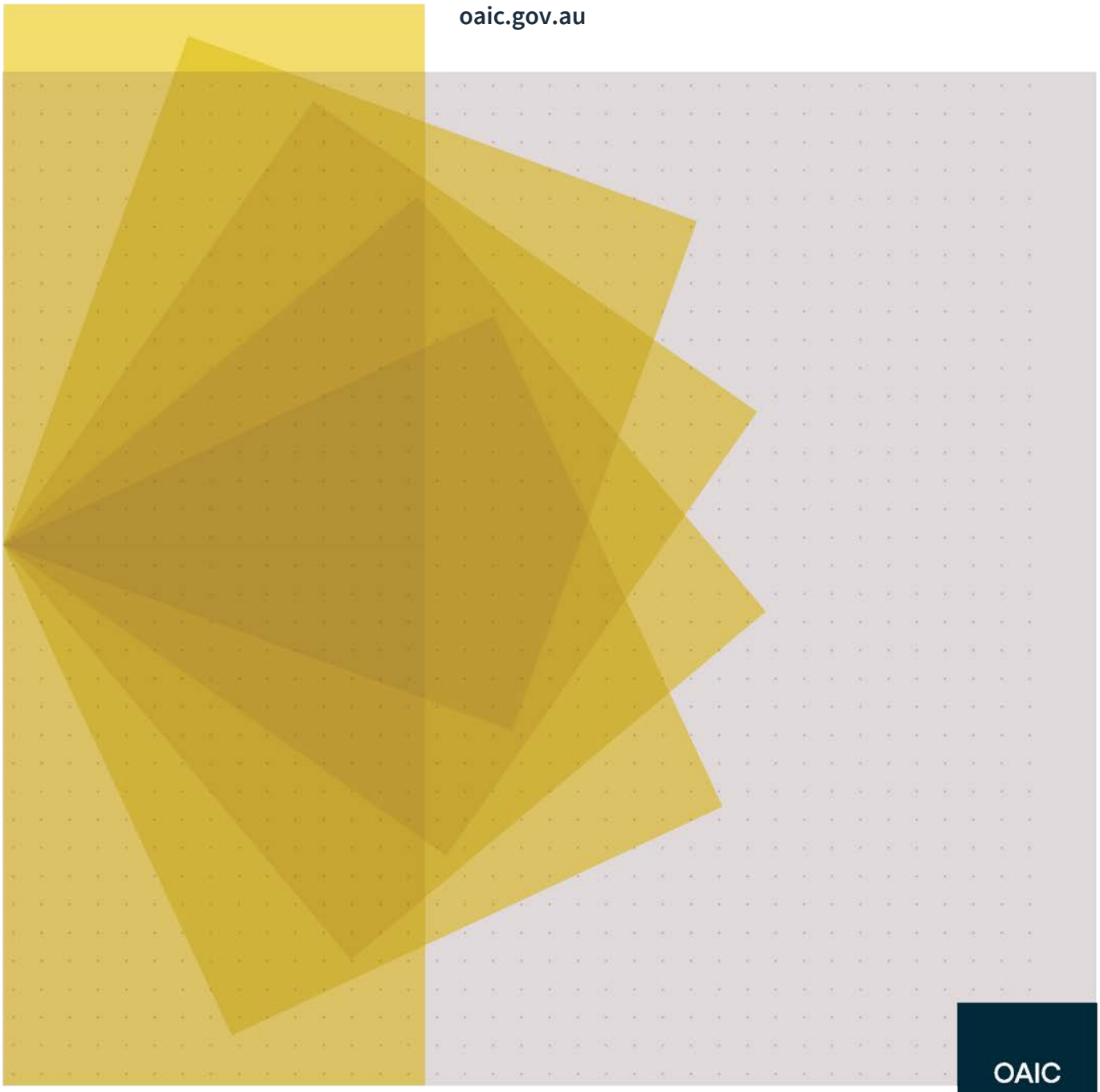
Australian Government

Office of the Australian Information Commissioner

Guide to Data Analytics and the Australian Privacy Principles

oaic.gov.au



March 2018

OAIC

Contents

Executive Summary

Overview

The use of data analytics is increasingly common across government agencies and the private sector. This has been driven by a fundamental shift in analytical processes, together with the availability of large data sets, increased computational power and storage capacity. The ability of data analytics to discover hidden insights has the potential to yield great benefits, including helping organisations to deliver better products and services, personalise people's online experiences, or develop stronger, evidence-based policies.

These activities, like all activities that use personal information, can have a significant impact on individual privacy. However, new data analytics processes, such as big data activities, can differ from 'traditional' data activities in some respects, and may therefore pose some specific privacy risks. For example, some data analytics activities have a tendency to:

- collate data from a wide variety of different sources, including from third parties
- generate new information through 'collection via creation'
- use data insights for a range of different purposes, including new purposes that may not have been anticipated, and
- retain data for a longer period of time than usual, in case it may be useful in future for an unspecified purpose.

This *Guide to data analytics and the Australian Privacy Principles* (the Guide) addresses some of these challenges. It is divided into two parts: Part One, which provides an introduction to the relevant key concepts when considering data analytics and privacy, and Part Two, which outlines how the Australian Privacy Principles apply to data analytics. Throughout this Guide, there are also a number of risk points and tips to help your organisation overcome some of the inherent challenges in conducting data analytics. Attachment 1 of this paper contains a consolidated summary of these risk points and tips.

A best practice approach to data analytics activities

When privacy is built into data analytics from the beginning, it not only helps organisations to comply with the *Privacy Act 1988* and Australian Privacy Principles, but can help drive innovation and build public and consumer trust. In this regard, the Guide makes a number of key recommendations to organisations to protect personal information when conducting data analytics processes:

✓ **Use de-identified data where possible.** Organisations should first consider whether a data analytics project requires the use of personal information. Successfully de-identified data is not personal information, meaning the *Privacy Act 1988* will generally not apply. See the De-identification section in Part One for further information.

- ✓ **Embed good privacy governance into your organisation by taking a privacy-by-design approach.** Integrate and embed privacy into your organisation's culture, processes and systems from the beginning through to the implementation of a project by adopting a 'privacy-by-design' approach. See the section on How to build privacy into your data analytics activities in Part One.
- ✓ **Conduct Privacy Impact Assessments for your data analytics projects.** Your organisation should consider conducting Privacy Impact Assessments for data analytics projects, to assist in identifying and addressing all relevant privacy impacts. A Privacy Impact Assessment should be treated as an iterative process. As the data analytics project progresses, new privacy risks may emerge, and your organisation should then consider how to address these emerging risks. See the section on How to build privacy into your data analytics activities in Part One.
- ✓ **Be open and transparent about your privacy practices**. Be upfront about your personal information handling practices, to help your organisation build trust and avoid being 'creepy'. Your APP Privacy Policy should clearly and simply describe the main functions and activities of your organisation, the general purposes that you put information to, and how your data analytics activities relate to this. See the Open and Transparent Management section in Part Two.
- ✓ **Know what you're collecting**. Using 'all the data' for 'unknown purposes' will expose your organisation to privacy compliance risks. Limit the collection of personal information where appropriate, and ensure that you are only collecting information which is reasonably necessary to pursue your legitimate functions and activities. See Collecting personal information in Part Two.
- ✓ **Be careful with sensitive information.** Be aware that data analytics may lead to the creation of and, consequently, the collection of, additional personal information. This will require particular care when sensitive information may be generated, based on inferred or derived data. If personal information is created which the organisation is not able to collect under APP3, it may need to be de-identified or destroyed. See Collecting personal information in Part Two.
- ✓ **Make your notices as clear and effective as possible.** Make your notices as dynamic, clear and user-friendly as possible. This will help you to establish the purposes for which data may be used at a later date (See Notices in Part Two). Key matters that your notice may set out could include:
    - o that analytics may be conducted on personal information for the purposes of determining which products or services your customers may be interested in
    - o that analytics may be conducted using information from a range of sources, such as information collected from third parties (and a description or list of those sources)
    - o any anticipated secondary purposes that data may be put to, and/or
    - o any anticipated disclosures of personal information to third parties (and a description or list of those entities).

- ✓ **Establish grounds for new uses of information.** Organisations should carefully consider whether uses and disclosures of personal information are compatible with the original purpose of collection - particularly when the information is collected from a third party. If not, organisations will need to rely on one of the exceptions in APP 6, such as having the individual's consent, or a reasonable expectation that the information would be used for that secondary purpose. See Using and disclosing personal information in Part Two.

- ✓ **Provide options.** When using privacy notices to inform individuals about a particular use or disclosure, organisations should consider how they might allow individuals to choose which uses and disclosures they agree to and which they do not. See Using and disclosing personal information and Notices in Part Two.

- ✓ **Ensure your marketing activities comply with APP 7.** Data analytics are often undertaken for the purposes of direct marketing. Ensure that your organisation provides clear 'opt-outs' and meets its other obligations under APP 7 when engaging in direct marketing, or when facilitating direct marketing for other organisations. See the section on Direct marketing in Part Two.

- ✓ **Ensure the accuracy of information.** In some circumstances, your organisation should take more rigorous steps to maintain the quality of information used for data analytics (see the section on Quality of personal information in Part Two). More rigorous steps may include:

    - ensuring that any third parties you deal with have good privacy practices in place to ensure the accuracy of the information they provide
    - verifying the accuracy of information which is not collected directly from the individual (particularly where information may be relied upon when making a decision which will affect the individual)
    - implementing procedures to monitor and record what type of personal information you are collecting, and
    - putting in place systems (including auditing and reviews) to check that the analytic techniques used (such as algorithms) are operating appropriately and are fit for purpose.

- ✓ **Protect information in line with your risk assessments.** Take reasonable steps to monitor and protect against the security risk posed by data analytics activities, noting that large, detailed datasets can become 'honey pots' of valuable and sensitive personal information. Undertaking an information security risk assessment may assist. See Security of personal information in Part Two.

Part 1: Introduction and key concepts

1.1 About this Guide

This *Guide to Data Analytics and the Australian Privacy Principles* (the Guide) provides guidance about the Australian Privacy Principles (APPs) and how they apply to data analytics activities, which include (but are not limited to) big data, data mining and data integration. The Guide is intended for both Australian Government agencies and private sector organisations (collectively referred to organisations in this Guide) covered by the *Privacy Act 1988* (Privacy Act).[1]

The aim of the Guide is to assist organisations to identify and take steps to address the privacy issues that may arise. The Guide should also be read in conjunction with the *Australian Privacy Principles Guidelines* (APP Guidelines) which outline the mandatory requirements of the APPs and how the Office of the Australian Information Commissioner (OAIC) interprets the APPs, together with guidance for best practice.

