

12 January 2017

Mr. Peter Harris  
Presiding Commissioner  
Data Availability and Use Inquiry  
Productivity Commission  
Locked Bag 2, Collins St East  
MELBOURNE VIC 8003

Dear Mr. Harris,

**Re: Australia Post Submission to the Productivity Commission's *Inquiry into Data Availability and Use***

Australia Post welcomes the opportunity to comment on the Productivity Commission's *Draft Report on Data Availability and Use*.

This submission makes comments on specific elements of the Draft Report that either affect Australia Post directly, or where Australia Post believes it can provide feedback based on our extensive experience with data.

Australia Post supports the increased availability of data and its use by Australian citizens. We agree with the finding that: *Australia's provision of open access to data is below comparable countries with similar governance structures — including the United States and the United Kingdom. There remains considerable scope to improve the range of datasets published (and, correspondingly, the diversity of agencies and research bodies publicly releasing data) and the usability of open data portals.*

There are significant opportunities for Australia Post, to be a trusted partner of public and private organisations in the collection, management and sharing of data for the public good.

However, in the context of the Commission's Draft Report, Australia Post submits that any changes to current policy settings should be occur with due consideration of the following issues:

- Potential implications for the privacy of individuals.
- The costs to organisations of collecting and organising data into a format that is accessible for consumers, particularly if the definition of consumer data is expanded.
- Commercial and competition risks to organisations of the proposal that some public and private datasets may be considered in the "national interest" and therefore made available.
- Retaining the attractiveness of public-private sector partnerships. For example, the assessment of datasets for their strategic significance and public interest value prior to a government project involving the private sector so agreements could include a provision for valuable datasets to be acquired back by the government could be unworkable and create uncertainty regarding private sector ownership of their intellectual property.

The attached document includes Australia Post's comments on specific findings and recommendations. Please contact me should you require any further information from Australia Post to support this important inquiry, and we would welcome any opportunity to discuss this further.

~~Yours~~ sincerely,

Andrew Walduck  
Executive General Manager  
Trusted eCommerce Solutions Business  
Australia Post

ADDRESSING SPECIFIC IMPEDIMENTS TO PUBLIC SECTOR DATA ACCESS	
Draft finding/recommendation/information request	Australia Post response
<p><b>Draft finding 3.1</b>  <i>Australia's provision of open access to data is below comparable countries with similar governance structures — including the United States and the United Kingdom. There remains considerable scope to improve the range of datasets published (and, correspondingly, the diversity of agencies and research bodies publicly releasing data) and the usability of open data portals.</i></p>	<p>Australia Post agrees with this finding.</p>
<p><b>Draft Recommendation 3.1</b>  <i>All Australian Government agencies should create comprehensive, easy to access data registers (listing both data that is available and that which is not) by 1 October 2017 and publish these registers on data.gov.au.</i></p> <p><i>States and territories should create an equivalent model where one does not exist and in all cases should make registers comprehensive. These should in turn be linked to data.gov.au. The central agencies responsible for data should:</i></p> <ul style="list-style-type: none"> <li><i>set measurable objectives, consistent with best practice, for ensuring that available data and metadata are catalogued and searchable, in a machine-readable format</i></li> <li><i>improve accessibility of data for potential data users.</i></li> </ul> <p><i>Limited exceptions for high sensitivity datasets should apply. Where they do, a notice indicating certain unspecified datasets that have been assessed as Not Available should be published by the responsible department of state, on the relevant registry.</i></p>	<p>If Australia Post is considered a government agency for the purpose of this draft recommendation, clarity would be needed as to whether this would apply to all Australia Post products and services, or just those which are part of our Customer Service Obligation (CSO) requirements.</p> <p>If it were to be applied to all products and services, this recommendation could have implications for Australia Post as resources would need to be allocated to meeting this requirement; and there is potentially commercial risk to Australia Post in providing information about datasets it does and does not have.</p> <p>Providing a list of data that are not available would also potentially cause significant frustration for some stakeholders.</p>

