National Freight Data Hub: Options Discussion Paper - Response Template

Please note: Submissions close on 11 September 2020 at 17:00, and should be emailed to freightdatahub@infrastructure.gov.au.

We welcome all responses. You may use this template, or simply email your response. You may address all questions, or you may choose to respond to selected questions of interest to you.

Your submission will be published on the website unless you request otherwise. Information collected during this consultation process may be provided to persons making an application under freedom of information laws. Personal details will not be published in any report.

Respondent details

<table>
<thead>
<tr>
<th>Organisation name</th>
<th>Australian Rail Track Corporation</th>
</tr>
</thead>
</table>
| Contact details (to whom any correspondence in relation to this submission can be addressed) | ARTC owns, leases and manages a rail network with over 8,500km of rail that is a key element of Australia’s freight supply chain. A significant constraint on ARTC’s ability to maximize rail’s competitive position in this freight supply chain is a lack of data on alternative transport modes. ARTC considers the National Freight Data Hub a necessary, and critical, innovation to ensure data is available which allows in depth analysis of the operation of the freight supply chain to ensure optimal policy and investment decisions are made. The development of a centralized and independent freight supply chain data hub will also provide a framework for the potential development of a national freight supply chain operator who is responsible for assessing the overall productivity of Australia’s freight supply chain, and not just a particular mode. The data hub will provide participants the ability to identify what freight is moving where; and more importantly:  
- What Mode that freight is using;  
- What route it is taking; and  
- The Weight of the freight  
The data hub should leverage the significant work of industry in developing the GSI system of global supply chain standards. The potential productivity and policy benefits to freight, and also to the wider Australian economy, from this development are significant. The benefits to government through better informed investment decision making will more than offset the cost of development of the hub. The development of this data hub should therefore be a significant policy priority for all levels of government and industry; with funding initially by government but with a requirement, similar to the ABS, that bespoke requests be met on a user pays basis. |
Questions for discussion – Design Principles

1. Please share your organisation’s perspectives on the proposed design principles, including any which are not represented here.

The provision of a complete data set in respect of the operation of Australia’s freight supply chain is critical to ensure informed policy and investment decisions are made by governments and participants alike. The critical gap in this data set currently is the lack of data in respect of road freight. That is, there is limited understanding of the size and performance of the road freight task which inhibits the ability for informed policy and competitive responses from competing transport forms. This lack of data also constrains the ability to understand the impact of road freight on road congestion and the consequent cost to the economy. The provision of a complete freight data set is therefore a critical initiative.

Whilst ARTC agrees that data is a collaboration, the value of the data to individual participants will vary. Therefore, a “give and get” model is unlikely to incentivise all parties to contribute data. A form of compulsion is therefore required to ensure all necessary, and consistent, data is collected – which in turn sets a requirement for a trusted, independent facilitator to mitigate any concerns with the treatment of the data. Data exchange therefore needs to be more than collaborative to ensure that the data set is not compromised by parties who choose not to collaborate given the cost of collaboration may outweigh any gains.

Ultimately, breaking down the barriers which exist in understanding the operation of Australia’s freight supply chain is critical and removal of the significant information asymmetries is an essential task to allow informed policy and investment decisions. This requires the Hub to lead by example and will ensure that high quality data is used to deliver high quality insights by both government and industry.

The data hub should leverage the existing industry standards, in particular the Australian Freight Labelling Guideline based on the GS1 Open Global Standards. This will ensure consistency with existing industry arrangements and approaches which should minimize implementation costs and maximize industry engagement.

ARTC therefore mainly agrees with the proposed design principles for the hub. The one area where it feels the principles should be strengthened is in respect of the collaborative nature of the data exchanges, where a degree of compulsion is required to ensure the symmetry of the data set is not compromised from a choice to abstain from the data provision.

The Hunter Valley Coal Chain Coordinator is an example of a system operator that gathers data from system participants, aggregates it and the presents it in a fashion that allows industry to analyse that data and use it in decision making to the benefit of the full coal supply chain. This development has been key in making the Hunter Valley coal chain one of the most efficient multiple user resource supply chains in the world.

