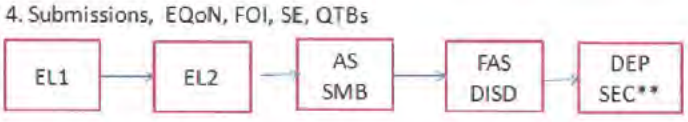
- Similar to the content clearance process, various aspects of the publication were amended during the review by the ISRG website team. The clearance process included consideration of Ministerial distribution standards, for which members of the ISRG website team also have responsibility.
- Amongst the various iterations of the publication, the chart on Page 10 of the publication was pasted into the MS Word version of the publication as a MS Office Excel chart object, as opposed to a picture. This resulted in the MS Excel data underpinning the chart being embedded into the publication. The Business Analyst involved could not precisely recall the specific option she selected on the computer to result in the MS Excel data becoming embedded.
- Communication, noting Director level approval from the ISRG Website team, was submitted to the Web operations, Publishing and Governance group mail box, where it entered a queuing system to be actioned.
- The publication was randomly allocated to a Web publisher by an EL1, in accordance with standard procedures.
- The Web publisher renamed, optimised and checked the properties of the MS Word and pdf versions of the publication in accordance with accessibility requirements of the Web Content Accessibility Guidelines 2.0 (WCAG 2.0).
- The publication was uploaded to a test environment by the Web Publisher and screenshot and document links emailed to the ISRG website team for review and approval.
- Approval was provided by the Area Web Coordinator before the publication was migrated to the live environment and uploaded to the internet. We understand that when the initial version of the publication was uploaded, an ISRG Reporting team member encountered difficulties with accessing hyperlinks, which were rectified before the publication was accessible online.

Quality assurance reviews were performed throughout the process of publishing the document however, in general, these were performed on a hardcopy of the publication and focused on ensuring the writing style, grammar and spelling were consistent with Departmental guidelines and that calculations were arithmetically correct. When an electronic version was reviewed, quality assurance checks focused on ensuring compliance with accessibility guidelines. Authors and approvers were generally unaware that embedding of MS Excel data within a MS Word document could occur and were therefore not mindful of checking for indicators of this risk.

Departmental policies and guidance regarding online publishing

ISRG reporting team clearance matrix

The ISRG reporting team applies a 'ready reckoner' clearance matrix, an internal working document used by that team only, in determining the requisite level of review for approving content with respect to the group's role in preparing non-financial reporting of detention services and related activities. The clearance process adopted in respect of the document subject of this review is summarised as follows:



We note the following in respect of the content clearance process:

- There is no underlying detailed guidance setting out specific quality assurance checks to be undertaken at each escalating level; and
- The matrix indicates that the ISRG reporting team will "walk documents around for clearance if extremely urgent".

Online style guide and web governance intranet guidance

Set out below is an overview of roles and responsibilities with respect to web publishing and governance and compliance with Privacy legislation, as set out in the DIBP's Online Content Governance procedures and as displayed on the DIBP's intranet.

Role	Online style guide compliance	Privacy principles
Division Head	**Expectected to....monitor** quality, relevancy, accuracy and **legality** of web content under their responsibility	
Branch Head	**Will....guarantee** their web content **meets the governance requirements** set out in this document	
Director *(responsible for viewing information prepared by authors)*	**Ensure** new web content and major changes...are usability and accessibility tested to **meet online publishing guidelines**	Nil
Area web coordinator	**Ensure** content for their division **meets the department's publishing standards**	Nil
Author *(responsible for preparing information for websites)*	**Ensure** that information submitted for approval: • ....follows the departmental style guide • **meets** Australian Government Online and **Departmental online publishing requirements and standards**; and • **does not infringe** copyright and **privacy legislation.**	

We note the following with respect to the above guidance material:

- Neither the content authors, nor the Director of the ISRG reporting team were aware of these responsibilities, nor had they received any training in their application;
- Whilst available on the intranet, it was not readily accessible for users to locate; and
- The online style guide, discussed in further detail below, is not something that either the content authors, or the Director of the ISRG reporting team, were aware of.

We have reviewed the DIBP's Online Style Guide and note the following relevant guidance which may have assisted those responsible for the creation, approval and publishing of the Document, to detect the embedded data.

Section	Guidance
Appendix D – Uploading documents	Use the tools in Microsoft Office applications or Adobe Acrobat to delete hidden metadata permanently before putting non-HTML documents (such as Word, PDF, Excel, PowerPoint) online.
	Remove hidden data and personal information *(underlined)*.
	Check that you have removed portions of embedded objects that are not visible in your document.
	Additional information provided in the manual as to why the above guidance is important: • Where a non-HTML document is created, edited or saved using Microsoft Office applications, hidden data (also known as metadata) is automatically added to the document: – When this documented is distributed electronically hidden data remains present, but often not visible, until it is deleted. Some data is easily seen and some can be viewed only by opening the document in a specialised program. – This document metadata can put agencies at risk if viewed or extracted, by exposing sensitive information that is not intended for public distribution.

