improving public transport through technology

Submission to Consultation:

Bus Services Act 2025: improving information about local services

RTIG is the UK community organisation for public transport real-time information technology. Our membership spans local transport authorities, bus operators and the supplier community, together with wider stakeholders in passenger information and transport data.

RTIG is not a trade association acting for a single interest. Our purpose is to improve the quality, interoperability and practical usefulness of public transport data across the sector, and our members sit on every side of the questions this consultation asks. This response reflects that.

It has been prepared through discussion with members, including in consultation workshops. Where members' interests diverge - most obviously on the publication of commercially sensitive operational and financial information, we have set out the full range of views rather than presenting a single position. Where the questions are technical, we have given RTIG's considered view, clearly and without hedging.

We have answered all 37 questions. Several are directed specifically at individual operators or authorities; for those we have summarised the range of member experience rather than answering as though RTIG were itself a registrant, and we encourage DfT to weight individual member responses accordingly. We have said so explicitly at each such question.

Scope note: our response addresses England outside London: in consistency with the scope of the consultation. Nonetheless, several of our recommendations - particularly those on identifiers and data standards, have consequences for Scotland, Wales and London, and we encourage DfT to coordinate accordingly.

RTIG Inform Ltd
36 Fields End,
Oxspring,
Sheffield,
S36 8WH,
United Kingdom

Tel: +44 3330 341 712
email: secretariat@rtig.org.uk

RTIG Inform Ltd
Company Limited by Guarantee
Registered in England, Number 14513703
Registered Office: 36 Fields End, Oxspring, Sheffield,
S36 8WH, United Kingdom

Question 1: Have you ever used the existing registration database? If so, for what purpose(s) have you used it?

Yes. RTIG members use it extensively, and RTIG uses it in its own work. The principal uses members report are:

- Local transport authorities: checking that what is registered matches what is being operated; identifying forthcoming variations and cancellations in order to plan tendered network support and, increasingly, to identify and monitor socially necessary local services; supporting BSOG and BSIP assurance; and cross-checking the registered position against what is published to BODS.
- Operators: confirming the status of their own registrations and monitoring competitors' network changes.
- Suppliers and data consumers: reconciling BODS timetable data against the registered position; building and maintaining service inventories; and diagnosing why a service appears in one source but not another.
- RTIG, consultancies and researchers: network-change analysis, sector evidence and member briefings.

The DfT should be clear what it means by “the existing registration database”: the “Find registered local bus services” search on the Vehicle Operator Licensing service, and the regional registration extracts published on data.gov.uk. For most of the purposes above, our members find the bulk extracts more useful than the search interface, because they can be processed programmatically. That preference is itself a finding, and it should shape the design of the replacement.

Question 2: Are there any features of the existing database you would like to see carried over into a new database?

Yes. The following should survive:

- Bulk, machine-readable extracts, downloadable without an account - or at least an easily created account, which allows for machine-to- machine connections.
- Coverage of every registered service, regardless of whether the operator also publishes to BODS. The registration record is currently the only complete inventory of registered local services, and that completeness is its most valuable property.
- The registration number as a persistent, meaningful identifier -, along with a relationship to the operator's licence number.
- Recording of the operator's licence number, and of the local transport authorities served; the ability to filter by authority area is heavily used.
- Effective dates, and a record of variations against a registration rather than only the current position.
- The service status vocabulary (registered, varied, cancelled, expired), which is useful, although its application is inconsistent (see Question 3).
- Free at the point of use, with no charge and no licensing barrier.

Question 3: Is there anything you would improve about the existing database referenced in Q1? If so, please explain.

There is significant opportunity to improve the existing database. In priority order:

1. **A machine-readable, versioned API with a change feed.** The current extracts are periodic regional files with inconsistent structure; no stable published schema; no per-record last-modified timestamp; and no way to detect what has changed since the last download other than by differencing whole files. This means that every consumer has to build the same inefficient workflows.
2. **Point-in-time query and history.** It is currently very difficult to establish what was registered on a given past date, or to see the sequence of variations that produced the current position. Authorities need this for evidence; analysts need it for any historical tracking.
3. **Geometry and stops.** Registration records describe routes in narrative form - “via” text and place names. There are no coordinates, no NaPTAN stop identifiers and no route geometry, so a registration cannot be related automatically to a timetable, a stop, a corridor or a place. This is the reason that a lot of cross-referencing work has to be performed manually..
4. **Identifier integrity.** Registration numbers, National Operator Codes, licence numbers and the service identifiers used in BODS do not reliably agree, and there is no authoritative link between them. Matching a registration to its published timetable is, in practice, a fuzzy-matching exercise. This is one of the biggest issues.
5. **Data quality and currency.** Cancelled and expired records persist inconsistently; short-notice changes are not always reflected promptly; and practice varies between traffic areas. Consumers cannot tell whether an absence means “not registered” or “not yet processed.”
6. **Completeness.** London is absent, and as the consultation rightly identifies - franchised services and services in EP areas with delegated registration are increasingly absent too. The gap grows as franchising and EP registration authorities expand.
7. **Licensing clarity.** An explicit Open Government Licence statement on the data, with no ambiguity about downstream reuse.
8. **Documentation.** A published field dictionary and schema. The explanatory headings file that exists today is helpful but is not a schema, and does not constrain what is published.

