Health data governance – enforceable undertaking for re-identification of MBS/PBS data

3 May 2018

Authors: Simon Lewis, Steph Baker

Explore more insights 365

In brief

The Australian Information Commissioner has concluded an investigation into the re-identification of Medicare service provider data within the de-identified Medicare Benefits Schedule and Pharmaceutical Benefits Schedule data published by the Commonwealth Department of Health on data.gov.au in 2016. Background and further information about the data published can be found in a LegalTalk Alert which was distributed on 11 May 2017. In the investigation, the Commissioner found that the Department of Health failed to take reasonable steps to protect personal information and to implement practices, procedures and systems to ensure compliance with Australian privacy laws.

The Department of Health provided the Commissioner with an undertaking, which included a requirement to establish an external review and audit into departmental policies and procedures for the release of data based on personal information.

The incident in 2016 has provided a valuable learning experience for Government agencies as they explore how to best realise the value of public data sets. Importantly, the Commissioner observed that the risk of re-identification may require limiting the sharing of some types of data to trusted recipients, and/or using secured environments to share information, rather than simply relying on de-identification techniques.

In detail

Commonwealth government agencies are increasingly encouraged to publish data on the data.gov.au website. In 2015, the Prime Minister released the Australian Government Public Data Policy Statement, which, amongst other things recognises Commonwealth data holdings as a strategic national resource and commits Commonwealth entities to publish anonymised government data to data.gov.au 'by default', in order to facilitate innovation and productivity improvements across the Australian economy.

The *Privacy Act 1988* (Cth) (the **Privacy Act**) establishes Australian Privacy Principles (**APP**) that apply to Commonwealth agencies and other public and private sector organisations. Relevantly,

- APP 1 requires that reasonable steps are taken to implement practices, procedures and systems that will ensure compliance with the APPs.
- APP 11 requires that reasonable steps are taken to protect the personal information held from misuse, interference and loss and unauthorised access, modification or disclosure.

Publication of Sample Health Data

On 1 August 2016, the Commonwealth Government's Department of Health (the **Department**) published a sample of de-identified Medicare Benefits Schedule (**MBS**) and Pharmaceutical Benefits Schedule (**PBS**) data on data.gov.au, constituting 10 per cent of the MBS data collected between 1984 and 2014, and separate PBS data collected between 2003 and 2014 (the **Sample Health Data**).

Several de-identification methods were employed by the Department for different types of data (for example, information concerning medical professionals and patient data). For medical professionals, sensitive Medicare provider identification numbers were subject to:

- encryption,
- collapsing of locations into four broad geographic regions, and
- randomly perturbing the date of service and date of supply for services provided to each patient.

The de-identification methods employed by the Department were in line with approaches previously used, including in 2015 when the Department provided a selection of PBS data to a group of approved research organisations.

Re-identification

On 8 September 2016, the Department was alerted to the vulnerability of the Sample Health Data when data researchers from the Department of Computing and Information Systems at the University of Melbourne were able to re-identify Medicare provider numbers contained in the Sample Health Data. As a result, the Department removed the Sample Health Data from data.gov.au that day. PwC wrote extensively about the event in a LegalTalk Alert published in May 2017.

Since the event, the Department of Prime Minister and Cabinet has published the Whole of Government Sensitive Unit Record Open Data Process. The process document, published in 2016, describes the method that Commonwealth agencies should follow to release sensitive unit record data sets as open data (refer to the public statement here).

Investigation

The Department notified the Office of the Australian Information Commission (**OAIC**) of the data privacy issue on 9 September 2016. The OAIC immediately commenced an investigation under section 40(2) of the Privacy Act (the **Investigation**) in respect of a number of matters, including whether the Department of Health:

- disclosed personal information of healthcare providers or patients in the course of publishing the Sample Health Data,
- had appropriate practices, procedures and systems in place to ensure that personal information proposed for de-identified release was not inappropriately used or disclosed, and was appropriately secured, and
- took reasonable steps to secure the personal information used to form the Sample Health Data prior to its publication.

Following an investigation, the OAIC issued a report detailing the findings of the Investigation on 23 March 2018 (**Investigation Report**), which included that, although the Department had acted in good faith:

- The Sample Health Data did contain personal information about Medicare service providers, and the risk of re-identification of providers was not sufficiently low.

- The publication of the Sample Health Data represented a use of the information of Medicare provider numbers for the 'secondary purpose' of making the information available to researchers and the public at large. This was used without the consent of the individuals, and for a purpose which the individuals would be unlikely to 'reasonably expect'. Accordingly, the Department had breached APP 6, which regulates the use and disclosure of personal information.

- Where an entity proposes to publish de-identified information to the public, it must do so appropriately and with the necessary practices, procedures and systems in place to ensure compliance with the APPs. In this case, the Commissioner found that both the protective measures employed by the Department and its decision-making process to release the data were not sufficiently rigorous, and therefore in breach of APP 1 and 11.

In contrast, in the context of patient information, the OAIC found that:

> *…whilst the information contained in the Sample Health Data relating to medical treatments, fees for appointments and the like would constitute information 'about' a patient, the information was not 'reasonably identifiable'. This was because, even if the patient PINs were decrypted, there were additional protections in place that could prevent identification of individual patients. In particular, outliers or unusual features had been removed from the data.*

Accordingly, the patient information within the data set was not considered to be personal information.

A key takeaway from the Investigation Report is the acknowledgement of the difficulty in de-identifying large and rich datasets for publication, including an observation that:

> *'At this time, it is uncertain whether de-identification of a unit level dataset of this size and detail is possible to an extent that would permit full public release, while still maintaining the utility of the data'.*

Undertaking

In response to the findings of the Investigation Report, the Department provided the Commissioner with an undertaking that it will:

- commission an external review of the compliance of departmental policies and procedures for the release of data based on personal information with APS 1 and 11, and implement the recommendations from the independent review following the incident;

- commission an audit of the implementation of the findings of the external review, and

- treat proposals to release sensitive unit record data sets as 'high privacy risk projects' for the purpose of *Privacy (Australian Government Agencies – Governance) APP Code 2017,* and follow the Whole of Government Sensitive Unit Record Open Data Process from time to time, including considering alternatives to public release of data sets (e.g. release only in secure environments).

The takeaway

De-identifying large, rich data sets with confidence when disclosing to the public is always difficult.

The incident has highlighted the need to more carefully consider the potential benefits of sharing data sets against the risk of re-identification and the legal and reputational risks. In some cases, the risk of re-identification may require that sharing of data be limited to trusted recipients and/or require the use of secured environments, rather than sharing data more broadly.

The incident has provided an important opportunity for the Government to develop more robust approaches to publishing data. In line with this, the Department of Prime Minister and Cabinet published guidance in December 2016 on the *Process for Publishing Sensitive Unit Record Level Public Data as Open Data*, to assist agencies in releasing datasets related to sensitive information.

From July 2018, government agencies will also be subject to additional obligations in relation to APP 1.2 compliance via the *Privacy (Australian Government Agencies – Governance) APP Code 2018*. These obligations include the appointment of privacy officers and champions, among other things.

Let's talk

For a deeper discussion of how these issues might affect your business, please contact:

Tony O'Malley, Sydney
+61 (2) 8266 3015
tony.omalley@pwc.com

Bryony Binns, Sydney
+61 (2) 8266 1107
bryony.binns@pwc.com

Simon Lewis, Sydney
+61 (2) 8266 2161
simon.lewis@pwc.com