We note the following in respect of the specific online style guide information set out above:

- The guidance is explained, in part, with reference to the concept of metadata which, depending on the technical proficiency of the reader, may be interpreted differently and the user may not necessarily recognise the applicability of the guidance to the other private data hidden in the document subject of this review. Nevertheless, the specific guidance notes as set out above, when read individually, directly refer to the presence of hidden data and personal information, which was in fact contained in the document subject of this review and it is reasonable to expect that the guidelines could have been interpreted sufficiently broadly as to be applied in this particular instance;
- In the absence of accompanying practical training demonstrating the range of ways in which these risks can occur, it is unlikely that less technically proficient readers would understand the risks and how to determine whether any underlying data was present; and
- Whilst the relevant guidance is set out in an Appendix relating to the uploading of documents, the associated risks with respect to hidden data and objects is also applicable to any circumstances under which an electronic document is transmitted.

Procedures relating specifically to the creation, review and publishing of online content

Each team involved in the creation, review and publishing of the document subject of this review, has procedural guidance available to team members. We set out below, our observations with respect to each team's respective guidance:

ISRG reporting team step by step guide to preparing the Monthly Detention and Community Statistics Summary

- When copying the statistical analysis from the MS Excel template to MS Word, the user is instructed to "Copy and paste each graph as a picture into the word document (each graph must be pasted as a picture/image). We note the following with respect to this instruction:
  - Adherence to this instruction would have prevented the detainee data being embedded in the MS Word document;
  - Despite being aware of this explicit instruction, no members of the ISRG reporting team involved in creating or reviewing the document understood the IT security context or the risks associated with failing to adhere to the instruction, for example that data could be embedded via a graph pasted into a MS Word document,; and
  - Without proper training, the instruction in itself is insufficient in its application as the outcome would visually appear to be the same, that is a picture / image appears, irrespective of the method used to copy the data from MS Excel to MS Word.
- Whilst the document sets out the clearance process escalation chain, it does not provide any guidance in respect of specific quality assurance checks to be performed at each level.

ISRG website team guidance

- The following specific procedural guidance was available to the ISRG website team, specifically the Area web coordinator and associated reviewers, at the time of the incident:
  - Making changes to the DIBP website Standard Operating Procedures (SOP)
  - Procedures for web content representatives
- Guidance regarding quality assurance checks to be undertaken makes no reference to the risk of embedded data or any associated measures designed to detect such data, with checks being limited to arithmetical accuracy, grammar and spelling, formatting errors and ensuring that pictures and tables contain "alt text" in line with accessibility guidelines. We note that checking each image for "alt text" could possibly have resulted in the detection of the embedded data.

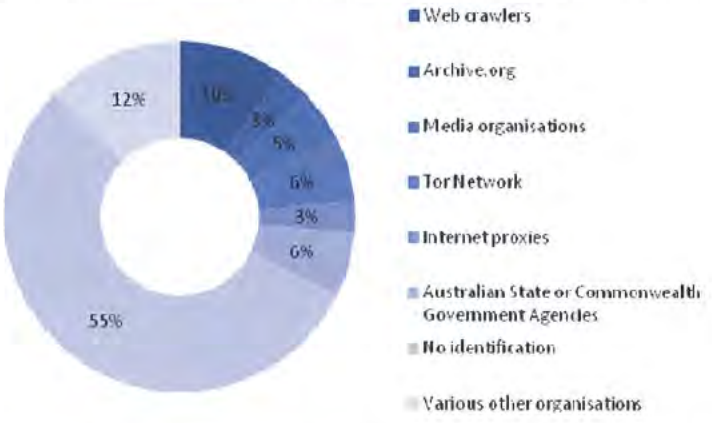
Web operations and governance guidance

- The Web operations, publishing and governance section maintains procedural documentation relating to internal processes around online publishing. This includes, inter alia, RACI matrices, process flows and both pre-publishing and quality assurance checklists. We note the following with respect to this documentation:
  - The document is stated as having been developed for the Web operations and governance team, and it is unclear whether the document or its associated learnings have been made available to other relevant personnel, such as content authors;
  - The pre-publishing checklists includes checks as to whether the content adheres to the online style guide and also specifically whether images included in the content are in appropriate web formats. It is possible that adherence to these checks may have resulted in the embedded data being detected in the document;
  - The QA checklist also contains references to accessibility and metadata checks of images which may, indirectly, have resulted in the embedded data being detected in the document; and
  - The checklists, collectively, reflect a focus more on compliance with writing style and accessibility with no explicit reference to consideration of IT security risks that may be associated with the publishing of content online.