This Guide is not legally binding. The OAIC will however refer to this Guide when undertaking its functions under the Privacy Act. The OAIC's Privacy regulatory action policy provides information on when and how we may exercise our functions.

This Guide assumes some knowledge of privacy concepts. The Guide includes a number of examples, which are provided for illustrative purposes only.

1.2 What do we mean by data analytics?

Data analytics describes processes or activities which are designed to obtain and evaluate data to extract useful information. The scope of data analytics is broad and covers several terms and concepts such as 'big data', 'data integration', data mining' and 'data matching' which are discussed below. However, data analytics is an evolving term, and the discussion below is not intended to be an exhaustive list of concepts included in the scope of this Guide.

Big data

Fundamental shifts in analytical processes, together with large data sets, increased computational power and storage capacity has led to the ability to bring about enormous social and economic benefits. There is no definitive definition for big data. However, Gartner's 'three Vs' definition is often used:

> […]high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight, decision making, and process optimization.[2]

[1] For more information on the jurisdiction of the Privacy Act, see our 'Privacy Act' webpage.

[2] Gartner, *The Importance of 'Big data': A Definition*, cited in Department of Finance and Deregulation, *The Australian Public Service Big Data Strategy* [PDF], 2013, p 8.

According to this definition, big data encompasses three dimensions: volume (the amount of data), velocity (the speed of data capture and processing), and variety (the use of different data types and sources). These dimensions have changed the way organisations use data to identify trends and challenges, by analysing large data sets, often from a variety of sources, quickly.

Data integration

'Data integration'[3] refers to the bringing together of multiple datasets, to provide a new dataset (usually for statistical or research purposes). Data integration refers to the full range of practices around the process, including data transfer, linking and merging the data and dissemination. 'Data linking' is an element of data integration, which is the process of creating links between data from different sources based on common features present in those sources.

Data mining

Data mining is the process of discovering meaningful patterns and trends by sifting through large amounts of data stored in repositories. Data mining employs pattern recognition technologies, as well as statistical and mathematical techniques.[4] For example, it is increasingly used by organisations to enable effective and targeted marketing campaigns, and develop products to increase sales and profitability.

Data matching

Data matching means the bringing together of at least two data sets that contain personal information, and that come from different sources, and comparing those data sets to produce a match.[5] Data matching is usually conducted by government agencies, and is performed for a range of purposes including fraud detection and facilitating debt collection.

Although this guide may be useful when conducting data-matching activities, you should refer to the OAIC's *Guidelines on Data Matching in Australian Government Administration* for specific guidance and good practice principles.[6] You should also be aware that the use of tax file numbers to detect incorrect payments is subject to the requirements of the *Data-matching Program (Assistance and Tax) Act 1990)* and relevant guidelines.

1.3 Benefits and challenges of data analytics

Through the amassing, aggregating and analysing of data to discover new relationships, data analytics activities have the potential to bring about enormous societal, economic and or/personal benefits, for example:

- Promoting strong policy outcomes and openness of government through the discovery and application of new data insights to the policy development process.

[3] National Statistical Service - Statistical Data Integration involving Commonwealth Data, 2017, Glossary.

[4] Gartner, 2017, 'IT glossary'.

[5] OAIC's *Guidelines on Data Matching in Australian Government Administration*.

- Improving our understanding of diseases by analysing medical records, which can in turn assist with the development of new medicines.
- Predicting and responding to disasters, where data can be analysed to predict where earthquakes might occur next, and patterns of human behaviour which can help aid organisations to provide emergency assistance to survivors.
- Making it easier for individuals to make consumer choices and save money, by better understanding their spending and patterns of consumption. For example, data can be analysed to help draw consumers' attention to relevant products or services when shopping online, or relevant content when using online media streaming channels.

Example: In 2015 the Humanitarian Data Exchange was used to assist in delivering effective relief efforts following the Nepal earthquake. A task force of about 2,000 people from 80 countries analysed 'millions of Nepal-related tweets to build several databases'. This data helped produce 'quick-and-dirty' maps to coordinate humanitarian relief efforts by the government, the UN, and NGOs.[7]

However, data analytics can also have significant privacy implications. The collection and generation of personal information in unexpected ways, and the use of complex or opaque algorithms can be 'creepy' for affected individuals.

Example: In 2014, Facebook conducted a 'happy-sad' emotional manipulation experiment, by splitting almost 700,000 users into two groups and manipulating their newsfeeds to be either 'happier' or 'sadder' than normal. The results were then analysed, and it was found that users tended to post positive or negative comments according to what was in their news feed. This led to significant user backlash, with users describing the research as 'creepy' and 'terrifying'. Facebook had relied on user consent for the research program, on the basis that research was included in their Terms and Conditions. Following these reactions, Facebook's Chief Technology Officer announced in a blog that the social network had 'mishandled the study'. Facebook has since instituted a new framework for handling research.[8]

1.4 Why are today's data analytic techniques different?

While organisations have undertaken data analytics activities for a long time, more recent trends in data analytics activities have some unique characteristics which make them different from more 'traditional' methods of data analysis. This includes the tendency to:

- **Apply algorithms to identify relationships —** Data analytics can use sophisticated analytics or algorithms, involving artificial intelligence and machine learning. By undertaking new analyses

[7] Mark Wilson, 2015, 'How The Candy Crush Of Data Is Saving Lives In Nepal', *Fast co designs*.

[8] Kramer ADI, Guillory JE and Hancock JT (2014) 'Experimental evidence of massive-scale emotional contagion through social networks', Proceedings of the National Academy of Sciences 111(24): 8788–8790.

of datasets using these techniques, new relationships and insights begin to emerge. This is different to how data was analysed in the past where particular hypotheses were tested.

- **Use complex processing which lacks human input —** The complexity of the data processing techniques used, and automated decision-making, can make it difficult to understand the decision-making rationale.
- **Collect 'all the data' —** Traditionally, data analysis involves a representative or random sample of the population. However, data analytics activities can now typically collect and analyse all of the data that is available. This practice is driven by the ability to more easily collect, store and analyse large volumes of data.
- **Use data for a secondary purpose —** Data analytics activities often use data for secondary purposes (that is, not the purpose for which data was originally collected). In addition, these activities often use data collected from a range of sources including third party organisations. This practice can be advantageous given the ability to use data analytics techniques to look for new insights and find correlations between extremely disparate datasets.
- **Use new methods of collecting data —** Developments in technology, such as the 'internet of things' (IoT) has resulted in large amounts of data being collected about people, which may be provided in a 'passive' rather than active way. Information is collected through constant monitoring by devices (such as mobile apps), or inferred, derived or created through analytics.

1.5 How does the Privacy Act apply to data analytics?

The Privacy Act regulates how organisations handle personal information, including sensitive information. It includes 13 APPs which set out standards, rights and obligations in relation to the handling of personal information.

Under the Privacy Act (s 6(1)), personal information is:

> **'Information or an opinion about an identified individual, or an individual who is reasonably identifiable:**
>
> **(a) whether the information or opinion is true or not; and**
>
> **(b) whether the information or opinion is recorded in a material form or not.'** [9]

What constitutes personal information will vary depending on whether an individual is reasonably identifiable in a particular circumstance. Common examples of what constitute personal information are included in the OAIC Guide on 'What is Personal Information'.

When conducting data analytics, organisations should remember that personal information includes opinions or inferences drawn about people from other data, whether or not these are accurate.

[9] See Privacy Act s (6)(1).

Example: An organisation infers information about an individual from their online activities, such as their tastes and preferences from online purchases they have made from their web browsing and/or transaction history. Even if the inferred information is incorrect, it is still personal information.

Data used for data analytics may include personal information, and the activities will therefore be subject to the Privacy Act. However, in many cases, these activities do not require the use of personal information. In these cases, organisations may choose to use de-identified information. Information that has undergone an appropriate and robust de-identification process is not personal information, and is therefore not subject to the Privacy Act.

Importantly, whether information is personal information (or de-identified) should be determined on a case-by-case basis, with reference to the specific circumstances and context of the situation. Some information may not be personal information when considered on its own. However, when combined with other information held by (or accessible to) an organisation, it may become 'personal information'. Information holdings can therefore be dynamic, and the character of information can change over time. See our Guide on 'What is Personal Information' for further information.

Example one: A government agency is planning on conducting data analytics activities to model the likely causes and impacts of fires in the future using datasets about fires managed by fire and rescue services. These datasets may not appear to contain any personal information when considered in isolation as they are appear to be about something non-personal. However, despite not being directly about people, fires often happen in people's homes. So while such fire data is not primarily 'about' people, it may be information 'about' an individual in some situations, primarily where the fire happens at a person's address. Accordingly, in this situation, the data custodan errs on the side of caution and treats the information as 'personal information'.

Example two: As part of its customer loyalty program, a company holds potentially identifying information about each of its members. The company wants to conduct data analytics on this information, so it removes some of the identifying details (for example name, address, date of birth, contact numbers) and instead assigns each customer file a unique customer identifier. The customer files are then given to a third party data analytics company for research purposes. In the hands of the third party data analytics company, this information may not be personal information. However, if employees within the company are able to match the unique identifier with the original customer record to identify the person, this information may be personal information when handled by the company.

1.6 De-identification

What is de-identification?

As discussed above, 'de-identified' information is information which has undergone a process of de-identification[10], and no longer falls within the definition of 'personal information' under the Privacy Act.

De-identification involves the removal or alteration of information that identifies a person or is reasonably likely to identify them, as well as the application of any additional protections required to prevent identification.

In line with this, a de-identification process generally includes two steps. The first is the removal of direct identifiers, such as an individual's name, address or other directly identifying information. The second is taking one or both of the following additional steps:

- removing or altering other information that may allow an individual to be identified (for example, because of a rare characteristic of the individual or a combination of unique or remarkable characteristics that enable identification), AND/OR
- putting controls and safeguards in place in the data access environment, which will appropriately manage the risk of re-identification.

For information to be de-identified, it must have a very low risk of re-identification, having regard to all the circumstances (and in particular, the context in which the information will be handled, including who will have access to the data, and what other information they might have access to). Put another way, information will be de-identified where there is no reasonable likelihood of re-identification occurring.

The OAIC and CSIRO Data 61 have released the *De-Identification Decision-Making Framework* to assist organisations to de-identify their data effectively. The De-Identification Decision-Making Framework is a practical and accessible guide for Australian organisations that handle personal information and are considering sharing or releasing it to meet their ethical responsibilities and legal obligations, such as those under the Privacy Act. The guide is an adaptation of the existing UK *Anonymisation Decision-Making Framework* for an Australian audience.

You can also refer to the OAIC's *De-identification and the Privacy Act* Guide, which provides general advice about de-identification and protecting privacy to maximise the utility and value of data while safeguarding privacy.

Risk point: Data used for data analytics may include personal information, and the activities will therefore be subject to the Privacy Act.