Question 4: Do you agree or disagree with fully digitising the registration, variation and cancellation process for local bus services, as well as including information about services which do not have to be registered with a traffic commissioner (franchised services and those running in an EP area with delegated registration) on a single system?

We have a range of mixed views amongst our members.

On digitisation: the current position, in which a majority of registrations are still made on paper, imposes avoidable costs on operators, traffic commissioners and every authority and information provider downstream. It also means that the authoritative record of a service is created as unstructured text, when structured data describing the same service already exists in the operator's scheduling system.

On coverage: a source of truth that excludes a growing proportion of the network is not a source of truth. As franchising and EP delegated registration expands, the national picture degrades. Bringing franchised and delegated services into the same record, to the same schema, is necessary if the database is to remain useful at all.
On variations: There is a difference between legal registration and information for passengers. There are situations where operational changes and decisions result in data updates which do not require a variation to the registration. Any new systems or processes should not inhibit the provision of updated timetable data which is helpful for passenger information but not of interest to regulators.

Expanding the scope of registration: A single source of bus service public transport information - no matter the legislation it operates under, is a key requirement for data consumers to make their data sourcing decisions as simple and effective as possible. This should include demand responsive and flexible services, scheduled community transport routes and other similar services. It should be mandated to ensure that it is effective.

Three qualifications, which we expand in later answers:

- Digitising a paper process is not the same thing as redesigning it. There is an opportunity to stop asking for information the department will already hold, not to simply reproduce form PSV355 on a screen. We would regard a digital system that requires the same details to be typed in as a missed opportunity (see Questions 20 and 21).
- “A single system” should be a single authoritative record, not necessarily a single monolithic application. The important commitments are one schema, one identifier spine and one published output, not that every function sits in one piece of software.
- The legal act of registration and the publication of information are different things, with different consequences when they fail. They should be linked, and

should share a “front door,” but they should not be confused with each other (see Question 20).

Question 5: Do you agree or disagree with making all of the information in question 4 publicly available on GOV.UK?

Agree, subject to a definition of what “publicly available” must mean, and to proper handling of the personal data within the registration particulars.

An exception could be “closed door” school services where they are door to door or semi door to door where for protection of PII, these may not be suitable for full public access. Those which are bus stop to school when “closed door” should be included to allow journey planner providers help parents to find the public transport options for a particular school.

For this to deliver the benefit intended, publication should mean all of the following:

- Open Government Licence, stated explicitly on the data itself, with no ambiguity about commercial reuse.
- A documented, versioned API, with no login requirement for read access and rate limits set *for* legitimate bulk use rather than *against* it.
- Bulk download of the complete dataset, plus a daily change file.
- A published, stable schema and field dictionary, versioned, with a deprecation policy and advance notice of change.
- Full history, with point-in-time query.
- A human-readable search interface built on the same API: so that what the public sees and what a machine can retrieve cannot diverge meeting accessibility requirements.

Publishing only an HTML search page would satisfy the wording of the proposal whilst missing almost all of its value. Most of the benefit this consultation seeks arises from reuse, and reuse requires machine access.

On personal data: the registered particulars for flexible services include the operator's address and the arrangements for booking; and for sole traders and small operators that address is frequently a residential one. DfT should complete and publish a data protection impact assessment covering the full published dataset, and should set a redaction policy before, not after, launch. We note that Section 24 of the Act expressly permits regulations about the use and disclosure of information provided; we encourage DfT to use that power to set the publication scope explicitly in regulations - rather than leaving it to be settled by system design.

Question 6: Do you agree or disagree with making the provision of information about services in EP areas with delegated registration and franchised services, also fully digitised?

Agree. If these services are provided by any other route, the single record is not single, and every consumer is back to reconciling multiple sources, which is the problem the consultation sets out to solve.

Three points on how, which matter more than the principle:

- For franchised networks, the underlying event is not “an operator registers a service” but “an authority specifies a network.” A franchised network change could affect dozens of services with a single common effective date. The digital process must support bulk submission of a whole network - staged and reviewed before it takes effect, and not dozens of individual applications.
- EP delegated registration exists in part to allow local flexibility over notice periods and process. Digitisation should carry that flexibility, not quietly remove it by encoding a single national workflow.
- The authority's system of record is typically a contract management, scheduling or network planning system, not a registration system. The route in must therefore include an API and a bulk file upload, not only a web form. Authorities will need lead time and a stable specification for their suppliers to build that.

Question 7: Do you agree or disagree with making the information provided by LTAs with delegated registration functions, and franchising LTAs, consistent with that provided by operators registering their services with the TCs?

Agree with the principle of a common information model: the same particulars, the same definitions, the same identifiers and the same schema, whoever provides them. Consistency of content is precisely what makes a combined database usable, and inconsistency is what makes the present position unusable.

We do, however, have two qualifications to this.

- Consistency of content is not the same as identity of obligation or process. A franchised service is not “registered” in the same legal sense as a commercial service; the authority is not seeking permission, and there is no traffic commissioner decision to make. The information should be the same; the workflow, the notice arrangements and the legal consequences legitimately differ, and the system should model that - rather than pretending otherwise.
- Consistency should be achieved through a single published schema with validation on submission and not through a shared form design. If the schema is the normative artefact, an authority can meet it from its own systems and an operator can meet it from its scheduling software; and both produce comparable records. If the form is the normative artefact, everyone is forced through the same interface regardless of fit.