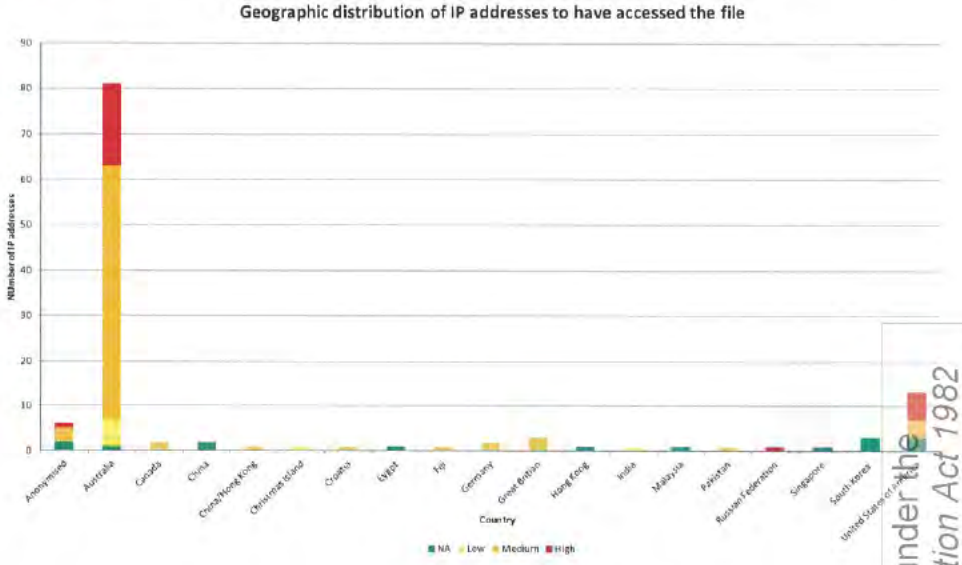
Forensic examination of the data publication

Our observations with respect to the forensic examination of the data publication are summarised as follows:

- 123 accesses via 104 unique internet protocol (IP) addresses attempted to retrieve the file at least once. Of these 104 IP addresses:
  - 88 unique IP addresses received the entire file at least once;
  - 1 unique IP address received a partial copy of the file which would have included part of the excel data[2]; and
  - 15 unique IP addresses attempted to access the file but either did not receive it or only received the content without the embedded spreadsheet.
- Summarised below, to the extent possible, are the types of organisation identified as being associated with the 104 Unique IP addresses that attempted to retrieve the file at least once.



- We attach at Appendix B, a matrix summarising our analysis of "hits" on the file and a risk assessment of the likelihood of the underlying data being obtained[3]. A chart of the geographic location of IP addresses identified as having accessed the file is depicted below:



We note the following with respect to the results set out in the chart above:

[2] For a typical user, it is unlikely this would have enabled the file to open and reveal data.

[3] Note that the file has 123 unique rows relating to the 104 IP addressed identified, as more than one hit can come from the same IP address, for a number of reasons.

- The predominant access country is Australia, though many other views or potential access originated from a range of other countries.
- 26 contacts downloaded the full document and therefore must be assumed to have obtained the underlying data. These are rated high risk, depicted in red in the chart above and Appendix B, as the document has clearly been accessed (as opposed to just browsed), archived or other indications exist that a copy of the full file has been saved in some form.
- 75 contacts related to the document being browsed over the internet through a computer based browser, capable of downloading the full document. We have rated these medium risk, depicted in orange in the chart above and Appendix B, as they potentially have a copy, or have saved a copy, though we could not confirm that a download had in fact occurred.
- The remaining contacts are considered low risk of obtaining the document and the data, due to the nature of the technology used to look at the document. These are depicted in green in the chart above and Appendix B.

With respect to the hits identified in Appendix B as relating to the IP addresses associated with "Archive.org", we have been in communications with that entity, on behalf of the DIBP, and been successful in arranging for the information obtained by Archive.org to be removed from publication.

Our analysis also identified that the file was downloaded 24 times from the DIBP's internal network, however, web proxy logs provide no further detail in this regard.

In relation to tracking and recovery of the data, our view is that, with very limited exceptions, the DIBP will not be able to track or recover the data. Some of the contacts have occurred through arrangements which anonymise the source of the contact. Our experience with Internet Service Providers is that they will not disclose the details of their customers, unless subject to a Court order or warrant. As such, the DIBP would be unable to identify who, in fact, accessed the document. The nature of some of the potential countries of origin for the contacts would make cooperation difficult. Once the data is downloaded, it could then be emailed to anyone or posted anywhere.