ADDRESSING SPECIFIC IMPEDIMENTS TO PRIVATE SECTOR DATA ACCESS	
<p><b>Draft finding/recommendation/information request</b></p> <p><i>Draft Recommendation 4.2</i>  <i>All Australian governments entering into contracts with the private sector, which involve the creation of datasets in the course of delivering public services, should assess the strategic significance and public interest value of the data prior to contracting. Where data is assessed to be valuable, governments should retain the right to access or purchase that data in machine readable form and apply any analysis that is within the public interest.</i></p>	<p><b>Australia Post response</b></p> <p>Australia Post is concerned about this draft recommendation because:</p> <ul style="list-style-type: none"> <li>• Assessing the strategic significance and public interest value of data prior to contracting could be difficult.</li> <li>• This has the potential to cause uncertainty for the private sector in partnering with government agencies, particularly in relation to safeguarding their intellectual property.</li> </ul>

THE CONUNDRUM OF PERSONAL DATA	
Draft finding/recommendation/information request	Australia Post response
<p><b>Draft Finding 5.1</b></p> <p><b>1.1.1</b> <i>The boundaries of personal information are constantly shifting, in response to technological advances and community expectations. The legal definition of personal information, contained in the Privacy Act 1988 (Cth), gives rise to uncertainty. This uncertainty will only increase in future, as new technology continues to emerge.</i></p>	<p>Australia Post believes that the boundaries of personal information need to be balanced with an individual's right to privacy. For example, providing information in return for a product or service is different to using a social persona to interact in social media or online.</p> <p>If this recommendation was implemented, the community should be asked whether they want organisations to be able to connect their transactional requirements (i.e. where they have accepted some kind of value exchange) with other potentially identifying activities.</p>
<p><b>Draft Recommendation 5.1</b></p> <p><i>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.</i></p> <p><i>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.</i></p>	<p>Australia Post's view is that practical guidance through a set of high-level principles would be beneficial, however these should not be so detailed as to preclude rapid changes in technology.</p> <p>It is also our view that the definition and delineation of and between "allowable" and "best practice" should be clearly stipulated because re-identifying data can be relatively easy, particularly by combining multiple datasets.</p>
<p><b>Draft Recommendation 5.2</b></p> <p><i>The Privacy Act 1988 (Cth) exceptions that allow access to identifiable information for the purposes of health and medical research without seeking individuals' agreement, should be expanded to apply to all research that is determined to be in the public interest.</i></p> <p><i>The Office of the Australian Information Commissioner should develop and publish guidance on the inputs required to establish a public interest case.</i></p>	<p>Australia Post notes that existing datasets should be taken into account if this recommendation is implemented, because current datasets may have been collected under agreements with consumers that would be in conflict with this recommendation.</p>
<p><b>Draft Finding 5.4</b></p> <p><i>Large volumes of identifiable information are already published online by individuals or collected by various organisations, with or without explicit consent.</i></p> <p><i>In this context, the incremental risk of allowing increased access to formerly identifiable data by public and private sector organisations, using security protocols and trusted user models, is likely very small.</i></p> <p><i>Breaches of personal data, often enabled by individuals' unwary approach to offering data, are largely dominated by malicious or criminal activity. By comparison, breaches due to sharing or release are far fewer in number and reach.</i></p>	<p>We note that according to leading email security experts Return Path, phishing attacks increased by 162% between 2010 and 2014.</p> <p>It is likely that the release of data that is personally identifiable or has previously been identified could increase the risk of this data being misappropriated.</p> <p>Australia Post is regularly targeted by malicious activity.</p> <p>We would welcome the opportunity to provide the Productivity</p>

	Commission with a confidential briefing with more information on this issue.
<p><b>Information Request</b></p> <p><i>The Commission seeks further views on the most practical ways to ensure improvements to linked datasets are available for subsequent dataset uses.</i></p>	<p>Australia Post agrees with the sharing of datasets in principle. This would need to be governed by a robust framework which would:</p> <ul style="list-style-type: none"> <li>• Clarify the attributes of datasets that are appropriate for sharing in terms of achieving a public good.</li> <li>• Ensure the original owner of a dataset provides complementary information to maintain its integrity and usefulness for subsequent uses.</li> </ul>
<p><b>Draft Recommendation 5.4</b></p> <p><i>To streamline approval processes for data access, the Australian Government should:</i></p> <ul style="list-style-type: none"> <li>• <i>issue clear guidance to data custodians on their rights and responsibilities, ensuring that requests for data access are dealt with in a timely and efficient manner;</i></li> <li>• <i>require that data custodians report annually on their handling of requests for data access;</i></li> <li>• <i>prioritise funding to academic institutions that implement mutual recognition of approvals issued by accredited human research ethics committees.</i></li> </ul> <p><i>State and territory governments should mirror these approaches to enable use of data for jurisdictional comparisons and cross-jurisdiction research.</i></p>	<p>Australia Post anticipates that there would be significant compliance costs in adhering to requests for data in a timely and efficient manner.</p> <p>If this recommendation was implemented, the Government would need to work with organisations to ensure data access requests were responded to within organisational capabilities, rather than an arbitrary timeframe.</p>