[10] A number of different terms are used in Australia to describe processes similar to de-identification, for example anonymisation and confidentialisation. Given the sometimes differing uses of terminology, it is a good idea to check in any given scenario or conversation that the terminology being used is understood consistently by all parties. Organisations should be aware that sometimes de-identification is used to refer to the removal of 'direct identifiers', such as name and address. In this guide, de-identification is used consistently with the meaning in the Privacy Act. It is therefore important to be aware that the removal of name, address or other identifiers alone would not result in de-identification for the purposes of the Privacy Act.

Privacy tip: Successfully de-identified data is not personal information, meaning the Privacy Act will generally not apply.

When should de-identified information be used in data analytics?

Organisations considering undertaking data analytics should consider whether de-identified personal information could be utilised as it allows organisations to use, share and publish information without jeopardising personal privacy. De-identifying information also lessens the risk that personal information will be compromised should a data breach occur.

De-identified data may be used in many different stages of a project involving data analytics:

- after the personal information is collected
- prior to analysis
- during the 'discovery' phase (for example as part of a big data project) to better assess risks to personal information or before the analytical outcomes are presented, or
- when data is shared externally or within organisations.

Ensuring data remains de-identified

It is important to remember that de-identification is not a fixed or end-state. The same information may be personal information in one situation, but de-identified information in another. Further, depending on the particular controls, information may be de-identified for some parts of an organisation, but remain personal information in others.

For example, suppose an organisation undertakes a de-identification process on a dataset, to enable an in-house big data project to be conducted using that data. However, the organisation retains a copy of the original dataset, which would enable them to re-identify the data subjects in the big data project if they wished to do so. To ensure that this particular use of the dataset is de-identified (and therefore outside the scope of the Privacy Act), additional controls may need to be put in place to prevent re-identification during the project. This may include technical and/or environmental controls to prevent those who are using the de-identified dataset from accessing the original dataset.

In this scenario, the in-house research team may be using data that is de-identified for the purposes of the Privacy Act, while those who handle the original, identified dataset within the same organisation would still be subject to Privacy Act obligations. This is acceptable, however organisations need to be aware of the context-dependent nature of de-identification and treat data accordingly. Organisations should take a risk-management approach when handling de-identified data which acknowledges that while the APPs may not apply to data that is de-identified in a specific context, the same data could become personal information in a different context.[11]

Conducting a risk assessment for re-identification

Data analytics activities may increase the risk of re-identification, because of the volume of data and the power of the analytics. In many (if not all) cases where a de-identification process is undertaken, the risk of re-identification will never be totally eliminated, and re-identification will

[11] For further information on this, see the OAIC's *De-identification and the Privacy Act* Guide.

remain technically possible. This means that, in practice, whether or not de-identification has been successful will turn on whether there is a 'reasonable' likelihood of re-identification occurring.

Where an organisation is proposing to de-identify personal information for a data analytics activity, they should therefore undertake a risk assessment to consider the risk of re-identification. For example, whether it will be used for data analytics activities within the entity, or whether the de-identified data will be disclosed to another entity for this purpose. This could be undertaken as part of a Privacy Impact Assessment for the proposed data analytics activity (see section on Open and Transparent management of information for more about conducting PIAs for data analytics activities).

In undertaking a risk assessment organisations should consider the variety of information that will be brought together, the algorithms to be applied, and how the outcomes will be used or disclosed. For example, where the de-identified information will be made available to other entities or the public generally, the relevant factors to consider may include the difficulty, practicality, cost and likelihood that the information may be re-identified. It is also important to assess both the risk of re-identification from legitimate access to the de-identified data sets, as well as the risk of unauthorised intrusion by external parties.

Following a risk assessment, appropriate mitigation strategies should be implemented. This may include using different or additional de-identification techniques. It may also include placing restrictions on the use of the de-identified information.

Where personal information is appropriately de-identified and mitigation strategies are implemented, the risk of re-identification should be low. If, however, personal information is re-identified, the Privacy Act regulates how it is to be handled.

Risk point: Where de-identification is not done properly, data analytics activities may lead to re-identification of personal information.

Privacy tip: Undertake a risk assessment to consider the likelihood of re-identification. Use appropriate de-identification techniques and implement risk mitigation strategies.

Case study: In 2000, a university student used publicly available health insurance information on workers employed by the state of Massachusetts. The information had the names, addresses, social security numbers and some other 'identifying' information of the workers removed. The researchers obtained the state voter rolls for the capital city of Cambridge. These provided the name, postcode, address, sex and date of birth of every registrant. The insurance data revealed that there were six people in the city of Cambridge who were born on the same day as the State Governor. Half of those were men. The voter data allowed the researchers to claim the Governor as the only one of those persons living in a particular postcode in Cambridge. The corresponding health-insurance data revealed the Governor's health information, including medical diagnoses and prescriptions.[12]

[12] Paul Ohm, 2010, 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization', CLA Law Review, Vol. 57, p. 1701.

1.7 How to build privacy into your data analytics activities

Privacy is not an obstacle to innovation. Rather, embedding strong privacy protections into your organisation's data analytics activities will not only benefit affected individuals, but will also be beneficial for your organisation. It will enjoy increased stakeholder trust, which in turn supports innovation. Some key tools and approaches that can help organisations to build privacy into data analytics include de-identification (as discussed in the section above), privacy-by-design and the use of Privacy Impact Assessments (PIAs). Organisations may also wish to consider developing their own approaches to consider their social responsibility that go beyond legal compliance in order to build relationships of trust with the public.

Privacy-by-design in data analytics

'Privacy-by-design'[13] is a holistic approach where privacy is integrated and embedded in an entity's culture, practices and processes, systems and initiatives from the design stage onwards. In practice, this means organisations should embed 'privacy-by-design' across and within their organisation, as well as for individual projects and activities. This includes taking a risk management approach to identifying privacy risks and mitigating those risks. Embedding 'privacy-by-design' will lead to a trickle-down effect where privacy is considered automatically throughout the project, resulting in better overall privacy practice and compliance. Other key principles of privacy-by-design include:

- managing privacy proactively, rather than retrospectively after any privacy issues come to light
- recognising it is possible to have both 'good privacy' and effective, innovative use of data
- keeping the activity user-centric by offering strong privacy defaults, appropriate notifications systems, and empowering user-friendly options, and
- end–to–end security throughout the full lifecycle of the project, ensuring that all personal information is kept securely from collection through to destruction.

Adopting a privacy-by-design approach can be extremely valuable when conducting data analytics activities involving personal information for the success of the project itself. This is because if a privacy risk with a data analytics project is identified, it can be an opportunity to find creative technical solutions that can deliver the real benefits of the project while also protecting privacy and enhancing trust and confidence in the project.

Conducting a Privacy Impact Assessment for data analytics

What is a Privacy Impact Assessment (PIA)?

A PIA is a practical tool which can help to facilitate 'privacy-by-design' because it encourages organisations to develop projects with privacy designed into the project, rather than being bolted on afterwards.

[13] Privacy-by-design was first developed in the 1990s by the former Information and Privacy Commissioner of Ontario, Canada, Dr Ann Cavoukian. Since then it has been adopted by both private and public sector bodies internationally. For further information, see the Information and Privacy Commissioner of Ontario's *Privacy by Design* resource at www.ipc.on.ca/resource/privacy-by-design.

A PIA is a valuable tool that:

- systematically assesses the privacy impacts of a project, and
- recommends strategies to manage, minimise or eliminate those impacts.

The OAIC recommends that organisations conduct PIAs as part of their regular risk management and planning processes when an entity is developing or reviewing a project that uses data analytics. However, a PIA is much more than a simple compliance check. It should 'tell the full story' of a project from a privacy perspective, going beyond compliance to also consider the broader privacy implications and risks, including whether the planned uses of personal information in the project will be acceptable to the community. It can also help to assess the overall proportionality of the data analytics project when considering whether the use of personal information strikes an appropriate balance between the objectives of the project and any impact on privacy. This is particularly important when information is collected on a compulsory basis and individuals do not have meaningful choice about whether to provide their information (for example by government agencies in exchange for essential payments or services).

Undertaking a PIA for data analytics activities

Generally, if personal information is involved in the data analytics project, some form of PIA will be necessary. Depending on how personal information is handled in the project, the PIA process might be brief, or complex. The greater the data analytics complexity and higher the privacy risk is, the more likely it will be that a comprehensive PIA will be required to determine and manage its impacts. In particular, when determining how high risk the data analytics project will be, some key questions to consider include:

- Does the project involve any new or changed ways of handling personal information?
- Is the project likely to have a significant impact on individuals?
- Is the project likely to be perceived as privacy intrusive or 'creepy'?

It can sometimes be challenging for an organisation to know when to start carrying out a PIA for complex data analytics projects (such as big data activities) due to the initial lack of clarity about the direction that the project will take. However, we recommend that organisations should start the PIA process as soon as possible to start describing their aims and to start thinking about the potential privacy impacts for the project. For example, if a company collects data for a particular purpose and, as part of the big data activity, it would use the data for another purpose (known as a 'secondary purpose'), the PIA might explore how this might be done in accordance with the APPs and how any privacy impacts will be addressed. It is important to remember that a PIA is an iterative process which will continue to develop as the project evolves. You should continue to review your PIA to ensure the privacy solutions are working as expected. As the objectives and purpose of the project shift, new privacy risks will emerge. Organisations will need to continue considering how they will address these emerging risks.

When conducting a PIA for data analytics:

- If the direction of a data analytics project seems unclear, you should err on the side of caution and begin the PIA process anyway.
- Where it is possible, clearly describe the predicted information flows. Where the purpose is not yet clear, one potential solution would be to de-identify the datasets. When correlations of

interest are discovered, the organisation would then be able to identify the aims of any further processing before handling any personal information.

- As the data analysis progresses, there may be new risks or privacy impacts which are identified. Organisations should continue to identify and record measures to address these risks.

More information about undertaking a PIA is provided in the Guide to undertaking privacy impact assessments.

Example: An insurance company is considering undertaking data analytics to find unknown correlations in their data. Initially, the company doesn't know what all the likely privacy impacts might be. However, the analyst expects that it is likely that the processing will show a correlation between an individual's risk behaviours and their premium levels.

The analysts decides to conduct a PIA to explain the potential insights/new data which may be generated from undertaking the analysis, and how the data will be sourced and managed for the activity. Through conducting the PIA, the company builds in privacy-enhancing practices such as the use of de-identification techniques and internal security measures (to keep data de-identified), as well as updating their notifications systems to provide customers with an opportunity to reflect their preferences about which purposes they would allow their data to be used for.

Risk point: PIAs can be more challenging for large scale data analytics projects (such as big data activities), as an organisation may not know exactly how it is going to use the data, or what data it will use during the initial 'discovery phase'.