The exceptions are cases where we can identify the downloader, e.g. Archive.org, in which case we have contacted them and arranged to have the data removed. Likewise, with organisations such as Guardian Australia, who we understand cooperated with destroying the information.

Although our analysis of earlier versions of this document that, were provided to us, did not contain the same link to the underlying data, the apparent limited knowledge of the process as set out in the web publishing and governance intranet guidance, on-line publication style guide and the process of reviewing and clearing publications in hard copy, creates a significant risk that other publications may not have been properly reviewed and published and therefore, would be at risk of having underlying data issues. The same would apply to any electronically transmitted information, not just on-line publishing.

Policies and management practices that contributed to the data publication

- The data set extracted from the DIBP's Compliance, Case Management, Detention and Settlement Portal (CCMDS) and Integrated Client Services Environment (ICSE) systems and various operational spreadsheets used to do the analysis required for the publication, included fields containing the personal information of detainees. It appears this portion of the data set, whilst useful in checking and preparing the data set for analysis, was not necessary to undertake the analysis required.
- Previous iterations of the publication were published online in pdf form only. Legislative requirements on Government to publish information in accessible form for the visually impaired led to a decision to publish the document in MS Word form also. It is unclear whether the DIBP IT security team was adequately consulted in respect of this decision to understand what additional risks this presented and how to best manage those risks.
- The creators, majority of reviewers and publisher of the document were unaware that it was possible to embed MS Excel data in a MS Word document. As a result, nobody who viewed an electronic version of the document checked to see whether there was embedded data.
- The majority of clearance reviews were performed on a hardcopy of the document. This was unlikely to result in the embedded data being detected.
- Although the DIBP's online style guide provides some guidance addressing the risk that data containing sensitive information can be embedded in documents, no procedural documentation adequately explains the privacy / security risks associated with copying and pasting charts and tables from Excel into an MS Word document. We have sighted email correspondence demonstrating confusion in the past amongst ISRG Reporting and Website teams regarding how to copy and paste charts and tables from MS Excel into a MS Word document in a way that achieves compliance with accessibility requirements..
- Though procedural guidance for creating the publication explicitly instructed the user to copy each chart into the publication as a picture, there was no step by step instruction for specific options to select in doing so. Further, staff involved in this process had no understanding of the context for why this needed to occur. The instruction was, in itself, insufficient in its application as the outcome would visually appear to be the same, that is a picture / image appears, irrespective of the method used to copy the data from MS Excel to MS Word.
- There is a general lack of understanding, amongst the various teams involved in online publishing, as to the clearance checks undertaken by each group and / or reviewer. In several instances, it was assumed that prior reviewers had undertaken more extensive checks than had actually occurred.
- There was a view expressed that the publication was handled by too many reviewers and that each additional layer of the quality assurance review process did not always add value.
- Both the ISRG reporting and website teams face significant time pressure in publishing the document. In respect of this publication, there are several contributing factors, including:
  - Whether the publication's assigned priority status, being urgent, is reasonable in the context of the DIBP's prioritisation matrix for online publishing, which is defined based on an impact and urgency assessment against specified criteria, and whether the matrix or communication protocols could be updated to better reflect the business context;
  - The amount of data cleansing and normalisation required to get the data set in the appropriate form to complete the analysis for the publication; and
  - The number of reviews required as part of the clearance process and the time required coordinating this.

- It may be a timing issue in this instance, but the majority of personnel involved in handling the publication had not recently undertaken any Privacy training. People acknowledged receiving this training on induction, however, had not participated in refresher training.
- The ISRG reporting team responsible for preparing the publication does not use MS Word extensively and do not have advanced knowledge of its functionality and operation.
- The primary author who prepared this edition of the monthly publication had not prepared this publication previously. Though this delegated responsibility was planned and supported by supervision and training, the author was unfamiliar with the MS Excel template and publication and the process was therefore more likely to be susceptible to human error.