Similarly, the Gas Market Bulletin Board is another great example where very specific information on the functioning of the natural gas market collated is collated and presented to allow participants, and policy makers, a detailed understanding of the market. Given differentials in the value of information across participants, this was a development that was imposed with a degree of compulsion but now provides significant benefits at an operational and policy level. This was not an example of an industry data hub referenced in the paper, and ARTC recommends analysis of its development over time as a case study in how industry data can be used to drive the optimal industry and policy responses.
Questions for discussion – Data

2. What specific benefits would each data priority provide to your organisation?

3. What level of data fidelity (i.e. transaction level data or aggregated data) and frequency (i.e. near real-time, weekly, monthly, quarterly) would be required to make the data priorities you’ve identified be of value?

The data which provides the most value is that which is focused on the unit of contestability in the freight supply chain between modes. Provision of this data will allow government and industry to analyse relative supply performance of modes. That is, the data will highlight what freight is moving:

- where; and
- on what mode and route; and
- at what weight

This data will therefore underpin the analysis required to make the necessary policy, strategy and investment choices to maximize the performance of their mode and the economic benefits to Australia.

Whilst ARTC believes that an understanding at the consignment level would be of value and would open up greater opportunities for strategic pricing decisions, this is not data that is available currently. The cost and complexity to collate that data could therefore be significant and so represents a second order priority compared to the value which can be created by focussing on container and vehicle data in the first instance.

The collation and display of infrastructure data and cost data is data that is already collected, at least for rail. The collation and presentation of capacity availability, utilization and cost would be of significant value to stakeholders and therefore is the first order of development for the data hub.

The utilization of infrastructure would be at the vehicle level, which suggests that should be the first priority for development. This is little different from container level data when considering truck movements, though there is a difference for rail. The contents of the containers would also provide significant value in how freight moves; however as above, this would require a substantial investment in technology and data measurement (and time) which should not delay the collection and presentation of higher-level data sets.

Therefore, ARTC believes that the first order focus should be on vehicle, infrastructure and cost data with a plan to roll the data sets down quickly to the container and then to the consignment level as the acquisition costs are made clearer. Leveraging the existing Australian Freight Labelling Guidelines based on GS1 Open Global Standards will assist in the pace and consistency of that extension into consignment data.

The largest benefit will come with high fidelity and high frequency data with mandated standards to ensure a consistency of data sets (and hence consistency of analysis).

This data should be publicly available as an aggregated data set – such that individual users are not identifiable. However, the aggregation should be such that the value of information is not lost. For instance, it should be clear what trucks are using a particular road segment at a point of time such that available capacity (and impact on congestion) is clear as well as the origin and destination of that truck and its weight.

Presentation and analysis of this data will be key to developing future policy actions, especially in respect of future heavy vehicle pricing and access frameworks to ensure that the externality impacts on congestion of heavy vehicles can be understood.
Questions for discussion – Technology

4. If a centralised or federated architecture model were pursued, what would be the benefits and challenges to your organisation to participate in the Hub?
5. What are the preferred methods and technologies to integrate with the data exchange platform?

As a starting point, a centralized architecture model would likely be the simplest to deliver; although noting this requires all individual data providers to connect into the system, such that it starts to look like a federated system anyway. The GS1 system embodies an open architecture approach that allows for modular expansion of the system with minimal disruption to existing applications. The data hub should leverage this existing investment in the most efficient manner. Therefore, the preferred architecture should reflect the best system to integrate with the existing industry framework. This ensures the optimization of the pace and cost of implementation and extension and delivers the best value for all stakeholders.

Regardless of the final system architecture, the system should be run by a nominated infomediary who owns the data and is responsible for confidentiality of that data, which should assist the engagement with all users and provides the simplest ability to manage data security and access levels. Similarly to how the Gas Market Bulletin Board is managed by AEMO, the development of this structure could be the forerunner to the development of an Australian Freight Market Operator to focus on the productivity of Australian freight across all modes rather than just on roads.