We would also note a practical point: some particulars have no easy equivalent in a franchised context, and a few have no equivalent for flexible or demand-responsive services. DfT should work these through explicitly with a small group of authorities and operators before the schema is fixed, rather than leaving them to be resolved in guidance.

Using the new data provision system

Question 8: As outlined earlier in this document, the government intends to create a new data provision system, bringing together information about franchised services, services operating in enhanced partnerships where the registration function has been delegated to the LTA, and services registered with a traffic commissioner. This will be published online.
If you are an operator of bus services, a franchising authority, an authority exploring franchising, or are responding on behalf of an LTA with a delegated registration function, what would you like the new database to allow you to do? If you are not responding on behalf of these bodies, please continue to question 10.

RTIG is answering on behalf of members in all of these categories. The priorities below are ordered by how frequently members raised them.

- **Bulk and programmatic access**. A documented REST API and full bulk download, with simple authentication required: only complex for write operations. Almost every member use case involves processing many records, not looking at one.
- **Submit once**. Derive the registration particulars from the TransXChange data that is already prepared and published, with the submitter confirming rather than re-keying. Where a variation is a timetable change, the TransXChange file is the change; asking for it again in narrative form adds cost and introduces divergence.
- **A forward view and change notification**. The need to see everything taking effect within a date window, filtered by authority area or operator, with a subscribable change feed. Authorities currently rely on manual checking and on email from operators, and things are missed.
- **Point-in-time reconstruction**. This is essential for monitoring socially necessary local services under section 14 of the Act, for BSIP evidence and reporting; and for any assessment of how a network has changed.
- **A reconciliation view across registration.** BODS timetables and location data in a single place that identifiy a service registered but not published, published but not registered, or registered and published with materially different content. No system can produce this today, and many consumer and data providers build a partial version of it.
- **Status tracking and a complete audit trail through the application lifecycle.** This includes short-notice applications and the reasons given, and traffic commissioner decisions — visible to the operator, the agent and the relevant local transport authority.
- **Delegated and agent access.** Many small operators use agents or consultants, and many authorities need several officers to act. Role-based access must support this properly, so that shared credentials are not the

workaround. The increased used of agents and bureau services is likely to be required to support any future requirements.

- **A test and sandbox environment**, so that submitters and their suppliers can validate before the statutory clock is running.
- **Publishing workflow and versioning**, Draft, staging and scheduled publication, and bulk operations — particularly for franchising authorities managing whole-network change, and for any submitter making a coordinated set of changes with a single effective date.
- **Integration with the authority's own publication duties,** so that an authority does not have to publish the same information separately on its own website in a format of its own devising.

Question 9: If you are a passenger, a bus data consumer, or an organisation which provides information to passenger, what would you like the new data provision system to allow you to access, and how?

A large part of RTIG's membership consumes this data or builds the systems that do. This is what our community is asking for:

- An open API, under the Open Government Licence: documented and, versioned, with a **deprecation** policy and generous rate limits.
- Bulk download of the complete dataset and a daily change file, so that consumers can maintain a local copy without polling.
- Stable identifiers that join to BODS. This is the single most valuable feature the new system could offer, and the one that cannot be retrofitted. If a registration record and its BODS datasets share identifiers by construction, the need for a whole category of manual reconciliation work disappears across the sector.
- Spatial data, NaPTAN stop identifiers and route geometry: so that a registration can be placed on a map, related to a place, and matched to a timetable automatically.
- A complete historical archive with point-in-time query, retained indefinitely.
- A machine-readable and human readable change feed supporting “what is changing in this area, in this window”.
- Precise, documented status semantics and effective dates: so that a consumer can determine with confidence whether a service is running on a given date.
- Published data quality and completeness metrics: as data - so that consumers can assess fitness for purpose programmatically and exclude known bad data sources - rather than silently propagating them to passengers.
- Accessible, human-readable pages built on the same API – to provide a clear public view of services being withdrawn or varied, which is the information sought most often by passengers and the one the current service addresses least well.
- Access to current operator vehicle lists to enable vehicle accessibility, equipment and facility information to be provided to passengers.

Using open data from BODS

Question 10: Have you ever used open data from the existing BODS? If so, for what purpose(s) have you used it?

Yes, extensively. BODS is central to the work of a large proportion of RTIG's membership. Uses reported include:

- Journey planning and passenger information, including national and regional journey planners and third-party apps.
- Real-time information at stops, in interchanges, on board and in apps.
- Network planning, service review and BSIP development and evidence.
- Performance and reliability monitoring by authorities and operators.
- Bus priority and traffic signal work, and highway scheme appraisal.
- Accessibility and equality analysis, including access to services from particular communities.
- Integration with wider multi-modal information.
- Academic research, consultancy analysis including the “Analyse BODS” tooling.

Question 11: Are there any features of the existing BODS you would like to see carried over to a new database?