Recommendations

- Handling of sensitive data
- QA processes
- Consultation with IT security
- Training
- Australian privacy principles
- Policies and procedures

Based on observations from our work completed to date, as well as consideration of suggestions from personnel consulted as part of the review, we recommend consideration of the following measures designed to prevent recurrence of this specific incident, or an incident of a similar nature, that may occur as a result of the vulnerabilities identified through this review:

- Consider the development and implementation of a procedure whereby any personal data extracted for the purpose of analysis is normalised and cleansed in a secure environment, to ensure that any private or sensitive data, not necessary for the analysis, is removed prior to any analysis being performed[4];
- Update online publishing quality assurance checklists to require approvers to confirm that the document has been reviewed in its native electronic form. Reviewers may, at their discretion also choose to review a copy of the publication in hardcopy, however, the electronic review should constitute the base line check;
- Hold online publishing workshops involving Director level representation from Information Technology (IT) Security, Web Operations and Governance, User Centred Design Competency Section and all Branches involved in the creation of material that may be published online. The objectives being to:
  - Identify and discuss the risks associated with handling sensitive data and publishing content online;
  - agree and delineate roles and responsibilities, including consideration of whether the area web coordinator facilitates consistency in a devolved online publishing model or whether removal of this role would streamline the process across the DIBP;
  - define an appropriate point in the online publishing process to undertake a quality assurance review, focused on information technology security, and assign accountability to an appropriately qualified and experienced position;
  - review the current online publishing clearance processes for both content approval and publishing to streamline the process for the purpose of creating efficiencies; and
  - consider the current prioritisation matrix for online publishing and update as appropriate to ensure that applicable criteria is in line with current business practices.
- Develop an IT security training program, to be delivered to all personnel operating in an area of the DIBP responsible for handling private or sensitive data, and include specific day-to-day scenarios covering typical risks associated with handling such data;
- Consider liaising with appropriate Commonwealth bodies regarding organisational readiness with respect to accessibility guidelines. In particular confirming the current status regarding acceptable publication formats and whether any recent technological advances in that regard alter the DIBP's current position with respect to online publishing preferences[5].
- Incorporate lessons learned from this review into Privacy training to be delivered in connection with the new Australian Privacy Principles; and
- Ensure that all policies, procedures and other guidance materials relating to roles and responsibilities of personnel involved in the creation, review and publishing of online content is updated on a timely basis and accessible to all areas of the DIBP.

In addition to the specific measures set out above, we recommend that the DIBP, in responding to this incident, also take the opportunity to consider its current practices and procedures with respect to handling of sensitive data and, in particular, the level of consultation with the DIBP's IT security team in managing this risk. Increasing demand for information from various stakeholders and interest groups and the proliferation of social media and emerging technologies presents an ever changing risk profile going forward and DIBP's IT security are well placed to assist the DIBP in managing this risk.

[4] In considering this, it is important to note the DIBP's primary function has an intrinsic human element and whether depersonalising data in all instances in any way diminishes that focus.
[5] It is important to emphasise that modifications to the online publishing process do not become the sole focus of remedial efforts going forward. This incident is symptomatic of a need to consider the handling and transmission of sensitive data more broadly.

Appendices

Appendix A

Australian Government
Department of Immigration and Border Protection

Form of Order for
Management Initiated Review

This order is placed pursuant to and subject to the terms and conditions of the Deed of Standing Offer between Department of Immigration and Border Protection (the department) and KPMG effective from 11 July 2013, for the provision of internal audit and related professional services.

Audit reference	MIR 11 (2013-14)
Title	Privacy Breach – Data Management
High level scope	Attachment A
Timeframe	3 March 2014

Service provider contact details

Service Provider **Vendor No** **ABN**	**KPMG** 105145 51 194 660 183
Contract Manager **Phone** **Mobile** **Email**	s 22(1)(a)(ii) [redacted]

Department contact details

Manager **Phone** **Email**	Michael Manthorpe PSM s 22(1)(a)(ii) [redacted]
Alternative contact **Phone**	
Contract Manager	s 22(1)(a)(ii) [redacted]

Fees

Indicative cost: $ 34,605 GST exclusive

Total fees are inclusive of:

- planning
- entry interviews
- fieldwork
- progress meetings, if required
- exit interviews
- developing and finalising deliverables
- travel and accommodation

Invoicing

Invoicing arrangements are to be agreed with the business sponsor(s) prior to the commencement of the assignment.

Deliverables

Weekly status updates as per template specified by the department.

Format of deliverables are to be agreed with the business sponsor(s) prior to the commencement of the assignment.

SIGNED on behalf of the
COMMONWEALTH OF AUSTRALIA AS REPRESENTED BY THE DEPARTMENT OF IMMIGRATION AND BORDER PROTECTION

BY

Signature: s 22(1)(a)(ii) [redacted]

Name: Michael Manthorpe.

Position: Deputy Secretary.

Date: 24.2.14

SIGNED on behalf of KPMG

BY

Signature: s 22(1)(a)(ii) [redacted]

Name:

Position: Partner.

