Amey

Written evidence

Transport Select Committee: Future of Transport data inquiry

August 2023

About Amey

Amey is a leading provider of full life-cycle engineering, operations and decarbonisation solutions, for transport infrastructure and complex facilities. Our purpose is to deliver sustainable infrastructure solutions, enhancing life, protecting our shared future.

We combine exceptional expertise in Consulting & Design, Advisory & Analytics, Transport Infrastructure and Complex Facilities to design, manage, and maintain clients' assets throughout their lifetime. Our industry leading data and analytics capability, combined with customer-focused advisory services, enables us to create transformative solutions that strengthen transport resilience and drive sustainable improvements that make a real difference to people's lives.

We welcome the opportunity to submit written evidence to the Transport Select Committee's inquiry on the future of transport data. Our expertise and investment in transport data, skills development and leadership in governance and standards is aligned with the Government's transport data strategy and recently updated geospatial strategy. Our recent investment in a new mobile network data capability alongside our long-established practice of bringing actionable, data drive insights to bear across infrastructure means that we are well placed to advise Government and all our clients through better understanding of travel patterns, leading to more user focused and sustainable transport outcomes.

We trust that our submission is useful to the Committee and would welcome the opportunity to discuss our expertise in this area in further detail.

Executive summary

The transport industry faces substantial challenges in releasing the benefits that better transport data can bring operators, government bodies and the travelling public.

In a world where the discussion on the impact of transport projects is becoming increasingly polarised, having easier, faster, and cheaper access to data will result in improved timescales for project assessments leading to quicker implementation of large transport infrastructure projects.

Data also plays a key role in connecting transport hubs, such as major seaports with hauliers. We previously used mobile network data in this space to help understand HGV movements between Northern Ireland and Great Britain to help with Brexit negotiations.

If the Government is serious about decarbonisation goals, then the availability of active travel data must be improved. This could be achieved by increased monitoring of active travel modes using mobile phone data and by making the data embedded in the National Travel Survey more readily available.

Data providers should be able to upload data to DfT owned sandboxes from which researchers and potential data users can innovate and test ideas with a wide range of data sources in a controlled environment. This will allow the UK to scale towards an integrated, comprehensive landscape for transport data.

How might planning and delivery of transport infrastructure and services be changed by greater sharing and use of transport data over the medium and long terms?

1. Overall, we agree with the observations in the Department for Transport's (DfT) Transport Data Strategy (2023) that the transport industry faces substantial challenges in releasing the benefits that open data can bring operators, government bodies and the travelling public.[1]

2. Many of the promised benefits of technological transport innovations, such as Mobility as a Service (MaaS), connected autonomous vehicles (CAVs) and even plain fleet electrification, can only be achieved if the data issue is resolved. It is unlikely that the private sector will be able to resolve this alone and we support the progress being made by the DfT on data standards and quality. The potential of MaaS can only be released when and if the commercial incentive of collaboration outweighs the commercial disincentive of the fear of modal and inter-operator competition.

3. In a world of making better use of existing assets in the transport system (as well as recognising the need for a substantial investment in infrastructure maintenance, both to deal with ageing assets but also to respond to or avoid major climate change impacts) data is essential. The recent two-month outage of the Nuneham viaduct between Oxford and Reading led to misery for thousands of rail passengers which could have been avoided by earlier and better monitoring and response. In addition, the traditional response of providing replacement bus services could be improved upon, for example by providing better data to rail users on alternative routes that were available, perhaps supported by financial incentives. When making better use of existing assets, the logical investment is in data rather than tarmac or steel.

4. In a future of connected and (possibly) automated vehicles, data on the state of the road transport system will become available instantaneously and continuously, and there will be multiple benefits, including:
    *   Immediate awareness of traffic conditions ahead, so that drivers or their vehicles can adjust speeds and routes accordingly.
    *   Information shared to operators, such as National Highways, to provide the necessary services to deal with incidents and their aftermath.
    *   Monitoring data, before and after, to improve future responses to incidents.