MAKING DATA MORE USEFUL	
Draft finding/recommendation/information request	Australia Post response
<p><b>Draft Finding 6.1</b>  <i>The lack of public release and data sharing between government entities has contributed to fragmentation and duplication of data collection activities. This not only wastes public and private sector resources but also places a larger than necessary reporting burden on individuals and organisations.</i></p>	<p>Australia Post agrees with this finding, noting that it can be difficult in the current environment to know if some source data can be trusted (e.g. whether datasets use the same parameters).</p>
<p><b>Draft Recommendation 6.1</b>  <i>Government agencies should adopt and implement data management standards to support increased data availability and use as part of their implementation of the Australian Government's Public Data Policy Statement.</i>  <i>These standards should:</i></p> <ul style="list-style-type: none"> <li><i>• be published on agency websites</i></li> <li><i>• be adopted in consultation with data users and draw on existing standards where feasible</i></li> <li><i>• recognise sector-specific differences in data collection and use</i></li> <li><i>• support the sharing of data across Australian governments and agencies</i></li> <li><i>• enable all digitally collected data and metadata to be available in commonly used machine readable formats (that are relevant to the function or field in which the data was collected or will likely be most commonly used), including where relevant and authorised, for machine to machine interaction.</i></li> </ul> <p><i>Policy documents outlining the standards and how they will be implemented should be available in draft form for consultation by the end of 2017, with standards implemented by the end of 2020.</i></p> <p><i>Agencies that do not adopt agreed sector-specific standards would be noted as not fully implementing the Australian Government's Public Data Policy and would be required to work under a nominated Accredited Release Authority (Draft Recommendation 9.6) to improve the quality of their data holdings.</i></p>	<p>The cost of achieving and maintaining this standard could be high for both Australia Post and other organisations.</p> <p>Australia Post would be uniquely placed to provide advice in this area because of our strong track record of working with government and our digital capabilities. We would appreciate the opportunity to participate in focussed consultation on this recommended measure.</p> <p>It is noteworthy that machine readable formats that are relevant to the function or field in which the data was collected or used do not currently allow for linked sets of diverse data, which could be an opportunity for innovation.</p>
<p><b>Draft Recommendation 6.2</b>  <i>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.</i></p> <p><i>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.</i></p>	<p>Australia Post supports self-regulation by the private sector in this context.</p>
<p><b>Information Request</b>  <i>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:</i></p>	<p>Australia Post notes that Application Programming Interface format (API) is a technology solution and may be superseded by advances in digital technology. We therefore caution</p>

<ul style="list-style-type: none"> <li>• In which sectors would consumers benefit from being able to access data in an API format?</li> <li>• What are the main costs and barriers to implementing APIs?</li> </ul>	<p>against a legislative presumption in favour of specific technology.</p> <p>Such a legislative presumption would also be dependent on the resolution of other issues such as:</p> <ul style="list-style-type: none"> <li>• An agreed definition of consumer data.</li> <li>• Ability to connect a consumer to an increased definition of consumer data (e.g. their online interactions).</li> <li>• Commercial impacts of allowing consumers access to this information (e.g. loss of IP, competitive advantage).</li> <li>• Ability to recognise and identify the consumer correctly and with an acceptable level of assurance (e.g. to prove their digital identity before releasing such information).</li> </ul> <p>Finally, APIs are a gateway or link to information and assumes consumer data has been collected, aggregated and organised in a format which consumers can access and understand.</p>
<p><b>Draft Finding 6.2</b></p> <p><i>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.</i></p>	<p>Australia Post notes that data is collected and stored in many ways (e.g. flat files, relational databases, transaction rows and unstructured data such as voice recordings and social comments), and that formats used by companies or organisations for analytics, reporting or insight are not necessarily in an accessible format for consumers.</p> <p>This means that organisations would have to allocate resources to convert data to a format that is suitable for consumers.</p> <p>To indicate the scale of data that may need to be converted to a suitable format for consumers, according to the <i>State of Salesforce 2016 Annual Report</i> 80% of companies' data is "dark" and untouched. This report is based on over 110,000 data points collected from 1,700+ Salesforce customers globally.</p>