Privacy tip: Even if the direction of a data analytics project seems unclear, err on the side of caution and begin the PIA process anyway. It is important that a PIA is treated as an iterative process, which continues to develop. As a project evolves, the potential privacy risks will become clearer and your organisation will be able to better address them.

Social responsibility in data analytics projects

Considering your social responsibility or having an ethics based approach can help to build trust and informed confidence with the public, which will ultimately deliver long term benefits for your organisation. Taking this approach to data analytics can help you to ensure that the processing of personal information as part of your organisation's data analytics is carried out in a fair, transparent, responsible and ethical manner.

For example, ask yourself - is the activity being done in a way that is respectful to the individual? Is the activity in line with community expectations? Will the activity have an adverse impact on individuals? Is the activity reasonable and proportionate in all the circumstances?

What type of approach you decide to establish will depend on how risky the data analytics being carried out are, the context of the project, and the quantity and type of personal information. There are a range of ways that ethics can be incorporated into a project, but examples include:

- **As part of your PIA** — which considers whether the planned uses of personal information in the project will be acceptable to the community.

- **Developing organisational values** — organisational values may set out the parameters for your project, and be included in your APP privacy policy (and/or APP privacy notice). This may assist readers to understand how your organisation will use their personal information and help to reassure affected individuals and build trust in your organisation's use of data. In addition, these values may be helpful for employees in any new data analytics activity to guide decision making and guide PIAs.
- **Use of an ethical framework** — an ethical framework generally sets out categories of ethical issues, standards or guiding questions when using and managing data, for example the Data Governance Australia Code of Practice.
- **Use of ethics committee** — Some organisations may have an internal area which considers community expectations or the social impacts of a project. Others may utilise external committees which bring people from diverse backgrounds to scrutinise projects and assess issues arising from data analytics.

Example: A government department is collaborating with researchers from a university on a data analytics project to improve health and education outcomes. It is exploring the idea of creating an automated tool that can predict the likelihood of the education and health outcomes of a newborn baby by looking at data on their parent's demographics and socio-economic status. The government department undertakes a comprehensive Privacy Impact Assessment, conducts an ethical review, and engages in extensive engagement with key stakeholders. Through these reviews, the government department identifies a range risks including stigmatisation of people identified as having high scores, risk that the tool may produce a number of false positives or false negatives, questions the actions and obligations of agencies in relation to high risk scores, and potential impacts on people's interactions with services and government agencies. The government agency then considers whether the risk of harm to individuals is proportionate to the policy objective it is seeking to achieve, and explores alternative options to mitigate the risks to achieve its objective.

Part 2: Australian Privacy Principles and data analytics

Some data analytics activities may challenge how key privacy principles, including notice and consent, data collection and retention minimisation, as well as use limitation, work in practice. The APPs are, however, technologically neutral and principles-based. This means organisations have the flexibility to tailor their personal information handling practices for data analytics.

While your organisation must consider all APPs when handling personal information, this Guide addresses how the following APPs apply when conducting data analytics:

- open and transparent management of information (APP 1)
- collection of personal information, and dealing with unsolicited personal information (APPs 3 and 4)
- notice of the collection of information (APP 5)
- use or disclosure of information (APP 6)
- direct marketing (APP 7)
- cross border disclosure of personal information (APP 8)
- quality of personal information (APP 10) and,
- security of personal information (APP 11).

The requirements in each of these principles interact with and complement each other.

The following part of the guidance takes you through each of the APPs listed above, and outlines the factors to consider when conducting data analytics. It also discusses risks points and challenges when applying the APPs, as well as strategies and privacy tips to address them.

2.1 Open and transparent management of personal information (APP 1)

The objective of APP 1 is to ensure that organisations manage personal information in an open and transparent way. It is the bedrock principle for the APPs. By complying with this APP your organisation will be establishing a culture and set of processes that will assist you in complying with all the other APPs, right from the start.

APP1 does this in two key ways:

- First, by requiring organisations to take reasonable steps to establish and maintain internal practices, procedures and systems that ensure compliance with the APPs (APP 1.2).
- Second, by requiring organisations to have a clearly expressed and up to date APP Privacy Policy describing how it manages personal information (required by APP 1.3).

Australian Government agencies should also be aware that as of July 2018, they will have specific obligations under APP 1.2 as set out in the *Privacy (Australian Government Agencies – Governance) APP Code 2017*.[14]

The complexity of data analytics can mean that the processing is opaque to the individuals whose data is being used. It may not be apparent to them their data is being collected, or how. Despite the challenges, with planning and foresight, transparency and good privacy governance in relation to data analytics can be achieved. Being open and transparent about how you will handle personal information (including the purpose of your algorithms), will help to ensure that you have a culture that respects and protects personal information. It also plays a key role in building public and consumer trust, improving the quality of data analytics, and encouraging innovation.

The section below sets out information about adopting good governance, systems and processes, and having a clearly expressed APP Privacy Policy within a data analytics context.

Good privacy management and governance

APP 1 requires your organisation to take reasonable steps to establish and maintain internal practices, procedures and systems that ensure your compliance with the APPs. In practice, this means that you should appoint key roles and responsibilities for privacy management and adopt governance mechanisms, such as regular staff training. You should also develop a program of proactive review and audit the adequacy and currency of your organisational practices, procedures and systems involving data analytics.

Below are some best practice tips to ensure good privacy management and governance:

- **Embed a culture that respects and protects personal information**
    - ✓ Appoint a senior member of staff to be responsible for the strategic leadership and overall privacy management.
    - ✓ Appoint a privacy officer to be responsible for the day to day managing, advising and reporting on privacy issues.
    - ✓ Integrate privacy training into induction processes and provide regular staff training to those who conduct data analytics.
    - ✓ Record and report on how datasets containing personal information are treated, managed and protected.
    - ✓ Have a privacy management plan, which includes information about how you will ensure any data analytics activities will comply with obligations under APP 1.2.
- **Establish robust and effective privacy practices, procedures and systems**
    - ✓ Use PIAs to inform data analytics.
    - ✓ If your entity is using de-identified information, ensure that you have strong processes in place to ensure that personal information is correctly de-identified.

[14] Further information about the development of the *Australian Government Agencies Privacy Code* is available at our Australian Government Agencies Privacy Code webpage.

- ✓ Have clear processes for reviewing and responding to privacy enquiries, complaints or requests for access to personal information.
- ✓ Develop policies and procedures for personal information used for data analytics, including clear APP Policies and Notices.
- ✓ Have a data breach response plan.

- **Regularly review and evaluate your privacy processes**
  - ✓ Systematically examine the effectiveness and appropriateness of the privacy practices, procedures and systems to ensure they remain effective and appropriate.
  - ✓ Measure your performance against your privacy management plan.
  - ✓ Create channels for both your staff and customers so you can continue to learn lessons from data analytics, privacy complaints and breaches, as well as customer feedback.
- **Enhance your response to privacy issues**
  - ✓ Use the results of your evaluations to make necessary and appropriate changes to your organisation's practices, procedures and systems.
  - ✓ Consider having data analysis externally assessed to identify areas where privacy processes may be improved.
  - ✓ Continuously monitor and address new security risks and threats to data held.

The OAIC has developed a range of tools to assist you to develop or review your privacy program and related governance structures, and to meet the requirements set out in APP 1. 2, including the Privacy Management Framework.

APP privacy policies

APP 1.3 requires organisations to have clearly expressed and up-to-date privacy policies describing how they manage personal information. An APP privacy policy is a key tool for ensuring open and transparent management of personal information.

An APP privacy policy should describe the main functions and activities of an organisation, and identify those that involve personal information handling.

An APP privacy policy should generally not be used as a substitute for an APP 5 privacy notice. It is more general in nature, and focuses on the entity's information handling practices.

What's the difference between a privacy policy and a privacy notice?

A privacy notice should provide specific information relevant to a particular collection of personal information. The purpose of the privacy notice is to provide an individual with enough relevant information to make an informed decision about whether to provide their personal information to an entity. A privacy policy is more general in nature about the entity's information handling practices.

Organisations undertaking data analytics activities should include general information about those activities in their APP privacy policy. For example, by including that they undertake data analytics for marketing or policy development.

Below are some tips to make it genuinely informative and manageable.

- **Think about your audience.** Don't treat the privacy policy as merely a legal document to manage legal risk. It should be a document that creates trust in your entity and speaks to your customers or clients.
- **Don't just repeat the words in the APPs**. Make the privacy policy specific to your business or operation.
- **Consult**. Seek input from all areas of your entity including your public relations department, which may have ideas about innovative formats for better communicating the policy, for example, through video or other mechanisms relevant to the communication channel (paper, telephone, email, online) that you are using.
- **Focus on what is important to the reader.** Do not try to cover everything in minute detail.
- **Keep it simple.** Use simple, clear language, and be as brief as possible.
- **Consider having more than one policy**. For large or complex organisations, consider whether you need to have more than one policy (for different parts of your operation or business, or different functions or activities).
- **Take a layered approach**. For example, for online publication provide a condensed (summary version) of key matters in the privacy policy, with a link to the full policy.
- **It is not an APP 5 notice.** The APP privacy policy is not meant to be a substitute for the notice requirements under APP 5. However, it may be used to help meet requirements in some circumstances.

More information about developing an APP privacy policy, including an APP privacy policy checklist, is provided in the Guide to developing an APP privacy policy.

Risk point: Data analytics activities may make it challenging to be clear in your APP Privacy Policy about how personal information will be managed by your organisation.

Privacy tip: You do not need to describe exactly how data is processed, or any of the technical details of data analytics activities in your policy. Instead, you should aim to clearly describe the main functions and activities of your organisation, the purposes that you put information to, and how your data analytics activities relate to this.

2.2 Collection of personal information (APP 3)

APP 3 outlines when personal information, including sensitive information, may be solicited and collected by organisations. It places obligations on organisations to:

- Collect personal information only where it is reasonably necessary for, or directly related to, the organisation's functions or activities
- Collect information only by lawful and fair means
- Collect information directly from the individual, unless it is unreasonable or impractical (or another exception apples), and

- Collect sensitive information only with the individual's consent (unless an exception applies).[15]

The above requirements of APP 3 may appear to challenge the goal of some data analytics activities to repurpose data for unspecified future uses, and collecting as much data as possible. Some tips to ensure compliance with the above requirements of APP 3 are discussed below.

It is also important to remember that personal information collected by an entity may generally be used or disclosed only for the primary (original) purpose for which it was collected, unless the individual consents or another exception applies (see Use and Disclosure of Personal Information). This means the way personal information is collected, and the notice given to the individual concerned, is key when conducting data analytics, as it will in part determine the scope of how the information can be used (see Notification section).

