



# **Response to the Productivity Commission's Draft Report on *Data Availability and Use***

December 12, 2016

## 1. Introduction

Envestnet Yodlee (“Yodlee”) appreciates this opportunity to respond to the Productivity Commission's Draft Report on "Data Availability and Use" (October 2016) (**Draft Report**). We wish to share our perspective regarding the opportunities, challenges and risks of data sharing. Yodlee operates a global platform that enables *consumer-permissioned financial account data sharing*. We have more than 1,000 commercial customers and, through them, over 20M active users on our Platform. We have been in business for over 16 years and operate in 17 countries, including Australia, many of which are engaged in similar Open Data initiatives.

Yodlee provides our account aggregation capabilities via both a hosted solution and open APIs on a business-to-business basis. Our customers are traditional financial institutions and financial technology (FinTech) companies. Our customers’ services include financial wellness solutions, which provide a single platform for consumers to track, manage, and improve their finance health across a host of different banks and financial institutions; real time account holder identity verification and in-line affordability checks required by responsible lending. Yodlee’s customers include top global banks in more than 20 countries as well as leading financial services innovators. *It is this perspective which is the context of our response.*

Our submission is comprised of our responses to:

- the Draft Recommendations in the Draft Report (at pages 3-7 of this submission);
- the Draft Findings in the Draft Report (at page 8 of this submission); and
- the Requests for Information (at pages 9-13) of this submission).

If you require any additional information about any aspect of our submission, please contact Chad Wiechers at [CWiechers@yodlee.com](mailto:CWiechers@yodlee.com).

## 2. Response to Draft Recommendations

### *Draft Recommendation 5.1*

#### DRAFT RECOMMENDATION 5.1

In conjunction with the Australian Bureau of Statistics and other agencies with data de-identification expertise, the Office of the Australian Information Commissioner should develop and publish practical guidance on best practice de-identification processes.

To increase confidence in data de-identification, the Office of the Australian Information Commissioner should be afforded the power to certify, at its discretion, when entities are using best practice de-identification processes.

**Yodlee agrees with this draft recommendation.** Standards and guidelines regarding de-identification processes are essential to build trust, and consumers will be more willing to authorise data sharing when they trust that their information will not be re-identified.

### *Draft Recommendation 6.2*

#### DRAFT RECOMMENDATION 6.2

The private sector is likely to be best placed to determine sector-specific standards for its data sharing between firms, where required by reforms proposed under the new data Framework.

In the event that voluntary approaches to determining standards and data quality do not emerge or adequately enable data access and transfer (including where sought by consumers), governments should facilitate this, when deemed to be in the public interest to do so.

**In the context of consumer financial data, Yodlee disagrees with the draft recommendation that sector-specific standards are best left for private sector participants to determine.** Leaving development to the parties directly involved would allow the largest financial institutions to dominate the standard-setting decisions because, as the dominant source of consumer financial data, the financial institutions can force third parties into accepting the types and formats of data that the financial institutions prefer. Without government intervention, the standards created will be biased in favour

of the major market participants, rather than objectively favouring the interests of consumers, who will greatly benefit from increased access to financial data. For example, financial institutions have an interest in limiting the extent of data-sharing because, as the Productivity Commission mentions, a financial institution can derive a competitive advantage from the data generated through providing financial services to consumers.

Yodlee recommends that the government follow the approach taken in the United Kingdom by establishing a working group to develop standards to facilitate the secure sharing of consumer financial data. The working group's membership should be representative of the diverse stakeholders in the financial data-sharing ecosystem, which includes financial institutions, aggregators, service providers, and consumer advocacy groups.

### ***Draft Recommendation 9.1***

#### DRAFT RECOMMENDATION 9.1

The Australian Government should introduce a definition of consumer data that includes:

- personal information, as defined in the Privacy Act 1988 (Cth)
- all files posted online by the consumer
- all data derived from consumers' online transactions or Internet-connected activity
- other data associated with transactions or activity that is relevant to the transfer of data to a nominated third party.

Data that is transformed to a significant extent, such that it is demonstrably not able to be re-identified as being related to an individual, should not, for the purposes of defining and implementing any Comprehensive Right, be defined as consumer data.

The definition of 'consumer data' should be provided as part of a new Act regarding data sharing and release (Draft Recommendation 9.11). Given the need for this definition to have broad applicability, it should also be included within the Acts Interpretation Act 1901 (Cth). Consequential amendments to other Commonwealth legislation would ensure harmonisation across federal laws.

**Yodlee agrees with the draft recommendation** that the definition of “consumer data” should include personal information, files posted online by the consumer, data derived from consumers’ online transactions and activities, and transaction or activity data relevant to third-party transfer.

Yodlee also agrees that the definition of “consumer data” should not include data that has been transformed to such an extent that it cannot be re-identified as being related to an individual. Balancing the usefulness of data with privacy protection is essential and nontrivial because, as the Productivity Commission points out, the risks associated with data sharing depend on the type, use, and benefits derived from the data. Yodlee, therefore, recommends that standards recognise a “spectrum of de-identification”<sup>1</sup> in which the extent of de-identification required is framed in the context of data utility.