Yes. BODS already gets a number of important things right:

- Open, no-login access under the Open Government Licence - to both APIs and bulk downloads.
- The separation of timetables, fares and location data, each expressed in a standard appropriate to it, rather than bespoke formats.
- The use of published UK and international standards: TransXChange with the PTI profile, NeTEx for fares, SIRI-VM for location rather than a departmental format. This is what allows an international supplier market to serve UK operators, though some suppliers have increasing concern about continued use of TransXChange though operators and authorities remain supportive.
- Derived GTFS and GTFS-RT outputs, which have materially widened the use of the data.
- The data catalogue files, which give consumers a useful source of what exists.
- The principle of validation at the point of publication, and the SIRI-VM validator.
- The AVL-to-timetable matching quality reports - though far from perfect, made a long-standing problem visible and measurable for the first time.
- The public changelog and release notes.
- Ingestion of Street Manager roadworks into disruptions data, though much more could be made of this.
- The organisational model anchored on the National Operator Code.

Question 12: Is there anything you would improve about the way open data are made available through the existing BODS? If so, please explain.

Yes. Our members' consolidated view, in no particular order:

- **Data quality** remains the single biggest issue, and matching quality above all. Location data that cannot be reliably matched to a timetable cannot produce an accurate prediction, and match rates remain materially below where they need to be. The published quality reports were an important step; the missing element is compliance follow-up.
- **Enforcement and accountability**. Publication is a legal obligation, but non-compliance and poor-quality compliance are rarely visible. Operators who invest in good data are, in effect, at a disadvantage to those who do not.
- **Reconciliation with the registered position**. There is no authoritative statement of what should be published, so there is no way to measure completeness. This consultation's proposals could resolve that, and in doing so would be among the most valuable outcomes available.
- **Identifier stability**. Some publishing toolchains regenerate service and journey identifiers on every republication, which breaks every downstream join and every cached relationship. Identifier persistence should be a validated requirement.
- **Timeliness**. Data is not always published sufficiently in advance of the effective date. Data published on the day a change takes effect has already failed the passenger.
- **Fares data**. Completeness and correctness are well below that of timetable data, and consumption remains very limited. We would encourage DfT to focus on making existing fares data usable before requiring more of it.
- **Disruption data**. Uptake is low, and the Create Transport Disruptions Service is a separate journey from BODS with a separate account and separate guidance. Convergence would help - as would ability of operators to access the system.
- **Versioning and archive**. Historical BODS data is not systematically retained and published, so a large class of analysis - including much that authorities need for BSIP evaluation, is simply impossible and/or costly for multiple suppliers to store.
- **Granularity of discovery**. Discovery is dataset-oriented rather than service-oriented; finding the current data for a particular service is harder than it should be.
- **Derived GTFS** outputs are treated as secondary. For a large part of the developer community they are primary. They should be validated, documented and covered by the same service commitments as the source formats.

- **Change management.** Specification and platform changes have at times been communicated with insufficient lead time for suppliers to respond across their customer base.

Question 13: One option for development of the new database is to publish the data contained within it via the existing BODS and link the 2 together, to create one source of information about local bus services across England outside London.

Yes - as a template for publication. The BODS publication pattern is proven and we would not want DfT to build something structurally different for its own sake. Specifically, the pattern worth reusing is: open licence, no-login API alongside bulk download, a data catalogue as a manifest, published validators, published quality reporting, and a public changelog. That combination is good practice and is well understood by the UK bus data community.

Three qualifications.

1. Template the pattern, not the implementation. Where BODS falls short, in for example, versioning, historical archive, service-level granularity of discovery, and quality enforcement - , the new database should be built to do better from the outset, not to inherit those characteristics because they are familiar.
2. Link the two; do not merge them. The registration record is a regulatory artefact with legal effect, a decision history and an audit trail; it needs retention rules, correction processes and evidential integrity that operational open data does not. BODS is an operational publication optimised for currency and reuse. Two stores with one identifier spine: cross-referenced in both directions, is the right architecture.
3. Publish the reconciliation between them. The most valuable new feature this pairing can produce is an authoritative characteristic public statement of which registered services have compliant open data and which do not. Neither system can produce that alone; together they can, and it would do more for data quality than any amount of additional guidance.

Records of registration

Question 14: If you are responding on behalf of a local transport authority (LTA) in an enhanced partnership (EP) area with a delegated registration function – in what format do you presently hold the information required under section 6I of the Transport Act 1985?

This question is properly for individual authorities, and we have encouraged members with a delegated registration function to answer it directly. RTIG's observation, from discussion with those members, is that practice varies widely and that no member reported holding the full particulars in a single structured system.

The typical pattern is a combination of:

- TransXChange files for route, stopping places and timetable – with the structured element held in a scheduling or passenger information system.
- Spreadsheets, or a case management or CRM system for the administrative particulars: operator identity, licence number, effective dates and notice tracking.
- Correspondence in email and PDF for the application and decision record.
- Narrative route descriptions and maps held as documents, frequently reproducing information that already exists in structured form in the TransXChange file.
- Separate publication on the authority's own website - in a format of the authority's choosing.

The practical consequences are that the same information is held more than once, not all of it is machine-readable, and there is no consistent structure between authorities. There will be a wide range of starting points and a common baseline should not be assumed when estimating the effort of transition.
We would encourage a survey of delegated authorities directly on this point before finalising the schema.

Question 15: If you are responding on behalf of a franchising authority, or an authority exploring franchising – do you already hold information about your services, similar or identical to that required under section 6I, centrally? If, so in what format?

RTIG's observation is that franchising authorities generally do (or will) necessarily hold this information centrally, since they specified the services but hold it in a different shape and for a different purpose.