Date: 24 Feb 2014

Appendix B

Summary of access to the file and risk assessment of the likelihood of the underlying data being obtained

IP Address	Country	Network Owner	Likely Device	File Downloaded ?	Retention Risk	Notes
193.111.141.30	Anonymised	Tor Node	Browser PC	Attempt_No_Excel	NA	German Tor Node
31.172.30.2	Anonymised	Tor Node	Browser PC	Attempt_No_Excel	NA	German Tor Node
202.55.151.4	Australia	Organisation	Search Engine Bot	Attempt_No_Excel	NA	Funnelback (Web Search Vendor)
182.118.20.219	China	ISP	Browser PC	Attempt_No_Excel	NA	China Unicom Henan province network
182.118.20.254	China	ISP	Browser PC	Attempt_No_Excel	NA	China Unicom Henan province network
41.129.120.57	Egypt	ISP	Browser PC	Attempt_No_Excel	NA	Link Egypt (Link.NET)
158.50.32.15	Hong Kong	Media Organisation	Browser PC	Attempt_No_Excel	NA	Agence France Press
175.139.243.13	Malaysia	ISP	Browser PC	Attempt_No_Excel	NA	TELEKOM MALAYSIA BERHAD
202.156.9.237	Singapore	ISP	Tablet	Attempt_No_Excel	NA	StarHub Cable Vision Ltd
121.189.37.16	South Korea	Search Engine	Search Engine Bot	Attempt_No_Excel	NA	ZUM Search Engine (Korean)
121.189.37.18	South Korea	Search Engine	Search Engine Bot	Attempt_No_Excel	NA	ZUM Search Engine (Korean)
121.189.37.20	South Korea	Search Engine	Search Engine Bot	Attempt_No_Excel	NA	ZUM Search Engine (Korean)
207.241.226.217	United States of America	Web Archive	Archive	Attempt_No_Excel	NA	Archive.Org
66.249.64.32	United States of America	Search Engine	Search Engine Bot	Attempt_No_Excel	NA	Google Bot, Document cached but Google presents a HTML version without the Excel Data
68.180.224.229	United States of America	Search Engine	Search Engine Bot	Attempt_No_Excel	NA	Yahoo, cached copy links to the Google HTML version
101.166.168.222	Australia	ISP	Tablet	Full	Low	Telstra
101.170.127.238	Australia	ISP	Tablet	Full	Low	Telstra
110.33.121.132	Australia	ISP	Tablet	Full	Low	Optus
120.151.212.149	Australia	ISP	Tablet	Full	Low	Telstra
124.183.240.241	Australia	ISP	Tablet	Full	Low	Telstra
165.118.1.50	Australia	AU_Gov	Browser PC	Partial_Excel	Low	WA Government Department of Finance and Deregulation
203.171.249.69	Christmas Island	ISP	Tablet	Full	Low	Christmas and Cocos Islands ISP
124.253.33.250	India	ISP	Tablet	Full	Low	QUADRANT TELEVENTURES LTD
31.172.30.3	Anonymised	Tor Node	Browser PC	Full	Medium	German Tor Node
5.199.142.195	Anonymised	Tor Node	Browser PC	Full	Medium	German Tor Node
94.242.251.112	Anonymised	Tor Node	Browser PC	Full	Medium	Tor Exit Node in Luxembourg
101.172.213.58	Australia	ISP	Browser PC	Full	Medium	Telstra
101.172.85.64	Australia	ISP	Browser Apple Computer	Full	Medium	Telstra
101.173.170.163	Australia	ISP	Browser Apple Computer	Full	Medium	Telstra
101.173.255.237	Australia	ISP	Browser Apple Computer	Full	Medium	Telstra
101.173.85.77	Australia	ISP	Browser Apple Computer	Full	Medium	Telstra
110.142.20.69	Australia	ISP	Browser PC	Full	Medium	Telstra
1.121.205.40	Australia	ISP	Browser PC	Full	Medium	Telstra
120.146.149.126	Australia	ISP	Browser PC	Full	Medium	Telstra
101.173.127.228	Australia	ISP	Browser Apple Computer	Full	Medium	Telstra
121.127.197.26	Australia	ISP	Browser PC	Full	Medium	iiNet (TRANSACT COMMUNICATIONS PTY. LIMITED, ACT)
101.173.170.150	Australia	ISP	Browser Apple Computer	Full	Medium	Telstra
101.173.42.166	Australia	ISP	Browser Apple Computer	Full	Medium	Telstra
121.45.205.103	Australia	ISP	Browser Apple Computer	Full	Medium	iiNet (Internode)
103.9.60.163	Australia	Organisation	Browser PC	Full	Medium	24/7 Distribution
110.174.21.238	Australia	ISP	Browser Apple Computer	Full	Medium	TPG
124.148.167.22	Australia	ISP	Browser Apple Computer	Full	Medium	iinet(Westnet)
110.23.86.158	Australia	ISP	Browser PC	Full	Medium	Optus
110.33.115.98	Australia	ISP	Browser PC	Full	Medium	Optus
117.120.16.131	Australia	Proxy	Browser PC	Full	Medium	Message Labs Proxy, IP anonymised and possibly origin country
117.120.18.133	Australia	Proxy	Browser PC	Full	Medium	Message Labs Proxy, IP anonymised and possibly origin country
124.190.199.41	Australia	ISP	Browser PC	Full	Medium	Telstra
119.15.97.138	Australia	ISP	Browser PC	Full	Medium	Infinite Networks
125.255.168.113	Australia	AU_Gov	Browser PC	Full	Medium	Australian Human Rights Commission
122.49.162.28	Australia	ISP	Browser Apple Computer	Full	Medium	iinet (ADAM INTERNET PTY LTD)
122.49.182.106	Australia	ISP	Browser PC	Full	Medium	iinet (ADAM INTERNET PTY LTD)