Limiting the collection of personal information

APP 3.1 states that organisations must not collect information unless it is reasonably necessary or directly related to one or more of its functions or activities. This principle may appear to challenge the concept of using 'all the data' for 'unknown purposes'. However, just because data analytics *can* discover unexpected or 'interesting' correlations, this does not mean that the new personal information generated is necessary to the legitimate functions and activities.

Example: A company conducts data analysis on its customer database for the purposes of discovering the most relevant products and services to market to their individual customers. As part of this, it may be revealed that consumers who are divorced are more likely to have children than people who have never been married. While this may be interesting, this information may not be relevant to the company's functions or activities.

While APP 3 does place restrictions on what data may be collected, this does not need to be a barrier for data analytics. APP 3 is intended to operate objectively and practically by allowing organisations to collect personal information that is reasonably necessary (from the point of view of a reasonable person) to pursue its legitimate functions or activities.

In practice, the challenge for organisations will be to determine early in the project why they need to collect and process particular datasets. Through the use of a PIA, organisations should map what they expect to learn by processing that data, and then assess whether the personal information is relevant and not excessive, in relation its legitimate functions and activities.

To help ensure that data is relevant and not excessive, Chapter 3 of the APP Guidelines provides information on how to determine whether a particular collection of personal information is permitted.

Risk point: Using 'all the data' for 'unknown purposes' will expose organisations to privacy compliance risks**.**

Privacy tip: Limit the collection of personal information that you hold where appropriate, and ensure that you are only collecting information which is reasonably necessary to pursue your legitimate functions and activities.

[15] More information about collection is provided in Chapter 3 of the APP Guidelines.

Collection of personal information via creation

The concept of 'collects' applies broadly, and includes gathering, acquiring or obtaining personal information from any source and by any means. This includes collection by 'creation' which may occur when information is created with reference to, or generated from, other information the entity holds.

Data analytics can lead to the creation of personal information. For example, this can occur when an entity analyses a large variety of non-identifying information, and in the process of analysing the information it becomes identified or reasonably identifiable. Similarly, insights about an identified individual from data analytics may lead to the collection of new categories of personal information. This was discussed above in relation to the definition of personal information (with an example given in relation to an individual's online purchasing behaviour).

The Information Accountability Foundation has described the generation of new personal information in three categories - observed, derived and inferred:[16]

- 'Observed data' is recorded automatically, for example through online cookies or mobile apps.
- 'Derived data' is generated from an original dataset in a simple way, for example by calculating customer's preferences based on the number of items in a store that they bought.
- 'Inferred data' is produced by using a more complex method of analytics to find correlations between datasets and using these to categorise or profile people, for example by predicting future health outcomes.[17]

Example: Mobile fitness devices and apps regularly 'create' new personal information about individuals through the monitoring of heart rates and pulse, the way individuals walk or sleeping patterns. Following analysis over a period of time, the organisation is able to create new insights about an individual's likely health outcomes, including the detection and prediction of disease.

This information may be different to information which has been provided 'consciously' by the individual in a more traditional way (for example, by filling in information on a form and providing it directly to the organisation). Where an entity collects personal information 'via creation' through data analytics, they therefore need to consider whether they could have solicited and collected the personal information (APP 3.1 and 3.2). More information about collection is provided in Chapter 3 of the APP Guidelines.

If personal information is created which the organisation is not able to collect under APP 3, it will need to be de-identified or destroyed, in a way similar to what is required by APP 4.[18] It is therefore important that organisations have practices, procedures and systems for identifying and dealing with such information.

[16] Abrams, Martin. The origins of personal data and its implications for governance [PDF]. OECD, March 2014.

[17] Inferred data tends to be less accurate and may create challenges for quality of personal information.

[18] Information about how to deal with unsolicited personal information is provided in Chapter 4 of the APP guidelines.

To manage the creation of new personal information, organisations should incorporate 'privacy-by-design' and conduct a PIA. This includes identifying where data comes from, how it is created, and ensuring compliance with the APPs. A PIA may also help your organisation identify any risks associated with the use of particular algorithms or data analytic processes (for example, where they may be generating personal information that organisations are not authorised to collect).

Risk point: Data analytics may lead to the collection 'via creation' of personal information.

Privacy tip: If personal information is created which the organisation is not able to collect under APP 3, it will need to be de-identified or destroyed. A PIA can help organisations to address what personal information may be collected via creation through data analytics.

Collection of sensitive information

Sensitive information is a subset of personal information that is afforded a higher level of privacy protection under the APPs. Sensitive information includes information about a person's political opinions, religious beliefs, sexual orientation and health information.[19]

Organisations can only collect sensitive information if the individual consents[20] to the collection, unless an exception applies, such as:

- where the collection is authorised or required by law, or
- where a permitted health situation exists (discussed below).[21]

As discussed above, when conducting data analytics (for example, through the use algorithms, IoT devices, or linking of data sets) organisations may inadvertently generate sensitive information that your organisation is not authorised to collect. If an organisation inadvertently collects sensitive information it is not authorised to collect, it will need to be de-identified or destroyed.

Permitted health situation

A permitted health situation exception applies only to private sector organisations, and not to government agencies. Examples of permitted health situations include where an organisation seeks to collect health information that is necessary for research relevant to public health or public safety, and the research purpose cannot be served by collecting de-identified information (and it is impracticable to obtain the individual's consent to collecting the health information).

In this situation, an organisation can collect health information in certain circumstances, for examples where the collection is either required by law, or is in accordance with the *Guidelines approved under Section 95A of the Privacy Act 1988* (s 95A Guidelines).[22]

An organisation seeking to rely on the s 95A Guidelines must be satisfied that the research for which health information is to be collected has been approved by a Human Research Ethics

[19] The full definition of sensitive information is provided in the APP Guidelines.

[20] Consent is defined as 'express consent or implied consent' (s 6(1)). One exception to the requirement for consent to collecting sensitive information is where a permitted health situation exists. Consent is discussed in more detail in Chapter B of the APP Guidelines.

[21] See the full exceptions to collecting sensitive information in Chapter 3 of the APP Guidelines.

[22] See s 16B(2) of the *Privacy Act 1988* for the full circumstances that apply to this permitted health situation.

Committee (HREC) in accordance with the Guidelines. HRECs assess proposals to handle health information by organisations for health research (without individuals' consent). They may approve a proposed research activity where they determine that the public interest in the research activity substantially outweighs the public interest in the protection of privacy.

The other exceptions to seeking the consent of the individual to collect sensitive information are discussed in Chapter 3 of the APP Guidelines.

Collection by lawful and fair means

An entity must collect personal information by lawful and fair means (APP 3.5). Collection that would not be lawful includes collecting in breach of legislation or contrary to a court order. A 'fair means' of collecting information is one that does not involve intimidation or deception and is not unreasonably intrusive. For example, it would usually be unfair to collect personal information covertly without the knowledge of the individual. More information about collection is provided in Chapter 3 of the APP Guidelines.

Collecting information from third parties

Personal information must be collected from the individual concerned, unless this is unreasonable or impracticable (additional exceptions apply to agencies).[23]

Whether it would be 'unreasonable or impracticable' may involve considering whether the individual would reasonably expect it to be collected from another source and the sensitivity of the information collected.[24]

However, personal information collected for data analytics may come from a variety of sources. Some will be collected directly from the individual, while some will be collected from other organisations (that is, third parties).

If your organisation wishes to collect personal information from a third party, you will still need to consider whether you are authorised to collect personal information in this way. One way to do this is to consider whether the third party has been transparent with individuals and ensured that they understood, and therefore would reasonably expect, that their personal information will be collected by your organisation. In practice, you may consider the third party's APP privacy policy and relevant APP 5 notices to ensure they describe the circumstances and purpose in which the information will be disclosed and used.

[23] Government agencies may also collect personal information from someone other than the individual if the individual consents, or the agency is required or authorised by or under an Australian law, or a court/tribunal order to do so.

[24] More information about when it would be unreasonable or impracticable is provided in Chapter 3 of the APP Guidelines.

Risk point: Personal information used in data analytics is likely to include information collected from third parties.

Privacy tip: Before collecting personal information from another organisation for data analytics, you need to ensure that you are authorised to do so. One way to do this is to consider whether the original privacy notice given to the individuals by the third party covers this further use and disclosure of their data.

Example: A retail company is considering collecting personal information from a third party organisation for the purpose of identifying general trends for advertising. To ensure that it is compliant with its collection obligations, it considers whether it would be unreasonable or impracticable to collect the information directly from the individual. In making this assessment, the company considers a number of factors including whether the individual would reasonably expect personal information about them to be collected by the organisation for this purpose. The retail company consults the third party's privacy policy and notices, which clearly state that it provides personal information to external parties for advertising purposes. The company therefore considers that an individual would reasonably expect for their information to be collected by them for this purpose.

2.3 Notification (APP 5)

When your organisation collects personal information, APP 5 requires that reasonable steps be taken to either notify the individual of certain matters, or to ensure the individual is aware of those matters. These matters include, for example:

- the fact and circumstances of the collection, and
- the purposes of collection.[25]

An APP entity must take these steps before or at the time it collects the information. If this is not practicable, reasonable steps must be taken as soon as practicable after collection.

However, providing notice effectively can be challenging for data analytics. For example, many people do not read privacy notices, particularly when they are long, and data may also be collected through observation, rather than through a specific transaction. Nevertheless, organisations still need to give individuals notification of the collection of their data.

Privacy notices therefore have a big job to do. They need to communicate information handling practices clearly and simply, but also comprehensively and with enough specificity to be meaningful.

While implementing these regulatory requirements in data analytics settings can be challenging, new technologies are also enabling opportunities to provide more dynamic, multi-layered and

[25] Information about the other matters to be notified is provided in Chapter 5 of the APP Guidelines.

user centric privacy notices. Innovative approaches to privacy notices can include 'just-in-time' notices, video notices and privacy dashboards.[26]

See the section above on Collection of personal information, for information on the interaction between an APP privacy policy and APP 5 notice.

How can the purposes of collection be communicated in a data analytics context?

Organisations may generally not use personal information for a purpose other than the primary purpose it was collected for (i.e. for a secondary purpose), unless an exception applies. Exceptions include where the individual has consented or the individual would reasonably expect the other related use (see below for more information about using personal information). If an entity plans to use or disclose personal information for purposes other than the primary purpose (known as a 'secondary purpose') these should also be included in the privacy notice.

Organisations will need to provide adequate information about the collection and potential uses while ensuring the notice does not become overly vague or overly detailed.

Where multiple uses are included in a notice, organisations should consider whether individuals have the opportunity to choose which collections, uses and disclosures they agree to and which they do not. PIAs are useful for informing the content of notices.