Typically, it sits in contract management, scheduling and network planning systems, and in the contract documents themselves. Route, stopping places and timetable exist as TransXChange or in a scheduling system's native format, because they must in order to produce passenger information. The differences from a registration record are structural:

- The data is organised by contract or lot, not by “registration”, and the boundaries do not always coincide with what would be registered as a service.
- Some particulars have no natural equivalent. For example, the identity of the operator is determined by contract award and may change during the life of the specified service - without the service itself changing.
- Change is managed through contract variation processes on the authority's own cycle, which does not map onto registration notice periods.
- Some of the information is commercially sensitive while a procurement is live.

Authorities exploring franchising will generally hold considerably less, and what they do hold will be in assessment and modelling artefacts - rather than operational systems.

The design implication is that the requirement should be expressed as an outcome; this information, to this schema, by this date with a bulk and API submission route - rather than as a process that replicates operator registration step by step.

Existing electronic bus service registration system (EBSR) and fully digitising registration, variation and cancellation

Question 16: Have you ever used the traffic commissioners' EBSR system to register, vary or cancel a bus service? If so, are there any existing features or capabilities of EBSR that you'd like to see carried over to a fully digitised system?

Members with EBSR experience report a mixed picture, but there are features that should certainly survive:

- Submission of the registration as structured TransXChange data rather than as a form. This is EBSR's core good idea, it is right, and it should be carried forward.
- Automatic derivation of the registration particulars from that file, so the operator describes the service once.
- The principle that the same underlying data can serve both the registration and the passenger information publication is not possible with a single submission, as operational data is rarely available at time of registration.
- Electronic notification to the relevant local transport authorities as part of the submission.
- A machine-readable acknowledgement carrying the registration reference.
- Third-party and agent submission on an operator's behalf.

What should not be carried over is the requirement, implicit in EBSR today, that an operator must own TransXChange-producing software in order to use the electronic route at all. See Questions 18 and 19.

Question 17: Is there anything that puts you off using the existing EBSR system?

The reasons members give: broadly in the order of how often they are raised:

1. It requires TransXChange-producing software. For an operator with a handful of services this is a significant cost with no other business justification, and it is the primary barrier.
2. The EBSR TransXChange profile differs from the BODS PTI profile as well as having different schema basis, so the same service must be prepared twice in slightly different forms. This is the clearest possible argument for integration.
3. Validation errors are hard to interpret and hard to resolve without supplier assistance, and the feedback loop is slow.
4. Perceived risk. A rejected electronic submission close to the 42-day deadline is a commercial problem. Paper feels safer.
5. The user interface and workflow show their age, and authentication is dated.
6. Limited visibility of status after submission, so the operator does not know where the application has got to.
7. Limited perceived benefit to the submitter. The effort falls on the operator; most of the benefit accrues to authorities and the department. That asymmetry is worth acknowledging in the design of the replacement.

Question 18: Is there anything you would improve about the existing EBSR? If so, please explain.

Since DfT proposes to *replace* - rather than *improve* EBSR, we have framed this as what the replacement *must do* differently.

- One profile, one file. Align the registration data profile with the BODS PTI profile so that a single TransXChange file serves both purposes. If the two profiles diverge, the duplication that deters operators today will simply persist under a new name.
- Revise the charging regime as the current approach directly causes some of the current BODS data issues - as it encourages compiling multiple lines into a single registration.
- Validate early and explain clearly. This requires field-level, plain-English validation messages linked to the relevant guidance, available before submission through a sandbox, so that an operator can resolve problems without the statutory clock running.
- Provide a route that does not require the operator to own scheduling software. A guided web journey that produces conformant TransXChange on the operator's behalf is, in our view, the single most important design decision in this programme. Without it, the large number of small operators will not move, whatever the regulations say.
- Status transparency and a full audit trail, visible to operator, agent, authority and traffic commissioner.
- Proper agent and delegated access, with role management, so that consultants and group head offices can act without sharing credentials.
- Modern authentication — GOV.UK One Login — and modern accessibility.
- An API for high-volume submitters, including group operators and franchising authorities.
- Pre-population from the current registered position, so that a variation is an edit rather than a re-entry. Most variations change one element of an otherwise unchanged service.

Question 19: If you currently use paper applications to vary, cancel or register a service, why do you choose this over EBSR?

This is a question for individual operators and we have encouraged members to answer it from their own experience. The reasons reported to RTIG are:

- No scheduling software, and no business case to buy it for a small number of registrations a year.
- The application fee is small relative to the cost of software or consultancy support, so the economics favour paper.
- Familiarity, combined with low frequency of change. An operator making one or two changes a year has little opportunity to become fluent in any system.
- Risk aversion close to deadlines - particularly where a rejected submission would jeopardise a commercial change.
- Agents and consultants handle the submission and use whichever route they prefer.
- Digital confidence and, in a small number of cases, connectivity.
- There is no compulsion. The paper route is accepted, so there is no forcing function.

The important implication for this consultation is that the barrier is the requirement to produce structured data, not the electronic channel itself. A digitised system that assumes the submitter has TransXChange capable software will move the problem rather than solve it. Provide a data-creation route within the service, and much of the paper problem resolves itself.