Summary of access to the file and risk assessment of the likelihood of the underlying data being obtained

IP Address	Country	Network Owner	Likely Device	File Downloaded ?	Retention Risk	Notes
123.2.119.147	Australia	ISP	Browser Apple Computer	Full	Medium	Dodo
124.149.41.56	Australia	ISP	Browser Apple Computer	Full	Medium	iinet(Westnet)
124.168.154.215	Australia	ISP	Browser Apple Computer	Full	Medium	iinet
14.202.141.28	Australia	ISP	Browser PC	Full	Medium	TPG
128.250.195.171	Australia	Organisation	Browser PC	Full	Medium	The University of Melbourne
129.78.233.210	Australia	Organisation	Browser PC	Full	Medium	University of Sydney (Web Proxy)
139.216.184.116	Australia	ISP	Browser PC	Full	Medium	Dodo
14.201.18.141	Australia	ISP	Browser PC	Full	Medium	TPG
14.203.97.18	Australia	ISP	Browser PC	Full	Medium	TPG
143.216.49.250	Australia	AU_Gov	Browser PC	Full	Medium	South Australian Government
152.91.9.167	Australia	Organisation	Browser PC	Full	Medium	Verizon Business Australia Pty Ltd
165.69.90.8	Australia	Media Organisation	Browser PC	Full	Medium	News Limited
202.14.81.49	Australia	AU_Gov	Browser PC	Full	Medium	Commonwealth Department of Finance
203.134.77.37	Australia	ISP	Browser PC	Full	Medium	M2 Telecommunications Group Ltd (Iprimus)
203.161.84.36	Australia	ISP	Browser PC	Full	Medium	AMNET INTERNET SERVICES PTY LTD
203.174.136.130	Australia	ISP	Browser Apple Computer	Full	Medium	AAPT LIMITED
203.2.218.145	Australia	Media Organisation	Browser PC	Full	Medium	Australian Broadcasting Corporation
203.20.130.22	Australia	AU_Gov	Browser PC	Full	Medium	Commonwealth Ombudsman
203.26.177.2	Australia	Media Organisation	Browser PC	Full	Medium	FAIRFAX MEDIA LIMITED
203.29.11.12	Australia	Commercial	Browser PC	Full	Medium	Copytech Engineering Pty Ltd
203.29.97.21	Australia	organisation	Browser PC	Full	Medium	Eclectech (Canberra)
203.30.93.5	Australia	Organisation	Browser PC	Full	Medium	KPMG (Melbourne)
203.41.198.36	Australia	ISP	Browser PC	Full	Medium	Telstra
203.62.63.51	Australia	AU_Gov	Browser PC	Full	Medium	ACT Department of Education
218.185.73.168	Australia	ISP	Browser PC	Full	Medium	Uecomm
59.167.200.170	Australia	ISP	Browser PC	Full	Medium	iiNet
60.225.90.26	Australia	ISP	Browser PC	Full	Medium	Telstra Internet
60.240.234.36	Australia	ISP	Browser PC	Full	Medium	TPG
60.241.244.66	Australia	ISP	Browser Apple Computer	Full	Medium	TPG
60.242.177.99	Australia	ISP	Browser Apple Computer	Full	Medium	TPG
61.14.115.68	Australia	Organisation	Browser PC	Full	Medium	Amnesty International Australia
142.214.192.68	Canada	Organisation	Browser Apple Computer	Full	Medium	Humber College Institute of Technology and Advanced Learning
99.248.118.30	Canada	ISP	Browser Apple Computer	Full	Medium	Rogers Communications Inc
185.10.104.132	China/Hong Kong	Search Engine	Browser PC	Full	Medium	Baidu (Hong Kong) Limited (Check 24/02/2014 revealed no cached copy)
161.53.74.122	Croatia	Organisation	Browser PC	Full	Medium	Savez studenata Fakulteta elektrotehnike i racunarstva (Croation Student Association?)
27.123.149.161	Fiji	ISP	Browser PC	Full	Medium	Vodafone Fiji Limited
31.172.30.3	Germany	Tor Node	Browser PC	Full	Medium	German Tor Node
5.199.142.195	Germany	Tor Node	Browser PC	Full	Medium	German Tor Node
193.108.78.10	Great Britian	Organisation	Browser PC	Full	Medium	HSBC Bank plc
217.20.24.210	Great Britian	Media Organisation	Browser Apple Computer	Full	Medium	Graphic News
86.176.215.226	Great Britian	ISP	Browser PC	Full	Medium	British Telecom
182.178.108.0	Pakistan	ISP	Browser PC	Full	Medium	Pakistan Telecommunication Company Limited
50.81.67.195	United States of America	Organisation	Browser Apple Computer	Full	Medium	Mediacom Communications Corporation
68.232.186.227	United States of America	Proxy/VPN provider	Browser Apple Computer	Full	Medium	London Trust Media Inc
72.194.80.186	United States of America	ISP	Browser Apple Computer	Full	Medium	Cox Communication
97.120.222.192	United States of America	ISP	Browser PC	Full	Medium	Qwest Communications International Inc
109.74.151.149	Anonymised	Tor Node	Linux scripted	Full	High	Slavakian Tor Exit Node
120.146.149.126	Australia	ISP	PC Microsoft Office	Full	High	Telstra
125.255.168.113	Australia	AU_Gov	PC Microsoft Office	Full	High	Australian Human Rights Commission
14.202.141.28	Australia	ISP	PC Microsoft Office	Full	High	TPG
203.134.77.37	Australia	ISP	PC Microsoft Office	Full	High	M2 Telecommunications Group Ltd (Iprimus)