ARTC believes the development of a body focused on the productivity and operation of the Australian freight supply chain is a significant policy imperative and therefore the creation of this centralized data architecture is an important precursor to such a development. Overtime, a transition to a federated system would be assessed.

The integration with ARTC’s systems would require further detailed analysis once more detail in the proposed approach is understood. However, given the likely movement to more automated train management systems, the ability to connect operational data from these systems into a centralized data collection system is improved from the analogue processes that have been historically utilized. ARTC has been able to integrate its Hunter Valley operational data requirements with the needs of the HVCCC, so the provision of data is manageable on a system wide basis prior to automation.

The GMBB and collation of utilization data from gas market and transmission operators is a useful case study to start from in how to develop the data collection systems and necessary interactions.

Questions for discussion – Governance

6. Which governance structure could enable the Hub to be established quickly and generate quick wins, and should it change over time?
7. Which governance structure is most likely to facilitate the greatest use and participation?

Given that both government and industry will share significant value from the data, ARTC supports a Shared governance approach with funding coming from both industry and government. This also reflects the ownership of road infrastructure especially is public.

This structure can be reviewed periodically in line with changes in industry structure and governance. However, given the policy benefits arising from access to the data sets will match the strategic and operational benefits, a default to shared governance should be in place.
Questions for discussion – Funding

8. What funding arrangements could ensure users gain the value they are seeking from the Hub?
9. What services could the Hub provide that could be paid for by users?

Funding should follow governance and be shared between government and industry with some percentage of user pays, but with government meeting the initial set up costs. The percentage of user funding could grow over time, based on developments in pricing and access reform (again, especially in the road space). Where significant investments are required in technology to access and transmit data which provide limited value to users, government should bear the brunt of this cost given the policy benefits available from the data. That is, the benefits to government of improved investment decision making (given the size of those investments) more than offset the initial costs of acquiring the data and setting up the data hub. This would ensure the cost of acquiring the necessary data is not an impediment (or excuse) for the provision of data by any user.

Once the system is set up, and as all users benefit from participating in the freight supply chain, the costs of managing the system should for a part of their participation in the market. Given that all users benefit from a more efficient

Where the hub provides data that can be specifically applied by Users, such as trucking companies using congestion data in their contractual pricing arrangements, this should be paid for by those Users only. That is, where users request bespoke data, the costs of providing that data should be met on a user pays basis (similar to how the ABS manages specific data requests)

A principal focus of the allocation of cost to primary benefit should therefore be applied. Where beneficiaries are unable to be determined, that is an indication the costs should be shared across all government and industry users of the hub.

Questions for discussion – Regulatory

10. To support the Hub’s governance, ability to collect and share data, setting of standards and funding model, which regulatory option is best suited?
11. Would there be significant costs or benefits for your organisation associated with each of the regulatory options?
12. Are there additional circumstances to those outlined above, that may warrant a change, introduction or removal of a regulatory mechanism?

To provide value from the data set, there must be confidence that the data set is both complete and consistent across all freight modes. Where Government has funded the set-up cost and hub development it is even more imperative that the data be provided in a consistent and complete fashion. This can only be achieved through a legislated approach based upon a consistent data standard, such as GS1 discussed above. This is similar to the GMBB approach.

This presents some risk to Users, but only if there is reluctance to provide the data. Errors will occur; therefore, the penalties should not be onerous for first offences – but they should increase with frequency to ensure compliance and the value of the data set is not eroded.
Ultimately, the cost of compliance is significantly less than the value of the data through more efficient investment and informed decision making. Therefore, this compliance cost should not be seen as a barrier to enforcing the need for consistent data collection.

This requires balance and therefore data security and privacy must also be mandated by legislative requirements to ensure participants have confidence that their information will be protected.

This will create costs in set up and compliance and the individual costs should be assessed and any required government contribution in the set-up costs be evaluated. However, once the system is operational, the ongoing provision, maintenance and compliance costs should be borne by the participants.