Question 20: One option for the development of a fully digitised system used to populate the new database is to integrate it with the data submission process for the existing BODS and link the 2 together to create a single registration and BODS data submission journey for local bus services across England outside London.
If the government proceeds with this approach, do you think the existing BODS system is a good template on which to base the submission of data for the new database? Or would a new or different approach be favourable?

Partly — and less straightforwardly than for publication, which we addressed at Question 13.

What is right about it: BODS already receives data from most of the relevant operators; already has publisher accounts anchored on the National Operator Code; already validates at the point of submission; and already holds the same underlying TransXChange data from which the registration particulars can be derived. A single submission journey is the right ambition.

What is different about registration, and must be respected:

- Registration has legal effect and runs against a statutory clock. Publication does not.
- Registration needs an evidential audit trail: what was submitted, when, by whom, what was decided and on what basis; and a record capable of supporting an appeal.
- Registration involves a third party as decision-maker: the traffic commissioner - with the local transport authority as a notified party. BODS has no equivalent.
- The failure modes differ in kind. A BODS validation failure produces poor passenger data. A registration failure means a service cannot lawfully operate.

Our recommendation is therefore: a single submission journey, yes; a single monolithic application, perhaps not. One front door and one file - feeding two stores with different guarantees, different retention and different management processes. One design rule follows from this and we suggest it is adopted: a publication-quality validation failure must never be able to block a legally valid registration. Separate the test “is this a valid registration” from the test “is this good passenger data;”, report both to the submitter; and let the registration proceed while the data quality issue is pursued through a different route. Coupling the two would make a data quality tool into a barrier to lawful operation, which would be both disproportionate and, we suspect, quickly become subject to challenge.

Question 21: What, in your opinion, would be the advantages of basing the submission of data for the new database on the existing BODS system and creating a single registration and BODS data submission journey?

- It eliminates the duplicate preparation of the same information in two different profiles, which is the single largest burden reduction available in this programme, and one that costs operators real money today.
- It guarantees consistency between what is registered and what passengers are told. Divergence between the two is the root cause of a substantial proportion of current data quality complaints, and this design makes such divergence structurally impossible - rather than merely discouraged.
- It creates the identifier spine automatically. If the registration and the published dataset originate from the same submission, they share provenance and can be joined with certainty. This resolves a problem that the sector has worked around for a decade and cannot resolve on its own.
- One account, one set of credentials, one support route, one body of guidance, one place to look.
- Compliance becomes measurable. DfT would be able to state definitively which registered services have compliant open data, which is currently impossible.
- Supplier effort is roughly halved: one profile to implement, one validator to satisfy, one integration to maintain. That reduces cost across the whole customer base, and it reduces the time it takes the sector to absorb future change.
- Timeliness improves. Publication happens as a consequence of registration rather than as a separate later task, which is currently where it is most often delayed or forgotten.
- Authority effort in reconciling the two sources largely disappears.
- Franchised and EP delegated services enter the same pipeline on the same basis, so the national picture is complete for the first time.

Question 22: What, in your opinion, would be the drawbacks of basing the submission of data for the new database on the existing BODS system and creating a single registration and BODS data submission journey?

1. It couples a statutory process to a data platform. Availability, release cycles and validation changes acquire legal consequences. Whilst this is manageable, it requires an explicit service level, a documented fallback route for when the service is unavailable close to a deadline, and the separation of registration validity from data quality that we set out at Question 20.
2. Scope and delivery risk. BODS is a relatively mature service with a large installed base; adding a regulated workflow is a substantial change and carries a real risk of destabilising something that currently largely works.
3. The user populations are not the same. Not everyone who registers currently publishes, and not everyone who publishes - registers. Internally within large operators, the users are not the same from registration to operation. As it stands, the long tail of paper registrants is precisely the group least equipped to engage with BODS so the integration must not assume BODS familiarity but rather, address the barriers.
4. Governance. Traffic commissioners are independent statutory office holders; DfT would own and operate the system through which part of their function is exercised. Roles, decision rights, change control and escalation need to be explicit from the outset.
5. Change management becomes harder and must become slower. Once a specification change affects a legal process, the acceptable rate of change drops and the required notice period lengthens. That is a real constraint on future improvement and should be acknowledged as such – rather than discovered later.
6. Profile compromise. Registration and publication may require different things from the same file. A single profile risks serving both imperfectly; this needs to be worked through in detail rather than assumed away.
7. Data protection. Registration records contain more personal data than open timetable data. Merging the journeys makes accidental over-publication easier and exacerbates the consequences of a mistake.
8. Stranded investment. Suppliers and larger operators have invested in EBSR integration. Migration needs time and, for smaller organisations, funding.
9. Devolved divergence. Scotland and Wales register through the same traffic commissioners but sit outside these provisions. A single England-only journey necessitates the creation of a second system for cross-border operators to maintain, which will increase the administrative burden for them.

Question 23: For LTAs in EP areas with a delegated registration function, authorities exploring franchising or franchising authorities who could be required to input information about their services via a fully digitised system – do you foresee any barriers or issues with inputting registration, variation and cancellation information regarding your bus services onto a GOV.UK online portal like BODS or EBSR? Please describe.