For example, the notice may set out:

- that analytics may be conducted on personal information for the purposes of determining which products or services your customers may be interested in
- that analytics may be conducted using information from a range of sources, such as information collected from third parties (and a description or list of those sources)
- any anticipated secondary purposes that data may be put to (for example, your personal information may be subject to data analytics which seeks to determine the likelihood of certain events occurring, which may be relevant to the service provided to the individual), and/or
- any anticipated disclosures of personal information to third parties (and a description or list of those entities).

If the organisation does identify new purposes that it wants to use personal information for, it should communicate this to individuals as soon as possible (or alternatively, de-identify the data).

More information about the use of personal information, based on the notified purposes of collection, is provided below. While more information about the other specific matters that need to be notified is provided in Chapter 5 of the APP Guidelines.

What does taking 'reasonable steps' to notify an individual involve?

An organisation must take reasonable steps to notify an individual under APP 5 or ensure the individual is aware of the APP 5 matters. What are reasonable steps for an entity will depend upon circumstances including:

[26] 'Just-in-time notices work by appearing on the individual's screen at the point where they input personal data, providing a brief message explaining how the information they are about to provide will be used.'

- the sensitivity of the personal information collected (more rigorous steps may be required when collecting 'sensitive information')
- the possible adverse consequences for an individual as a result of the collection
- any special needs of the individual (for example, whether the individual may have a visual or other impairment), and
- the practicability, including the time and costs involved. However, an entity is not excused from taking particular steps by reason only that it would be inconvenient, time-consuming or impose some cost.

In some situations an entity may decide that not taking any steps is reasonable in the circumstances. For example, this may be the case when an individual already knows the APP 5 matters because the personal information is collected from them regularly by the entity.

More information about what might be reasonable circumstances, and when it may be appropriate to take no steps, is provided in Chapter 5 of the APP Guidelines.

Examples — how should notice be given?

When designing a privacy notice, your organisation should take the individual's perspective into account to ensure that it is user-centric and that the information provided is generally understandable from the point of view of the target audience.

Where possible, privacy notices should be multi-layered to assist with readability and navigability. This means that brief notices are provided which are supplemented by longer notices. The timing of notices can also occur more dynamically to ensure information is given in context, at the right time, in a way that is easy to read. It may also be helpful for organisations to consider consulting with users and seeking their input when designing notices, or pilot testing or using focus groups to ensure that individuals understand the content.

Organisations may include information about how, when and from where the personal information was collected. These details are particularly important when the entity collects the personal information:

- from a third party
- where the personal information is collected via creation, and
- where the individual may not be aware that their personal information was collected.

An individual may be notified or informed of APP 5 matters through a variety of formats. The way the personal information is collected, or later used, may suggest a particular form for the privacy notice. This includes whether the information is being collected directly from the individual or from a third party. How the personal information is collected (whether over the phone, by completing online forms, attending shopfronts, or through cookies) also impacts on how the notice may be given.

Organisations should use a PIA to consider how best to give notice of collection and the purpose of collection, especially for secondary uses. A PIA can consider the information lifecycle and help identify what information will be needed for which functions and activities of the entity. It can also identify how the personal information will be collected.

Example

A telecommunications company is preparing a privacy notice to let individuals know that it will be sharing their information with third parties in some situations, including for the purposes of conducting data analytic projects.

A good example of a privacy notice is one which clearly and simply informs individuals about the purposes their personal information will be put to, the reasons for these planned uses, and the choices available to the individual. It will present the information effectively, for example by using graphics/ colours to draw the individual's attention to particular aspects of the notice. The notification may also provide a genuine opportunity for the person to either agree to particular uses of their information, or to opt-out of particular uses. The individuals may also be provided with a convenient way to change their preferences at any time in the future through a 'privacy dashboard'.

A poor example of notification would involve the use of confusing and legalistic language, inclusion of a lot of unnecessary information, without including information of most relevance to individuals.

Risk point: Research shows many people don't read privacy notices.

Privacy tip: Organisations should use privacy impact assessments to inform what information to include in their notices and then provide it in easy to read, dynamic and user centric ways. For example, organisations may consider 'just-in-time' notices, video notices and privacy dashboards.

2.4 Using and disclosing personal information (APP 6)

APP 6 outlines when an entity may use or disclose personal information. It provides that personal information may only be used or disclosed for the purpose for which it was collected (known as the 'primary purpose'), or for a secondary purpose if an exception applies. This principle may appear to present a challenge when conducting data analytics, as the ability to analyse data for different purposes is an important characteristic (and benefit).

In practice, your organisation will need to be able to determine whether the uses and disclosures of personal information to a third party are compatible with the original purpose it was collected for, and the privacy policy and/or notice given to the individual.

If the use or disclosure of personal information is not compatible with the primary purpose, you will need to rely on one of the exceptions set out in the APP 6 Guidelines.[27] The most common exceptions for the secondary use of personal information for data analytics include, where:

- the individual has consented

[27] The other exceptions are discussed in Chapter 6 of the APP Guidelines.

- the individual would reasonably expect the entity to use or disclose their personal information for the secondary purpose (and that purpose is related or directly related to the primary purpose of collection), or
- a permitted health situation exists.

You may also choose to update your privacy policy and notices accordingly, ensuring that people are aware of likely secondary uses and disclosures of personal information (including data analytics projects). This may help to establish that an individual would likely expect the use or disclosure, or in some cases help to establish that an individual has provided informed consented to the use or disclosure of their information for a secondary purpose.[28] Organisations should also consider how they might allow individuals to genuinely choose which uses and disclosures they agree to and which they do not.

Risk point: Secondary uses and disclosures of personal information are common in data analytics.

Privacy tip: Organisations should carefully consider whether uses and disclosures of personal information for data analytics activities are compatible with the original purpose of collection (particularly when the information is collected directly from a third party). If not, organisations will need to rely on one of the exceptions in APP 6.

Privacy tip: When using privacy notices to inform individuals about a particular use or disclosure, organisations should consider how they might allow individuals to choose which uses and disclosures they agree to and which they do not.

Common APP 6 exceptions

The individual has consented

An entity may use or disclose personal information for a secondary purpose where the individual has consented to that use or disclosure. Consent is defined as 'express consent or implied consent' (s 6(1)). The four key elements of consent are discussed in Chapter B of the APP Guidelines. The four key elements of consent are discussed in Chapter B of the APP Guidelines.

The individual would reasonably expect the use or disclosure

An entity may use or disclose personal information for a secondary purpose if the individual would reasonably expect the entity to use or disclose the information for that secondary purpose, and:

- if the information is sensitive information, the secondary purpose is directly related to the primary purpose of collection, or
- if the information is not sensitive information, the secondary purpose is related to the primary purpose of collection.

[28] However, generally, it should not be assumed that an individual has given consent on the basis alone that they did not object to a proposal to handle personal information in a particular way. An organisation cannot infer consent simply because it provided an individual with notice of a proposed collection, use or disclosure of personal information. For further information, s B.36-B.42 of Chapter B: Key Concepts of the APP Guidelines.

This exception creates a two-limb test which focuses both on the reasonable expectations of the individual, and the relationship between the primary and secondary purposes.

The 'reasonably expects' test is an objective one that has regard to what a reasonable person, who is properly informed, would expect in the circumstances. This is a question of fact in each individual case. It is the responsibility of the entity to be able to justify its conduct. The relationship between the primary and secondary purpose must be more than a tenuous link.

More information about the meaning of 'reasonably expects' and the relationship between the primary and secondary purpose is provided in Chapter 6 of the APP Guidelines.

Example: When an individual signs up for a loyalty card which records all relevant transactions they make, in exchange for certain discounts or other offers, there would likely be a reasonable expectation that the company will be using this data to gain a better understanding of their customers' spending behaviour and using this information for marketing purposes. However, it may not be within reasonable expectations for the same company to track its customers' movements through analysis of their mobile phone data.

Permitted health situation

An entity may wish to use personal information for the secondary purpose of research. For organisations, the relevant APP 6 exception is where a permitted health situation exists. An organisation may use or disclose health information that is necessary for the secondary purpose of research relevant to public health or public safety if:

- it is impracticable to get the individual's consent
- the use or disclosure is conducted in accordance with the s 95A Guidelines approved by the Information Commissioner, and
- for disclosure, the organisation reasonably believes the recipient will not disclose the information, or personal information derived from the information.[29]

Whether it is impracticable to seek consent will depend on the particular circumstances of the case. An organisation relying on this permitted health situation will need to justify why it is impracticable to obtain an individual's consent. Incurring some expense or doing extra work to obtain consent would not itself make it impracticable to obtain consent.

An organisation seeking to rely on the section 95A Guidelines must be satisfied that the research for which health information is to be used or disclosed has been approved by a Human Research Ethics Committee (HREC) in accordance with the Guidelines. They may approve a proposed research activity where they determine that the public interest in the research activity substantially outweighs the public interest in the protection of privacy.

Agencies seeking to handle personal information for medical research in a way that may be inconsistent with the APPs should refer to s 95 of the Privacy Act.

[29] See section 16B(3) of the *Privacy Act 1988*.

Section 95 permits acts that would otherwise breach the APPs where those acts are done in the course of medical research and in accordance with *the Guidelines under Section 95 of the Privacy Act 1988* (s 95 Guidelines).

The s 95 Guidelines provide a framework for the conduct of medical research using information held or collected by agencies. An agency seeking to rely on the s 95 Guidelines must be satisfied that the research for which the personal information is to be handled has been approved by an HREC for the particular purpose in accordance with the Guidelines. In making a decision under these Guidelines, a HREC must consider whether it is reasonable for the research to proceed without the consent of the individuals to whom the information relates. In addition, the proposed handling of information must be done in the course of medical research.

More information about use and disclosure is provided in Chapter 6 of the APP Guidelines.

Risk point: Where health or personal information is being handled for data analytics activities it may be impracticable to obtain individuals' consent.

Privacy tip: Entities undertaking health or medical research should ensure they are familiar with the s 95 or s 95A Guidelines.

2.5 Direct marketing (APP 7)

Direct marketing is where an organisation directly promotes goods or services to an individual, and can encompass any communication made by or on behalf of an organisation to an individual.

One of the key purposes of data analytics is to assist organisations to improve their marketing strategies. Where organisations use or disclose individuals' personal information to tailor the direct marketing communications (such as online advertisements) they send to and target at those individuals, they should consider the requirements of APP 7.

APP 7 sets out when and how organisations can use and disclose personal information they hold for direct marketing purposes. Organisations can use and disclose personal information for direct marketing if:

- the organisation collected the personal information directly from the individual and the individual would reasonably expect their personal information to be used or disclosed for direct marketing
- the individual has consented to their personal information being used or disclosed for direct marketing, or
- it is impractical to get the individual's consent to their personal information being used or disclosed for direct marketing.