### **Draft Recommendation 9.2**

#### DRAFT RECOMMENDATION 9.2

Individuals should have a Comprehensive Right to access digitally held data about themselves. This access right would give the individual a right to:

- continuing shared access with the data holder
- access the data provided directly by the individual, collected in the course of other actions (and including administrative datasets), or created by others, for example through re-identification
- request edits or corrections for reasons of accuracy
- be informed about the intention to disclose or sell data about them to third parties
- appeal automated decisions
- direct data holders to copy data in machine-readable form, either to the individual or to a nominated third party.

Individuals should also have the right, at any time, to opt out of a data collection process, subject to a number of exceptions. Exceptions would include data collected or used as:

<sup>1</sup> See, e.g., Jules Polonetsky, Omer Tene & Kelsey Finch, *Shades of Gray: Seeing the Full Spectrum of Practical Data De-Identification*, 56 SANTA CLARA L. REV. 593 (2016).

- a condition of continued delivery of a product or service to the individual
- necessary to satisfy legal obligations or legal claims
- necessary for a specific public interest purpose (including archival)
- part of a National Interest Dataset (as defined in Draft Recommendation 9.4).

The right to cease collection would not give individuals the capacity to prevent use of data collected on the individual up to the point of such cessation.

**Yodlee agrees with this draft recommendation.** Having a range of secure means to view, understand and engage with one's financial information is an important precondition to financial health. Stronger relationships and greater transparency among the financial institutions, data aggregators and third-party applications that participate in the ecosystem would enable the development of more secure means of sharing consumer data and further facilitate the creation of new, innovative products and services.

To evolve the current ecosystem to a state acceptable to all stakeholders, Yodlee recommends:

- A Consumer Bill of Rights be adopted to that codifies a consumer's right to their financial data. This should also define the consumer's role and obligations in the ecosystem, including the responsibility to make informed decisions, follow reasonable practices, participate in awareness/education activities and otherwise engage with their service providers and financial institutions as reasonably requested.
- Financial institutions should be required to allow a consumer to access their data directly or to provide permission for a qualified service provider or aggregator to do so on their behalf. This access should be available via a standard API, batch file and screen-scraping so that smaller financial institutions are not unduly burdened with the cost of implementing and maintain the API.
- Consumers should be able to access a minimum data set via any of these mechanisms that consists of all the account and transaction level data held by a financial institution. In other words, any data that is available to the consumer on his statement and/or via the native online interface(s).

- Industry and/or regulatory standards should be crafted, published and required for any aggregator and service provider to access the data held by financial institutions which give consumers comfort in relation to security issues.
- Standard contract terms should be defined between all stakeholders to provide the legal framework for participation and enforcement of the above standards, including dealing with liability issues.

Yodlee does not agree with the Productivity Commission's conclusion regarding data ownership; it is Yodlee's position that consumers own their data no matter who holds it. This right of ownership allows the consumer to access, use and share the data as they see fit. Ownership also comes with obligations for acceptable use and participation in programs of safeguards that will be determined by standards.

### 3. Response to Draft Findings

#### *Draft Finding 6.2*

##### DRAFT FINDING 6.2

Data standards should aim to ensure that the content produced is usable by those who seek access to their own data. To achieve this, available data needs to be published in machine readable and commonly used formats that are relevant to the function or field in which the data was originally collected or will likely be most commonly used.

**Yodlee supports this finding.** Publishing data in a standardised format is necessary to truly support Open Banking. A complete and extensible data model will directly benefit consumers via their chosen financial wellness solution and support innovation as new models and insights are developed once the financial data-sharing ecosystem is launched.



## 4. Response to Information Requests

### *Information Request: APIs*

---

#### *INFORMATION REQUEST*

*The Commission seeks more information on the benefits and costs of a legislative presumption in favour of providing data in an application programming interface (API) format, specifically:*

- In which sectors would consumers benefit from being able to access data in an API format?*
  - What are the main costs and barriers to implementing APIs?*
- 

In order to fully appreciate the benefits and costs of favouring (or mandating) APIs for data access, the alternatives must also be evaluated: batch files and screen-scraping.

Batch files are a tried-and-true mechanism and can enforce the data model. However, these considerations have mixed value. Batch files can become too large to transmit, handle and process efficiently, thereby raising the cost and underlying security risk. It is also difficult to uniformly change the data model across the data sharing ecosystem.

Screen-scraping, when done correctly, can be an effective solution. However, much of the technology for responsible screen-scraping is intellectual property of aggregators, like Yodlee, and therefore must be licensed which negates the cost benefit. Bulk screen scraping requires that all the data on the website is transferred and sifted by the receiver, not just the approved data model. This is both operationally intensive and potentially violates the data subject's explicit consent for its use. Therefore, bulk screen-scraping is not the best choice for a truly open data ecosystem. That said, responsible screen-scraping should be considered as an option for smaller data sources that cannot quickly participate in the data sharing ecosystem.