5. Such data are not a fortunate by-product – they are a key element of successful system operation. Outside of the operation and user experience of the transport system, there are further benefits to be gained by having easier, faster, cheaper access to data. For example, this will lead to a faster turn-around time for project assessment (modelling and appraisal) leading to quicker implementation of large transport infrastructure projects. National Highways has already recognised this by developing its six Regional Traffic Models, using mobile phone derived travel patterns, which Amey's CitiLogik portal can also provide.[2]

6. In a world in which the discussion about impact of transport projects is becoming more and more polarised, this data derived from mobile phones, produced in near-real time using innovative data analysis techniques including artificial intelligence, provides an opportunity for better and quicker post opening assessment. This also helps in adjusting solutions quickly as a scheme is being implemented and continuously monitored – which would be helpful in the pro- and anti-debate around, for example, low traffic neighbourhood interventions and controversial cycling or parking schemes. Finally, this type of data provides an opportunity to understand unexpected travel responses to projects (which might not be picked up by traditional, more focused evaluation approaches), avoiding repeating the same mistake twice.

---

[1] https://www.gov.uk/government/publications/transport-data-strategy-innovation-through-data
[2] https://www.citilogik.com/portal

7. We recommend that the committee looks at how other governments approach the use of transport data, for example, the Victorian Government in Australia or the Netherlands National Data Portal Road Traffic. [34]

How might the travelling public, and local communities, experience the benefits of better use of transport data? What unintended consequences might there be?

8. The ultimate beneficiaries of better use of transport data will be the travelling public, if the data enables faster and cheaper travel, with increased reliability and resilience. Solid transport data enables the user to understand the full set of alternatives available to them to make a trip, be that a choice of route, destination, travel time or mode. Mobility as a Service can, in principle, provide travellers with the confidence of reaching their desired destination at their desired time, without being prescriptive up-front about how they will get there – also responding to unexpected incidents en route – for example changing route or mode without complicating payment. To do so, data is key.

9. This also illustrates the risk of unintended consequences. First of all, by providing all travellers with optimum advice on destinations, routes and modes, operational resilience will be removed from the system. A short-term benefit for travellers may be followed by a sustained period of reduced system reliability – basically as suboptimal behaviour has been rooted out.

10. A second and not unrealistic unintended consequence is that through push notices and targeted advertising, the travelling public may be encouraged to choose options that are either not optimal for them, or that may go against wider decarbonisation, social equity and health policy objectives.

11. Local communities are already experiencing some unintended consequences. For example, anecdotal evidence already exists that modern satnav systems suggest routes through residential areas that would previously would not have been known or used by drivers. Examples that reach the headlines include buses and lorries getting stuck at inappropriate pinch points. Governance of the data, and how it is used, remains essential, which relates back to the key drivers of success for a data strategy:
    - Identify data sources and encourage others to share.
    - Standards that will ensure that data is comparable and compatible (note – this is a firm commitment under the DfT Data Strategy).
    - Quality, ensuring that data that is provided for use, has been cleaned and verified.
    - Store the data securely and aligned with local laws and regulations.
    - Provide the data, including agreement on pricing and compensation where relevant.
    - Integrate data where the sum of the two adds value (but with an awareness of the risk of, for example, reverse engineering negating anonymity).
    - Governance of the use of data for unintended purposes, with an option to withdraw or restrict data streams when issues arise.

How will it benefit the freight sector and the supply chain?

12. Although much of our data experience in Amey relates to the transport of people, we recognise that major benefits can also be realised for freight and logistics. For example, the sharing of and access to better freight movement data (vehicle locations in time/space but also in terms of the types of goods being moved), should enable the identification of key routes for either combining loads, avoiding empty return journeys, or creating opportunity for multi modal interchanges for freight.

13. Data can also play a part in better connecting transport hubs, such as major seaports with hauliers, helping to further smooth out the supply chain but also better handle border crossings and border controls. Amey CitiLogik have previously used mobile network data in this space to help understand HGV movements between Northern Ireland and Great Britain to help with Brexit negotiations. [5]

---

[3] https://discover.data.vic.gov.au/organization/department-transport-planning
[4] https://www.ndw.nu/
[5] https://www.economy-ni.gov.uk/publications/analysing-hgv-movements-between-ni-and-gb-using-mobile-network-data

14. We are convinced, however, that it is less the availability and access to freight vehicle locations and movements (which is already possible using mobile phone data), and mainly the availability of data about the goods being moved, and haulier competition, that needs resolving.