Yes. Members identify the following, and we would encourage DfT to treat them as design constraints rather than as transition risks to be managed later:

- Resourcing and skills. This is a new statutory duty landing on LTA teams that are already stretched, and in franchising authorities it lands at the same time as the standing up of substantial new functions. It has been suggested that without new burden funding being made available, it will compete with delivery - particularly in smaller non-metropolitan authorities.
- Volume and timing. A franchised network change affects the whole network at once, with a single effective date. A per-service web form is unworkable at that scale. Bulk upload, API submission, draft and staging capability, and coordinated effective dates are essential rather than desirable.
- System of record mismatch. Authority data exists in contract management, scheduling and asset systems, not in a registration format. “Extract and transform” work is needed, and suppliers need a stable published specification and lead time to build it.
- Concept mismatch. Several registered particulars do not map cleanly onto a franchised network; and the boundaries of “a service” in a contract may not match the registration concept.
- Access management. This will involve multiple officers, multiple teams, and delegated access for the authority's operators and contractors with an approval step - before anything becomes public.
- Process alignment. EP scheme rules and contract variation cycles are set locally. A national workflow that assumes the traffic commissioner registration process will not fit them without configurability.
- Commercial confidentiality. Network specifications may be commercially sensitive before contract award, and a national portal must not force early disclosure.
- Assurance and accountability. An authority publishing to a national database will want a documented internal review and sign-off step, and a way to correct errors quickly, once published.
- Cost, and whether funding will follow the duty.

We would encourage DfT to work on design with a small group of franchising and delegated authorities from the outset, and to allow phased onboarding rather than a single national commencement date. RTIG would be glad to help convene that group.

Question 24: For operators of bus services who currently register, vary and cancel them with the traffic commissioners via the paper registration process – how much lead-in time do you anticipate you will require to move from paper to electronic registration, variation and cancellation applications?

This is properly a question for individual operators and we encouraged members to answer it directly with reference to their own circumstances. From a sector point of view, the answer depends almost entirely on one design decision: whether the new service provides a route to create conformant data within the service, or whether it assumes the operator will obtain scheduling software.
RTIG's recommendation on a reasonable overall lead-in:

- At least 12 months before the date on which electronic submission becomes mandatory.and publication of the final, frozen technical specification.
- A sandbox and validator available from at least 6 months before the deadline, preferably 12 months, not at the end of it.
- A parallel-running period of at least 6 months during which both routes are accepted.
- At least one complete change cycle of live experience for small operators before the paper route closes. An operator making two changes a year needs the transition period to span more than one of them.

If DfT provides a guided data-creation route, 12 months is achievable for most of the sector. If operators must procure and learn scheduling software, 18 to 24 months is more realistic for the large number of smaller operators, some of whom will need direct support rather than time.

Functionality of the fully digitised system

Question 25: What functionalities would you find most useful in a new GOV.UK portal? This can be with reference to either administering bus registrations, variations and cancellations, or navigating the new portal once it is available to access information about bus services.

Administering registrations, variations and cancellations

- Submit once: a single TransXChange file serving both the registration and the open data publication, with the particulars derived from it.
- A guided, no-software route that produces conformant data for operators without scheduling systems.
- Pre-submission validation against a sandbox, with field-level, plain-English error messages linked to guidance.
- Pre-population from the current registered position, so that a variation is an edit rather than a re-entry.
- A visible statutory clock: the earliest permissible effective date calculated and displayed, notice periods tracked, and short-notice applications supported with structured reasons.
- Automatic notification to the relevant local transport authorities: with the authority able to record its response in the same place.
- Full status tracking and audit trail, visible to operator, agent, authority and traffic commissioner.
- Bulk and API submission for group operators and franchising authorities - with coordinated effective dates across many services.
- Draft, staging and scheduled publication.
- Proper agent and delegated access with role management, and GOV.UK One Login.
- Machine-readable receipts and decisions.
- Integrated fee payment where fees apply.
- Withdrawal and correction routes that do not require starting again.

Navigating the portal and accessing information about services

- An open API and bulk download with a documented, versioned schema.
- Point-in-time query and complete history.
- A change feed, with subscribable alerts by area, operator or service.
- Map-based search built on NaPTAN stops and route geometry.
- Filtering by authority area, operator, status, effective date and service type — commercial, franchised, EP delegated, school, demand-responsive.
- Direct links in both directions, between a registration record and its corresponding BODS datasets.
- A compliance and reconciliation view showing where registration and published data disagree.

- Data quality and completeness dashboards, published as data as well as pages.
- Accessible human-readable pages meeting accessibility requirements, built from the same API as the machine interface.

Data transparency

Question 26: Do you agree with the general concept of increased data transparency around local bus services (going beyond existing initiatives)? Do you foresee any opportunities or challenges across any of the following areas:

- **areas where passengers would benefit from additional information being openly available, such as likely occupancy, accessibility or environmental features (for example, to aid journey planning), as selected operators already provide in their apps**
- **areas where additional information being openly available would help monitor effective bus operation and local network performance or make analysis more meaningful**

RTIG supports the principle. Better data about bus services has consistently produced benefits our members can point to, and the case for more of it is strong. We would offer the following on the two areas the question identifies.
Information that would benefit passengers include occupancy, accessibility and for some, environmental features
The opportunities are real, particularly for disabled passengers, and they support the accessibility duties elsewhere in the Act. The challenges are, however, serious and we would ask DfT to consider them fully at the design stage - rather than in guidance.