However, individuals' sensitive information can only be used and disclosed for direct marketing if the individual has given their consent. Specific requirements for direct marketing are set out under APP 7. This includes providing individuals with a simple means of opting out of future direct marketing communications and stopping their direct marketing where an individual asks them to stop.

Depending on the type of direct marketing communications organisations use to direct market to individuals, they may have other obligations that apply to their direct marketing communications, including the *Spam Act 2003* or the *Do Not Call Register Act 2006*.

It is also important to note that organisations that facilitate other organisations' direct marketing (such as data list brokers) also have specific obligations under APP 7. This includes no longer using or disclosing individuals' personal information where an individual has asked them to stop.

More information about APP 7, including the specific requirements, is provided in Chapter 7 of the APP Guidelines.

Risk point: Data analytics activities are often undertaken for the purposes of direct marketing. Organisations that *facilitate* other organisations' direct marketing have additional obligations under APP 7.

Privacy tip: Organisations should have a good understanding about how they use data analytics for direct marketing, and if this includes facilitating other organisations' direct marketing, they need to comply with additional obligations.

Using and disclosing personal information for direct marketing

Due to the high volume of data organisations may collect for data analytics to inform direct marketing, and the range of information sources they may use, organisations should:

- ensure they put in place monitoring processes to identify the types of information they are collecting. This will reduce the risk of using or disclosing sensitive information for direct marketing purposes without individuals' consent
- consider individuals' expectations about how their information will be used and disclosed in light of the original purposes for which their information was collected and any notices they were provided, and/or
- think about how to implement simple and effective ways that individuals can use to opt out of receiving direct marketing communications, or ask that their information is no longer used and disclosed for direct marketing purposes.

Organisations should also be mindful that even where APP 7 may not prevent them from using or disclosing customers' personal information for particular direct marketing purposes, it is still important to build a good relationship with their customers based on transparency and trust.

Having good privacy practices generally (as outlined earlier in this guide) will assist in building trust and transparency, and avoid creepy behaviour. For example, by having clear privacy policies and notifications systems, allowing opt-outs for certain collection or uses of information, and conducting PIAs with a community consultation focus (so that new ideas can be tested before they are implemented).

Organisations should also stay up to date with relevant media sources, particularly when data breaches or privacy incidents occur to get a sense of the community's attitudes to privacy. All of these activities will help organisations to predict what individuals want and expect in terms of the management and use of their personal information.

Risk point: The 2017 Community Attitudes to Privacy survey found that the majority of Australians are annoyed when they receive unsolicited marketing.

Privacy tip: Organisations should be transparent with their customers by explaining that their data is being collected, how and why their interests are being protected and giving them a choice. It is also important to think about the experience of the customer by considering whether the activities will be perceived as 'creepy', unexpected or harmful. [30]

Case study: Target developed an algorithm which could predict pregnancy in its customers, based only on which goods they bought. Target then sent coupons for "pregnancy products" to these customers. In one case, the distribution of such coupons to a family home revealed a young woman's pregnancy (her health information) to the rest of her family.[31]

Using and disclosing personal information to facilitate direct marketing

There are a number of organisations that collect and analyse personal information on behalf of other organisations, or on-sell that information to organisations for use in their direct marketing activities.

It may be difficult to keep track of each individual's personal information. However, these organisations should be aware that individuals can ask that they stop using or disclosing their personal information to facilitate the direct marketing of other entities under APP 7, and consider implementing systems that will enable them to more easily meet this obligation.

Risk point: It is common for third parties to collect and analyse personal information on behalf of other organisations, or on-sell that information to organisations for use in their direct marketing activities.

Privacy tip: Ensure that your organisation provides clear 'opt-outs' and meets its other obligations under APP 7 when engaging in direct marketing, or when facilitating direct marketing for other organisations.

2.6 Sending personal information overseas (APP 8)

APP 8 and s 16C of the Privacy Act apply when an entity discloses personal information overseas.[32] An organisation is considered to 'disclose' personal information where it provides the information to an overseas recipient, but does not maintain control over how that information is subsequently handled by the recipient.

---

[30] See 'Best Practice Guideline: Big Data' [PDF], Association for Data-driven Marketing & Advertising.

[31] Charles Duhigg, 16 February 2012, How companies learn your secrets, The New York Times Magazine

[32] APP 8 does not apply where the entity maintains effective control over the information so that it would be considered to be 'using' rather than 'disclosing' the information. Instead, an entity 'using' personal information overseas will be accountable for its information handling under the APPs that apply to 'use'. In practice, the steps that an APP entity is required to take and their accountability when sending personal information overseas can be similar regardless of whether the information is being used or disclosed.

APP 8.1 provides that, subject to certain exceptions set out in APP 8.2,[33] before an entity discloses personal information about an individual to an overseas recipient, the entity must take reasonable steps to ensure that the recipient does not breach the APPs in relation to that information.

These provisions are intended to enable individuals to seek redress in Australia, if their information is mishandled overseas.

Data analytics often involve the use of overseas cloud (or internet) based platforms. The APPs do not prevent the sending of personal information overseas or engaging an overseas cloud service provider. However, entities will need to carefully consider steps that may need to be taken to ensure compliance with the APPs.

Before an entity uses an overseas cloud service to store its data, or employs an overseas based cloud platform to perform data analytics, it should consider whether it can achieve its aims using de-identified information. De-identification is discussed in Part One.

Where it is necessary to disclose personal information overseas, an entity is required to take reasonable steps to ensure that the overseas recipient does not breach the APPs. Information about how to comply with APP 8 when sending information overseas is provided in our Privacy business resource 8: Sending personal information overseas and Privacy agency resource 4: Sending personal information overseas.

Where entities are likely to disclose personal information to overseas recipients they should include information about that disclosure in their Privacy Policy and Privacy Notice.

Risk point: Where an organisation discloses personal information to an overseas recipient (unless an exception to APP 8 applies) it will be accountable for an act or practice of the overseas recipient that would breach the APPs. This is the case even if the organisation has taken reasonable steps under APP 8.1.

Privacy tip: Entities should undertake due diligence before disclosing personal information to overseas recipients. This will help them identify risks and take steps to mitigate them.

2.7 Quality of personal information (APP 10)

Entities must take reasonable steps to ensure that the personal information they collect is accurate, up-to-date and complete (APP 10.1). Similarly, entities must take reasonable steps to ensure that the personal information it uses or discloses having regard to the purpose of the use or disclosure is accurate, up-to-date, complete and relevant. Guidance about the meaning of the terms 'accurate', 'up-to-date', 'complete' and 'relevant' is provided in Chapter 10 of the APP Guidelines.

Large scale data analytics may appear to present some challenges to the principles of accuracy and relevance of data. For example, these activities typically seek to collect large amounts of data from many diverse sources, with little opportunity to verify the relevance or accuracy of the information. Further, some data analytics techniques such as automatic algorithms have the

[33] See Chapter 8 of the APP Guidelines.

potential to create personal information with an inherent bias, that are discriminatory or that lead to erroneous or unjustified results.

Ensuring accuracy and quality in data analytics is particularly important where information may be used to make decisions about an individual, such as an administrative decision by a government agency. In these situations, it would be prudent for organise to take additional and more rigorous steps to ensure the quality of both the personal information collected, as well as any additional personal information created by the algorithms that process the data. Some examples of reasonable steps are set out below.

Risk point: Data analytics techniques such as automatic algorithms have the potential to make decisions that are discriminatory, erroneous and unjustified.

Privacy tip: Consider conducting regular reviews of your data analytic processes (such as algorithms used), to ensure that they are fit for purpose and promote the accuracy of information.

Example: A recruitment company conducts analytics on candidate data with the aim of identifying and hiring the most suitable candidate for an available role. However, where this involves automated decision-making, the organisation should ensure that the information used to do this (including inferences drawn from data analytics) is accurate.

Examples of reasonable steps to ensure quality of personal information

The steps which will be reasonable will depend on the specific circumstances of each case and are discussed in Chapter 10 of the APP Guidelines. Examples of steps which may be appropriate to take include:

- Where possible and appropriate, verifying the accuracy of information which is not collected directly from the individual. For example, checking that third parties, from which personal information is collected, have implemented appropriate practices, procedures and systems to ensure the quality of personal information. It may also be useful to put in place procedures to monitor and record what type of personal information you are collecting.
- Putting in place systems, including auditing and reviews, to check that the analytic processes used (such as algorithms) are operating appropriately and are fit for purpose, and not creating biased, inaccurate, discriminatory, or unjustified results
- Be as transparent as possible about the purpose of your organisation's analytic techniques (including algorithms), to better help individuals understand why recommendations or decisions have been made about them. An internal document may be more appropriate for commercially sensitive techniques.
- Ensuring that your organisation has taken reasonable steps to ensure that individuals may seek access to their personal information (APP 12) and seek to correct this information (APP 13).[34]

[34] Guidance about how individuals may seek access to the personal information entities hold about them and how they can seek correction of that information in certain circumstances is provided in Chapters 12 and 13 of the APP Guidelines.

Further discussion about the typical steps entities take is provided in Chapter 10 of the APP Guidelines.

Risk point: Where an organisations collects personal information from a third party and not directly from the individual, there may be a higher risk that the information may not be accurate, complete and up-to-date.

Privacy tip: Organisations may need to take more rigorous steps to ensure the personal information collected via creation is accurate, complete and up-to-date. For example, by checking that third parties, from which personal information is collected, have implemented appropriate practices, procedures and systems to ensure the quality of personal information.

2.8 Security of personal information (APP 11)

APP 11 requires entities to actively consider whether they are permitted to retain personal information. When entities retain personal information, they must take reasonable steps to protect it from misuse, interference and loss, as well as unauthorised access, modification or disclosure.

Guidance on the terms 'misuse', 'interference', 'loss', 'unauthorised access', 'unauthorised modification' and 'unauthorised disclosure' is provided in Chapter 11 of the APP Guidelines.

Retaining personal information

APP 11 requires entities to actively consider whether they are permitted to retain personal information. When an entity no longer needs personal information for any purpose for which it may be used or disclosed under the APPs (and if the information is not contained in a Commonwealth record or legally required to be retained by the entity) the entity should destroy or de-identify the information. An entity must take reasonable steps to destroy or de-identify the personal information. De-identification is discussed in Part 1.

Where personal information is retained, entities should be able to justify their retention of the personal information. More information about the retention of personal information is provided in Chapter 11 of the APP Guidelines.

Risk point: Entities can only keep personal information they need for permitted purposes under the APPs. The onus is on entities to justify their retention of personal information.

Privacy tip: Entities need to be able to justify why they have retained personal information and for what permitted purposes. Entities can also consider de-identifying personal information so they can keep the data for future uses. A PIA can be a useful tool for this purpose.