Summary of access to the file and risk assessment of the likelihood of the underlying data being obtained

IP Address	Country	Network Owner	Likely Device	File Downloaded ?	Retention Risk	Notes
203.20.130.22	Australia	AU_Gov	PC Microsoft Office	Full	High	Commonwealth Ombudsman
203.29.11.12	Australia	Commercial	PC Microsoft Office	Full	High	Copytech Engineering Pty Ltd
203.41.198.36	Australia	ISP	PC Microsoft Office	Full	High	Telstra
203.62.63.51	Australia	AU_Gov	PC Microsoft Office	Full	High	ACT Department of Education
103.9.63.163	Australia	Organisation	PC Microsoft Office	Full	High	24/7 Distribution
117.120.16.131	Australia	Proxy	PC Microsoft Office	Full	High	Message Labs Proxy. IP anonymised and possibly origin country
139.216.184.116	Australia	ISP	PC Microsoft Office	Full	High	Dodo
14.203.97.18	Australia	ISP	PC Microsoft Office	Full	High	TPG
143.216.49.250	Australia	AU_Gov	PC Microsoft Office	Full	High	South Australian Government
152.91.9.167	Australia	Organisation	PC Microsoft Office	Full	High	Verizon Business Australia Pty Ltd
203.2.218.145	Australia	Media Organisation	PC Microsoft Office	Full	High	Australian Broadcasting Corporation
203.30.93.5	Australia	Organisation	PC Microsoft Office	Full	High	KPMG (Melbourne)
218.185.73.168	Australia	ISP	PC Microsoft Office	Full	High	Uecomm
60.241.192.243	Australia	ISP	Linux scripted	Full	High	TPG
31.23.71.164	Russian Federation	ISP	Archive	Full	High	OJSC Rostelecom Macroregional Branch South
207.241.226.36	United States of America	Web Archive	Archive	Full	High	Archive Org
157.55.32.142	United States of America	Search Engine	Search Engine Bot	Full	High	Bing Search Engine web crawler (Microsoft) File not present in cache as viewed 24/02/14
157.55.34.177	United States of America	Search Engine	Search Engine Bot	Full	High	Bing Search Engine web crawler (Microsoft) File not present in cache as viewed 24/02/14
66.249.74.220	United States of America	Search Engine	Search Engine Bot	Full	High	Google Bot. Document cached but Google presents a HTML version without the Excel Data
66.249.75.96	United States of America	Search Engine	Search Engine Bot	Full	High	Google Bot. Document cached but Google presents a HTML version without the Excel Data
207.241.226.231	United States of America	Web Archive	Archive	Full	High	Archive Org

Contact us

kpmg.com.au

KPMG
cutting through complexity