15. For the logistics sector, there are benefits for individual providers to have better access to transport data, but because of the competitive nature of the sector, the societal benefits of integration and sharing data (through reduced congestion, reduced emissions, reduced accident risk) will be harder to realise.

What are the potential uses of data for understanding usage and condition of assets like roads, rail track, charging points, vehicles and the kerbside?

16. Here we suggest you clearly distinguish between usage and condition monitoring. Usage data of roads, rail track, charging points, vehicles and the kerbside is already possible by using mobile phone derived information, particularly the large scale, ubiquitous datasets that are available through the need for any device to be connected to cell towers at all times. Your inclusion of the kerbside is interesting, in that this is an asset in urban situations that is both undervalued and often misused.

17. Typically, usage is linked to individual movement, and thus governed by use of personal data and how this should be handled given GDPR requirements. Condition of assets does not necessarily have the same challenges around personal data.

18. A judicious metering of kerbside usage, and associated pricing, might be of great value for smoothing traffic flow by removing temporary obstacles, deal with illegal parking with impacts for safety for vulnerable groups, and provide income. The appyway app may be of interest, as an example.[6]

19. Understanding the condition of roads, tracks and vehicles can be achieved by unlocking the potential of other features of modern mobile phones, which can collect data almost continuously using a fine granular GPS positioning, altometer and/or accelerator. Although not everybody will have these features switched on, or be willing to transmit these, we expect that there are sufficient devices in use to provide credible information. An innovation challenge may lead to the development of appropriate tools, which could be combined with the more ubiquitous movement data obtained from mobile phones, to improve data density.

20. Another type of data to monitor asset condition includes car manufacturers' monitoring data. Information on weather conditions as well as acceleration, use of wipers, fuel consumption etc. This could be improved further by measuring shock absorbers' performance as a measure of roughness. Measured over time and across the fleet, this could be used as a digital, data-driven way to determine when maintenance activities may need to take place on a specific asset.

What privacy, ethical, security, resilience and intellectual property issues arise in relation to gathering and sharing transport data, including location-based data about journeys and data with commercial value? How should the Government seek to manage and regulate these?

21. The only regulation affecting the collection, processing and use of mobile phone data is GDPR. While we respect the principles of data privacy, GDPR is often viewed as a practical hindrance to the open sharing of transport-related data.

22. In our experience, many private sector data providers also interpret GDPR requirements differently and as a result, the quality of transport data that they provide differs widely – for example in terms of minimum sample sizes, spatial granularity and other aggregation thresholds. A clear consistent approach towards anonymisation, and aggregation to protect individuals and privacy would help. The Government can help here, and particularly in the determination of quantified minimum levels of aggregation required for any personal data provided by the private and public sector. https://www.amey.co.uk/media/press-releases/2023/january/new-mobile-data-capability-will-help-amey-support-clients-deliver-better-transport-solutions

---

[6] https://appyway.com/

23. In addition, often public contracts have standard and inappropriate IP clauses, which can be seen as a risk for data providers especially when direct data is not provided, but first processed to become meaningful for use in planning. Mobile network data is a perfect example of this – data is processed both by the network provider and in our case the data scientists within Amey's analytics team, but local authorities' contract conditions can request access to the raw data, which cannot be provided for commercial and legal reasons. The Government can help by the development and provision of reasonable conditions that protect suppliers and the taxpayer equally.

24. As an aside, there is the feeling that data should be free by many. Whilst data (particularly if collected for public sector projects) could be free, the cleaning, processing, validation, storage, and access points have clear associated costs and need to be paid for.

What are the biggest gaps in available data about transport networks and travel? What kinds of policy, planning or maintenance questions cannot currently be answered that we could answer with new, or more accessible, data?

25. Available transport data is highly road and rail orientated. This is probably related to the extent of both networks, the value invested in these systems, and the ease with which data on movement can be collected at these systems, e.g., through loop counters and ticketing systems. At the same time, the impact of failure on either system is such that their condition is regularly and automatically monitored, as opposed to other parts of the transport system such as bus stops, cycle paths or the pavement used by pedestrians.

26. This discrepancy in monitoring between modes may well have an impact on investment and maintenance. According to the National Travel Survey in 2019, i.e. pre-COVID, 26% of all trips were made by walking[7], but we know a lot less about who walks, why, where and how often.