VALUING AND PRICING DATA	
Draft finding/recommendation/information request	Australia Post response
<p><b>Draft Recommendation 2.1</b></p> <p><i>In determining datasets for public release, a central government agency with policy</i></p>	A rigorous and transparent decision-making process would need to be established given the broad range of organisations

<p><i>responsibility for data should maintain a system whereby all Australian governments' agencies, researchers and the private sector can, on an ongoing basis, nominate datasets or combinations of datasets for public release, with the initial priority being the release of high value, in-demand datasets.</i></p> <p><i>A list of requested datasets should be published. Decisions regarding dataset release or otherwise, and access arrangements, should be transparent. Agencies should provide explanations where priority datasets are not subsequently released on legitimate grounds. Where there are not legitimate reasons for withholding requested data, remedial action should be undertaken by the Australian Government's central data agency to assist agencies to satisfy data requests.</i></p> <p><i>Existing government data initiatives, such as data.gov.au, should be leveraged as part of this system.</i></p>	<p>across the public and private sectors that might nominate datasets for public release.</p>
---	---



<p><b>Draft Recommendation 7.1</b>  <i>Beyond achieving a 'fit for release' standard (Draft Recommendation 6.1), government agencies should only value add to data if there is an identified public interest purpose for the agency to undertake additional value adding, or:</i></p> <ul style="list-style-type: none"> <li>• <i>the agency can perform the value adding more efficiently than either any private sector entities or end users of the data; and</i></li> <li>• <i>users have a demonstrable willingness to pay for the value added product; and</i></li> <li>• <i>the agency has the capability and capacity in-house or under existing contract; and</i></li> <li>• <i>the information technology upgrade risk is assessed and found to be small.</i></li> </ul>	<p>If this recommendation is accepted and implemented, Australia Post and other government agencies would presumably need to reallocate resources internally to enhance agency data to ensure it is 'fit for purpose' to make it more valuable for use by a third party.</p> <p>Should this measure be implemented, Australia Post may be able to provide customised products and services to assist governments comply with this.</p>
<p><b>Draft Finding 7.1</b>  <i>There is no single pricing approach that could act as a model for guiding public sector data release decisions. The identification by agencies of the grounds for undertaking each release will have a direct bearing on the choice of price approach. Cost recovery, long considered to be the default option in the public sector, is only one of a range of approaches and not necessarily to be preferred.</i></p>	<p>If cost recovery is not the recommended option for funding public sector data release decisions, agencies would be required to fund such releases from their own resources. This could not only lead to significant costs for agencies but would provide no incentive for requests for data to be focused on only what is required by those applying for data releases.</p>
<p><b>Draft Recommendation 7.3</b>  <i>Minimally processed public sector datasets should be made freely available or priced at marginal cost of release.</i></p> <p><i>Where there is a demand and public interest rationale for value-added datasets, agencies should adopt a cost recovery pricing approach. Further, they should experiment with lower prices to gauge the price sensitivity of demand, with a view to sustaining lower prices if demand proves to be reasonably price sensitive.</i></p>	<p>The definition of "minimally processed public sector datasets" would need to be clarified. It should be noted that minimally-processed public sector data sets would still need to have some standards and consistency applied to them.</p> <p>Australia Post earns revenue through the sale of our Postcode Address File (PAF) data. This could obviously be compromised should this recommendation be implemented.</p> <p>Another concern is the risk to Australia Post of data which has been collected by Australia Post being used improperly. We would suggest that any legislation, policy or guidelines assist organisations in managing this risk but also provide some protection to government organisations from this risk.</p>