Reasonable steps for ensuring security of information for data analytics

Entities that engage in data analytics often hold larger amounts of data and for longer periods of time. Entities need to consider what security risks exist and take reasonable steps to protect the personal information they hold. This includes internal and external risks.

For example, holding larger amounts of personal information for longer may increase the risk of unauthorised access by staff or contractors. While 'honey pots' containing vast amounts of valuable data may increase the risk that an entity's information systems may be hacked.

It is expected that entities handling large amounts of personal information for data analytics purposes will conduct an information security risk assessment (also known as a threat risk assessment) as part of undertaking a PIA. This will enable the entities to identify and evaluate security risks, including threats and vulnerabilities, and the potential impacts of these risks to personal information.

Undertaking an information security risk assessment will assist the entity to identify reasonable steps to take to protect personal information. This would generally include (but not be limited to):

- limiting internal access to personal information to those who require access to do their job (i.e. providing access on a 'need to know' basis, and conduct regular reviews or audits of those with access)
- maintaining a chronological record of system activities, such as an audit log, by both internal and external users, for reviewing activity on an ICT system to detect and investigate privacy incidents
- implementing network security measures, such as intrusion prevention and detection systems to identifying and responding to known attack profiles, and network segregation to reduce cases of unauthorised access
- updating patches, automatic and disable unnecessary add-ons and programs from computers
- undertaking penetration testing of enterprise data warehouses to identify internal and external vulnerabilities to mitigate against
- email filtration, anti-virus and anti-malware if data sets can be accessed on devices with web access
- utilising encryption to ensure that information is stored in a form that cannot be easily understood by unauthorised individuals or entities
- ensuring reasonable steps are taken to destroy or de-identify the personal information when it is no longer needed, and/or
- in the event of a data breach, having a response plan that includes procedures and clear lines of authority which can assist an entity to contain the breach and manage their response. Our *Guide to managing data breaches in accordance with the Privacy Act 1988 (Cth)* provides guidance for organisations when responding to a data breach involving personal information.[35]

More information about reasonable steps, including further examples of what may be reasonable steps, is provided in the *Guide to securing personal information*.

[35] The passage of the *Privacy Amendment (Notifiable Data Breaches) Act 2017* established a Notifiable Data Breaches (NDB) scheme in Australia. The NDB scheme requires organisations covered by the Privacy Act to notify any individuals likely to be at risk of serious harm by a data breach. See the OAIC's NDB web page for information about notifying individuals about an eligible data breach.

Risk point: 'Honey pots' of valuable and sensitive personal information may be targets for hacking.

Privacy tip: Undertaking an information security risk assessment will enable entities to identify reasonable steps to take to protect the personal information they hold.

Attachment 1: Privacy tips and risk points when conducting data analytics

Personal information

Risk point: Data used for data analytics may include personal information, and the activities will therefore be subject to the Privacy Act.

Privacy tip: Successfully de-identified data is not personal information meaning the Privacy Act will generally not apply.

Re-identification of personal information

Risk point: Where de-identification is not done properly, data analytics activities may lead to re-identification of personal information.

Privacy tip: Undertake a risk assessment to consider the likelihood of re-identification. Use appropriate de-identification techniques and implement risk mitigation strategies.

Open and transparent management of personal information (APP1)

Risk point: Privacy Impact Assessments can be more challenging for large scale data analytics projects (such as big data activities), as an organisation may not know exactly how it is going to use the data, or what data it will use during the initial 'discovery phase'

Privacy tip: Even if the direction of a data analytics project seems unclear, err on the side of caution and begin the PIA process anyway. It is important that a PIA is treated as an iterative process, which continues to develop. As a project evolves, the potential privacy risks will become clearer and your organisation will be able to better address them.

Risk point: Data analytics activities may make it challenging to be clear in your APP Privacy Policy about how personal information will be managed by your organisation.

Privacy tip: You do not need to describe exactly how data is processed, or any of the technical details of data analytics activities in your policy. Instead, you should aim to clearly describe the main functions and activities of your organisation, the purposes that you put information to, and how your data analytics activities relate to this.

Collecting personal information (APP 3)

Risk point: Using 'all the data' for 'unknown purposes' will expose entities to privacy compliance risks.

Privacy tip: Limit the collection of personal information that you hold where appropriate, and ensure that you are only collecting information which is reasonably necessary to pursue your legitimate functions and activities. A privacy impact assessment is a useful tool for this process.

Risk point: Data analytics may lead to the collection 'via creation' of personal information.

Privacy tip: If personal information is created which the organisation is not able to collect under APP 3, it may need to be de-identified or destroyed.

Risk point: Personal information used in data analytics activities is likely to include information collected from third parties.

Privacy tip: Before collecting personal information from another organisation for data analytics activities, you need to ensure that you are authorised to do so. One way to do this is to consider whether the original privacy notice given to the individuals by the third party covers this further use and disclosure of their data.

Notification of collection of personal information (APP 5)

Risk point: Research shows many people don't read privacy notices.

Privacy tip: Entities should use privacy impact assessments to inform what information to include in their notices and then provide it in easy to read, dynamic and user centric ways. For example, your organisation may consider 'just-in-time' notices, video notices and privacy dashboards.

Use and disclosure of personal information (APP 6)

Risk point: Secondary uses and disclosures are common in data analytics activities.

Privacy tip: Organisations should carefully consider whether uses and disclosures of personal information for data analytics activities are compatible with the original purpose of collection (particularly when the information is collected directly from a third party. If not, organisations will need to rely on one of the exceptions in APP 6.

Privacy tip: When using privacy notices to inform individuals about a secondary use or disclosure, entities should consider how they might allow individuals to choose which uses and disclosures they agree to and which they do not.

Risk point: Where health or personal information is being handled for data analytics activities it may be impracticable to obtain individuals' consent.

Privacy tip: Entities undertaking health or medical research should ensure they are familiar with the s 95 or s 95A Guidelines.

Direct marketing (APP 7)

Risk point: Data analytics activities are often undertaken for the purposes of direct marketing. Organisations that facilitate other organisations' direct marketing have additional obligations under APP 7.

Privacy tip: Organisations should have a good understanding about how they use data analytics for direct marketing, and if this includes facilitating other organisations' direct marketing, they need to comply with additional obligations.

Risk point: The 2017 Community Attitudes to Privacy survey found that the majority of Australians are annoyed when they receive unsolicited marketing.

Privacy tip: Organisations should be transparent with their customers by explaining that their data is being collected, how and why their interests are being protected and giving them a choice. It is also important to think about the experience of the customer by considering whether the activities will be perceived as 'creepy', unexpected or harmful.

Risk point: It is common for third parties to collect and analyse personal information on behalf of other organisations, or on-sell that information to organisations for use in their direct marketing activities.

Privacy tip: Ensure that your organisation provides clear 'opt-outs' and meets its other obligations under APP 7 when engaging in direct marketing, or when facilitating direct marketing for other organisations.

Sending personal information overseas (APP 8)

Risk point: Where an organisation discloses personal information to an overseas recipient (unless an exception to APP 8 applies) it will be accountable for an act or practice of the overseas recipient that would breach the APPs. This is the case even if the organisation has taken reasonable steps under APP 8.1.

Privacy tip: Entities should undertake due diligence before disclosing personal information to overseas recipients. This will help them identify risks and take steps to mitigate them.

Quality of personal information (APP 10)

Risk point: Data analytics techniques such as automatic algorithms have the potential to make decisions that are discriminatory, erroneous and unjustified.

Privacy tip: Consider conducting regular reviews of your data analytic processes (such as algorithms used), to ensure that they are fit for purpose and promote the accuracy of information.

Risk point: Where an organisation collects personal information from a third party and not directly from the individual, there may be a higher risk that the information may not be accurate, complete and up-to-date.

Privacy tip: Organisations may need to take more rigorous steps to ensure the personal information collected via creation is accurate, complete and up-to-date. For example, by checking that third parties, from which personal information is collected, have implemented appropriate practices, procedures and systems to ensure the quality of personal information.

Security of personal information (APP 11)

Risk point: Entities can only keep personal information they need for permitted purposes under the APPs. The onus is on entities to justify their retention of personal information.

Privacy tip: Entities need to be able to justify why they have retained personal information and for what permitted purposes. Entities can also consider de-identifying personal information so they can keep the data for future uses. A privacy impact assessment can be a useful tool for this purpose.

More information

De-identification Decision-Making Framework

The OAIC and CSIRO's Data61 have released the *De-Identification Decision-Making Framework* to assist organisations to de-identify their data effectively. The *De-Identification Decision-Making Framework* is a practical and accessible guide for Australian organisations that handle personal information and are considering sharing or releasing it to meet their ethical responsibilities and legal obligations. The guide is an adaptation of the existing UK version, the *Anonymisation Decision-Making Framework* and is consistent with the OAIC's De-identification guidance, which should be read together with this resource.

Guide to De-identification and the Privacy Act

Provides general advice about de-identification, to assist entities to protect privacy when using or sharing information containing personal information. Guidance is provided on when de-identification may be appropriate, how to choose appropriate de-identification techniques, and how to assess the risk of re-identification.

Guide to 'What is personal information'

This resource aims to assist entities to understand and apply the definition of 'personal information' in section 6(1) of the Act. Specifically, this guidance aims to take you through the factors that you may wish to consider when determining whether information is personal information.

Guide to undertaking privacy impact assessments

The *Guide to undertaking privacy impact assessments* provides assistance to entities on designing, conducting and acting on a privacy impact assessment.

Privacy Impact Assessment eLearning

The eLearning program complements the *Guide to undertaking privacy impact assessments*, and aims to give you information on conducting a PIA in an easy-to-understand format so that you can have the confidence to do a PIA in your organisation or agency.

Guide to developing an APP privacy policy

APP 1 requires entities to have a clearly expressed and up-to-date privacy policy describing how they manage personal information. An APP privacy policy is a key document to ensure personal information is managed in an open and transparent way.

Guide to securing personal information

Provides guidance on the reasonable steps entities are required to take under the Privacy Act to protect the personal information they hold from misuse, interference, loss, and from unauthorised access, modification or disclosure. It also includes information about the reasonable steps entitles are required to take to destroy or de-identify personal information they hold once it is no longer needed (unless an exception applies).

Data breach preparation and response — A guide to managing data breaches in accordance with the Privacy Act 1988 (Cth)

Assists Australian Government agencies and private sector organisations prepare for and respond to data breaches in line with their obligations under the *Privacy Act 1988*.

Guidelines on Data Matching in Australian Government Administration

Aims to assist Australian Government agencies to use data matching as an administrative tool in a way that complies with the APPs and the Privacy Act, and is consistent with good privacy practice.

The examples provided in this resource are for illustrative purposes only. APP entities will need to consider how the Privacy Act applies to their particular situation.