27. Few datasets provide robust information on active modes (for example, it is widely expected that NTS, and all other household travel surveys tend to underestimate short walk trips). In addition, most datasets do not provide sufficient context to the trips they capture, to fully support planning purposes. For example, although trips can be captured, the reason for travelling (trip purpose) can only be inferred. Person characteristics, and particularly those that support the Government's objectives in terms of social exclusion and health are often not available, sometimes by a misguided interpretation of GDPR regulations as described earlier. There is also the issue of local behaviour which is not covered well by NTS due to its sample size and research methods, which tend to overlook and under-report these short trips.

28. If the Government is serious about its decarbonisation goals (for example its target that half of all journeys in towns and cities will be cycled or walked by 2030)[8] then the lack of data on travel movements made by active modes needs to be improved. In our opinion, this can be achieved by increased monitoring of these modes using mobile phones, but also by making the data embedded in the National Travel Survey more easily available to all those working in transport.

29. By merging mobile network data from travel movements (such as embedded in Amey's CitiLogik portal) with NTS data on the number of trips made, when, where and how, by different person categories, we can derive proxies for many of the explanatory variable that are at the moment hidden in the anonymised phone traces. This would be worthy of deeper investigation and funding through e.g., Innovate UK.

How can the UK scale up from pilots, pockets of innovation and existing single-mode data sets towards an integrated, comprehensive landscape for transport data?

---

[7] https://www.gov.uk/government/statistics/walking-and-cycling-statistics-england-2020/walking-and-cycling-statistics-england-2020

[8] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1009448/decarbonising-transport-a-better-greener-britain.pdf

30. We strongly believe that the DfT should be the data aggregator and integrator. Its Transport Data Strategy illustrates its understanding of the importance of data (collation, storage, quality control, provision and governance). An effective way to achieve the ambitions set out in the report is by creating repository and data lakes to store data collected at all different levels of spatial granularity (local authorities, agencies, nationally).

31. Data providers should be able to upload data to DfT owned sandboxes from which researchers and potential data users can work and test ideas with a wide range of data sources. Transport for London (TfL) has seen enormous benefits from making its data available.[9] In 2017, it was estimated that TfL open data directly support around 500 jobs that would not have existed otherwise, whilst a further 230 indirect jobs in the supply chain and wider economy have also been created.[10]

32. We suggest that DfT promotes this data through organised events such as workshops, conferences, and even hackathons to trigger new ideas from data being made available. At a much smaller scale, DfT already provides data for transport modelling and appraisal, through its TAG Databook. Although the technical challenge of wider collation, governance and provision of much larger quantities of transport data is quite different from the spreadsheet approach used in TAG, DfT is not averse to taking responsibility where it sees the need and benefits.

How should data availability, and sharing by transport operators, suppliers and other bodies, be encouraged, facilitated and regulated?

33. As above.

What skills and capacity do operators, infrastructure providers and local transport authorities need in order to manage their own data well and get the most value out of available data? What help do they need to anticipate and cater for future requirements?

34. Technically, and based on Amey's in-house experience, we suggest that the minimum skills and capacity requirements are access to data analysts and full stack developers. In addition, organisations also require a good understanding of developing and operating in cloud environments.

35. In general, these capacities will be unaffordable for local transport authorities to attract and maintain in-house. Rather than decentralising this need for skills and staffing, we suggest that facilities and staffing would be provided centrally, by DfT itself or outsourced.

36. We envisage a minimum data environment that enables operators, infrastructure providers and local transport authorities to upload their data, standardised as per the standards set by DfT via a portal, which can then be visualised and analysed, and merged with other datasets if so desired. The functionality of the tool would be determined via stakeholder engagement.

37. Then, data interrogation and visualisation should be facilitated by the use of web-based access portals with front ends and simple query tools, but also more advanced functionalities for the more expert user to interact with data. This could take the form of APIs for example. This means that UI/UX technologists must be brought in to help a wider audience accessing the data.

38. We suggest the Government takes a greater interest in what is developed under the DAFNI (Data & Analytics Facility for National Infrastructure) initiative representing an £8 million investment from the UK Collaboratorium for Research on Infrastructure and Cities (UKCRIC) to provide world leading infrastructure systems research capabilities and enhance the quality of outputs.[11] A mainly academic initiative, we believe that it could and should provide greater value for transport data practitioners.