APIs have none of these challenges. They can be published as an open standard. They allow for an evolving data model that can be asynchronously adopted by the members of

the ecosystem. They provide efficient data transfer and a mechanism for responsible consumption of the data.

For these reasons, Yodlee is wholly supportive of the development of open API standards to facilitate the secure sharing of consumer financial data and a legislative presumption in favour of their use.

In our view, the open API presumption should require a standard set of APIs that provide consistent access to the data set(s). In the absence of this consistency, providers will find it difficult to comply with both privacy and security requirement and, most importantly, some consumers could encounter disparate data sharing and/or data access experiences which could negate the benefit of their chosen solution and their financial wellbeing.

While a consistent API is necessary to harmonize access, the underlying data model must also be standardized. A complete and extensible data model will directly benefit consumers via their chosen solutions provider(s) and support innovation as new models and insights are developed.

In our experience, the primary barrier to publishing via a standard API is that many data rich organizations do not have the internal connections required to make all their data available via an external API. In the case of banking, many financial institutions cannot create a single online view of their customer as the deposit system isn't connected to the loan system. Many health care providers and governments are in the same situation. Therefore, the incentives to these organizations must be sufficient for them to address their internal systems in order to participate in the open data sharing ecosystem.

## **Information Request: Access to Consumer Data**

---

### *INFORMATION REQUEST*

*Further views are sought on the effects of providing access to consumer data, as defined. In particular, views are sought on the potential creation of incentives for deliberate de-identification of data holdings to avoid providing access, and whether effective and low cost remedies to such behaviour could be introduced.*

---

In the context of Yodlee's response, deliberate de-identification of consumer data is essential to obtain the desired value from the data *and* to meet requirements for safeguarding the data and the consumer who owns that data.

Consumer financial data can be used in a consumer-permissioned model between the consumer and their service providers. For each of these relationships, the level of de-identification of the data can be selected to protect the consumer from inadvertent or malicious data leakage without impacting the value of their solution. For example, a financial health application can use bank and credit card account transactions to categorize spending without requiring the bank or card details. A consumer advocacy study can use de-identified transaction details grouped by postal code to assess living wage statistics without knowing names and addresses. In all these cases, the responsible use of the data is fully supported by the correct level of de-identification or anonymization.

Who, therefore, should be given an incentive to provide the de-identification? Is it the entity consuming the data, the source of the data or a third-party? This is determined by the use case. In general, Yodlee believes that the required governance over data sharing puts the default obligation for de-identification on the entity that initially collected the data from – or in respect of – the consumer, unless the consumer has consented to allow another collector of the data downstream from the initial collector (e.g., a service provider to the initial collector) to take on this responsibility. Data de-identification is not overly burdensome or costly, so guidelines and incentives need not be either.

## **Information Request: Methods of Disclosure**

---

### *INFORMATION REQUEST*

*The Commission seeks views on what methods of disclosure would be most likely to result in consumers making a meaningful choice about how their personal information is being used, and how these disclosure requirements might best be implemented.*

---

Consumer disclosures must be clear, concise and complete. These requirements are difficult to accommodate, especially when dealing with complex data sources, and are further complicated by user experiences on mobile devices, tablets, watches and Internet-of-Things devices. The complexity of this issue is directly proportional to the number of elements in the data set. Some academic consumer protection advocates would like to see the consumer permission each data element in the set. While this satisfies the requirement for “complete,” it prevents “concise” and risks the user making an uninformed decision. Other suggestions include creating standard use cases to map sets of data elements. Yodlee believes this is actually harmful to the consumer as it stifles innovation and the value of innovative improvements to data analysis models.

Yodlee recommends requiring members of the financial data ecosystem to provide disclosures in a standard format based on a template that sets out the information required by law (e.g., the service being delivered, how the data is used, the safeguards, third parties involved, information on secondary uses) and high level groupings of data elements for the consumer to consent to sharing (such as consent to sharing bank account and transaction data elements but not information relating to loans).

Presenting privacy disclosures in a simple and consistent format, similar to the Model Privacy Form required in the United States under the Gramm-Leach-Bliley Act, makes it easier for consumers to understand the entity’s information practices and the consumer’s options to limit use and disclosure of their information. Research commissioned by seven US agencies found that the format of privacy disclosures significantly affects consumers’ comprehension of the disclosure’s content and ability to compare different disclosures,

and that privacy disclosures presented in a table format are far easier for consumers to understand than other disclosure formats.<sup>2</sup>

This is combined with clear data governance requirements so that all recipients of the data have clear and enforceable requirements for safeguarding the data.

---

<sup>2</sup> Final Model Privacy Form under the Gramm-Leach-Bliley Act, 74 Fed. Reg. 62890, 62894-95 (Dec. 1, 2009), *available at* [https://www.ftc.gov/sites/default/files/documents/federal\\_register\\_notices/final-model-privacy-form-under-gramm-leach-bliley-act-16-cfr-part-313/091201gramm-leach.pdf](https://www.ftc.gov/sites/default/files/documents/federal_register_notices/final-model-privacy-form-under-gramm-leach-bliley-act-16-cfr-part-313/091201gramm-leach.pdf).