FUNDAMENTAL REFORM IS NEEDED	
Draft finding/recommendation/information request	Australia Post response
<p><b>Draft Finding 8.1</b>  <i>It is important governments and businesses maintain a social licence for their collection and use of data. This can be built through enhancement of consumer rights, genuine safeguards, transparency, and effective management of risk. Community trust and acceptance will be vital for the implementation of any reforms to Australia's data infrastructure.</i></p>	<p>Australia Post agrees with this finding, noting that this needs to be guided by clear objectives, the widespread acceptance of standards and appropriate risk mitigation.</p> <p>Consumers will need to be educated about the benefits of how the sharing of their data contributes to community benefits, including through the sharing of successes. They will also need transparent and simple avenues of inquiry about how their data is being shared, who it is being shared with and the purpose of sharing.</p>
<p><b>Draft Finding 8.2</b>  <i>There is no shared vision amongst public sector data holders in Australia on how to consistently deliver widespread data sharing and release. The community — current and future — is entitled to expect such a vision. Comprehensive reform of Australia's data infrastructure is needed to signal that permission is granted for active data sharing and release and that data infrastructure and assets are a priority. Reforms should be underpinned by:</i></p> <ul style="list-style-type: none"> <li>• <i>clear and consistent leadership</i></li> <li>• <i>transparency and accountability for release and risk management</i></li> <li>• <i>reformed policies and legislation</i></li> <li>• <i>institutional change.</i></li> </ul>	<p>As above.</p>
<p><b>Draft Finding 8.3</b>  <i>By applying a risk-based approach to data access, government agencies can establish a sound basis for where further risk mitigation effort is necessary and for moving early to the sharing or release of low risk data, while building and retaining the trust and confidence of users and the wider community.</i></p>	<p>There are several important considerations in how this finding might be applied to public policy:</p> <ul style="list-style-type: none"> <li>• What are the standards for a risk-based approach?</li> <li>• Who will set them?</li> <li>• Who will endorse them?</li> <li>• Who will govern them?</li> </ul>
<p><b>Draft Recommendation 9.1</b>  <i>The Australian Government should introduce a definition of consumer data that includes:</i></p> <ul style="list-style-type: none"> <li>• <i>personal information, as defined in the Privacy Act 1988 (Cth)</i></li> <li>• <i>all files posted online by the consumer</i></li> <li>• <i>all data derived from consumers' online transactions or Internet-connected activity</i></li> <li>• <i>other data associated with transactions or activity that is relevant to the transfer of data to a nominated third party.</i></li> </ul> <p><i>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</i></p>	<p>Australia Post is concerned about the consumer privacy implications, red tape and cost burdens if this recommendation is implemented.</p> <p>Community engagement via research is vital to explore whether consumers want their data linked together in exchange for purchase or use of a product or service.</p> <p>It is also noteworthy that data derived from online transactions or internet activity is likely to be valuable intellectual property which organisations may not wish to disclose to an individual or</p>

<p><i>implementing any Comprehensive Right, be defined as consumer data.</i></p> <p><i>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).</i></p> <p><i>Consequential amendments to other Commonwealth legislation would ensure harmonisation across federal laws.</i></p>	<p>other third party.</p>
---	---------------------------

**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:*

- *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.*

While in principle individuals should have a comprehensive right to digitally-held data about themselves, the following should be considered before any changes are made:

- The need for an agreed and appropriately-focused definition of consumer data.
- The commercial impact of allowing consumers access to this information (e.g. loss of IP, competitive advantage).
- The ability to recognise and identify the consumer correctly and with an acceptable level of assurance (e.g. to prove their digital identity before information is released).
- The risk of compromise once data has been provided to the consumer in a machine or electronic format.
- The need to guard against data being intercepted or re-identified maliciously.

For this to be implemented efficiently, consumer data will also need to have been collected, aggregated and organised in an accessible format for consumers to access and understand.

This could be very difficult for Australia Post which like many organisations does not have a single customer view of our data. We have multiple data sources (e.g. parcel event data, customer feedback surveys, vast amounts of unstructured data e.g. call recordings, survey verbatims, social posts using social personas which are unlikely to be linked to a customer record) and not all consumer data is captured against this.

We have also not collected this data with a view to sharing, so would need to inform and either mandate or request permission to share data. Depending on the level of data being requested, consumers may feel that their information is at increased risk the more it is shared.