Is the UK's digital infrastructure sufficient to allow the greatest value to be derived from transport data?

---

[9] https://tfl.gov.uk/info-for/open-data-users/our-open-data?intcmp=3671
[10] https://content.tfl.gov.uk/deloitte-report-tfl-open-data.pdf
[11] https://www.dafni.ac.uk/about-2/

39. We believe that the UK's digital infrastructure is more than sufficiently developed to allow greater value to be derived from transport data; the constraining factors are much more related to a lack of sufficient expertise in data engineering and data analytics, in the UK overall, and in the transport sector more generally.

40. As part of the DfT's data strategy, identification of those courses that produce graduates with appropriate skills, and the promotion of attendance at Master's courses where these skills can be further honed, will be of immediate and lasting value.

How effectively does the Government use data in appraising and prioritising transport investment?

41. Appraising and prioritising transport investment in the UK is governed by the Department's Transport Analysis Guidance (TAG). TAG is carefully aligned with the Treasury's Green Book requirements and is considered internationally as best in class.

42. However, as TAG's guidance on the use of data in modelling and appraisal was developed some 20 years ago, well before many of the current data sources were available, derived from mobile phones and other connected devices, its data-related advice lags current best practice. A good example is that its Unit M2.1 (Data Sources and Surveys) is wholly focused on traditional surveys, such as traffic counts and roadside interviews – the latter being considered expensive and disruptive, and mainly replaced in practice by mobile network data. Using mobile network data (MND) is only described in a separate report, and even then, only focused on the development of trip matrices, ignoring the wider potential of this data to replace and improve on techniques currently advised to practitioners.

43. Appraisal also tends to be strongly model, rather than data driven. This again may be largely due to the lack of data when TAG was developed, and the subsequent investment by the appraisal profession in strategic transport models.

44. Proper pursuit and success in the delivery of the DfT's Transport Data Strategy will offer opportunity for a greater role of data in forecasting and appraisal. Apart from the added richness that observed data can bring over data that has been forced into a TAG-compliance straitjacket, better use of data in the early stages of appraisal (particularly the Strategic Outline Case (SOC) and Outline Business Case (OBC)) will save time and money. We are convinced that objections that hinder progress on key schemes, can be at least reduced if data rather than models illustrate what the need for change is, and how this change may impact different parts of the population.

45. In scheme development and investment prioritisation, millions are spent on model development and application. A larger proportion should be invested in data – both up-front when developing the tools and data platforms to ease access, and in their subsequent acquisition and use. This will also provide a natural route into monitoring and evaluation.

46. There should also be more of a focus on modelling. One of the great benefits of data, is that it can help monitor if a scheme is performing as intended (and modelled) after opening. There are many benefits to this approach:
    - Help confirm user responses to a new scheme identify flaws in design which may need to be addressed in future similar schemes.
    - Help gather data which could be used for future scheme assessments.
    - Help educating wider stakeholders on true observed benefit of specific schemes.
    - Help improve forecast modelling by discovering incorrect modelled behaviour responses.

What milestones and ambitions should the Government set in this area? How effectively has the Government's Transport Data Strategy identified barriers to sharing and getting value from transport data, and the actions needed to overcome those barriers?

47. Having analysed the DfT Transport Data Strategy and the wider government's wider National Data Strategy, we believe that the barriers have been correctly identified. The actions required are to

overcome these barriers are set out in the two documents, and what would be most helpful is monitoring whether these ambitions are followed through, and if not how to take corrective action.

What is the emerging best practice internationally, in terms both of developing standards and frameworks for sharing and using transport data, and supporting specific innovations? How does the UK compare, and how can it help to shape international standards?

48. The transport data landscape differs substantially internationally. Defining best practice is worthwhile and should reflect the intended use and expected benefits. A focused fact-finding project would be of value. We suggest that such a project should separate at least the following three components: people, movement, network; and report on the earlier mentioned elements of a successful data strategy and its implementation:

- Available data sources and their equivalent in the UK
- Standards
- Quality control
- Storage and security
- Data provision portals and user-friendliness
- Examples of data integration and added value, including how to deal with anonymisation
- Governance and the role of government vs the private sector