- Accuracy matters more than availability, and for accessibility data it matters absolutely. Wrong accessibility information is worse than no accessibility information: a wheelchair user who decides to travel because the data said there was a wheelchair space, and finds there is not, has been actively harmed by the data. Any accessibility data requirement must therefore, come with a duty of accuracy, an accessible correction route, and published provenance and timestamping so that consumers can judge reliability.
- Improved coordination with other teams within the DfT which are developing accessibility requirements as well as those collecting data about bus services from LTAs and operators.
- Vehicle-level attributes are only meaningful if the vehicle is *reliably* associated with the journey. Fleet allocation changes daily and at short notice and so a static statement that “this route has USB charging” is frequently false in practice. The useful version needs to be dynamic: which vehicle is operating this journey, now; and it depends entirely on the matching problem discussed at Questions 27 and 31 being solved first. Mandating the attributes without fixing the matching would produce data that looks informative and is not.
- Occupancy data quality varies enormously with the counting technology used, and there is no agreed UK definition of “likely occupancy”. Publishing an unqualified figure invites misuse and will erode trust when it is wrong.

- Definitions must be standardised before collection. “Wi-Fi” and “USB charger” mean materially different things to different operators. NeTEx already carries a facility and accessibility vocabulary, and DfT should adopt it rather than creating new enumerations.

Information that would help monitor operation and performance, or make analysis more meaningful

The opportunity is substantial. Consistent, comparable data on network change, service delivery and reliability would improve BSIP evaluation, support the identification and monitoring of socially necessary local services, and make value-for-money assessment of public support far better founded than it is today.

The principal challenge is comparability, which requires common definitions and a common methodology; without them, published figures will be compared when they are not comparable. The sector will then spend its time disputing the numbers rather than acting on them. We would ask that provenance, methodology and quality metadata be published alongside every dataset, as data.

One general point. DfT already receives a good deal of this information through BSOG and BSIP returns, the annual bus statistics survey, and existing section 141A returns. Every new requirement should be tested against what is already collected, and should replace, rather than supplement it - wherever possible.

Question 27: DfT's view is that there could be greater transparency on the vehicles used to operate a local bus service (including their on-board facilities and technical specification), the type of service being operated, the costs of operating the services they run, the number of staff engaged in doing so, the distance travelled by vehicles, the number of journeys made and the number of passengers on services.
Additionally, it sees opportunities to aid journey planning through the surfacing of journey cancellations and diversions in real time to improve passenger confidence.
With reference to the list of data which we propose mandating operators to provide and which would be published by the department, is there specific data you think DfT should, or should not, be requiring operators to provide or make public?

We deal with the technical items and the commercial items separately. On the technical items, RTIG takes firm positions. On the publication of commercially sensitive financial and patronage data our members have divergent views and RTIG deliberately takes no position on whether - only on how.

Technical items: RTIG's positions
We strongly support new technical information to aid matching of timetables to live location feeds and identification of vehicles, and we would put this first in the whole list.

This is the single most valuable item proposed and is the enabler for most of the rest. It also addresses a problem the sector has documented for years: only a proportion of BODS location data can be matched to timetable data using the recommended strategy, and match rates remain a material constraint on prediction quality today. Consistent VehicleJourneyRef, BlockRef and DatedVehicleJourneyRef values, and an unambiguous persistent vehicle identifier, would transform arrival prediction and would make vehicle-level attributes usable for the first time. We recommend that exact-match consistency between equivalent SIRI-VM and TransXChange fields be made a validated, enforced requirement rather than a recommendation; and that BlockRef in particular be mandatory - as without it, predictions for a journey can only begin once that journey has started, which is precisely when passengers waiting at the first stops most need them.

On the vehicle registration mark specifically, we support a mandatory persistent vehicle identifier but would ask DfT to consider carefully whether that identifier needs to be the registration mark in the public feed. A vehicle registration mark broadcast continuously with location is capable of identifying an individual driver's movements over time, and is likely to constitute personal data. A stable pseudonymous fleet identifier in the public feed, with the mapping to the registration mark held by DfT and made available to authorised users, achieves the operational purpose without that risk. A published data protection impact assessment should precede any mandate.

Cancellations and diversions in near real-time strongly support the outcome; do not build a third mechanism.

Much of this capability already partially exists. BODS has released basic non-standardised operator cancellation publication functionality, and the Create Transport Disruptions Service already produces SIRI-SX disruption data for non-rail modes in England and ingests Street Manager roadworks. The gap is uptake, consistency and consumability - rather than capability. We recommend that DfT mandate the use of the existing routes; converge the two submission journeys; define the semantics precisely (a fully cancelled journey, a part-cancelled journey, a short working, a diversion with stops not served, a stop temporarily unavailable); and ensure that the outputs are available in both the SIRI family and GTFS-RT.
Sub-five-second automatic vehicle location — whilst we support an increase in cadence, we do not support this in isolation - as substantial changes are needed elsewhere for this to be truly useful.

Frequency is not necessarily the right lever. The accuracy of an arrival prediction is dominated by journey identification and matching quality, not by position update rate: a five-second feed that cannot be matched to a timetable is worth considerably less than a fifteen-second feed that can. The costs are substantial and the impact should be considered, as they will fall on every vehicle as mobile data volumes rise in direct proportion to cadence; telematics contracts are frequently priced on message volume; and ingest and distribution load rises for BODS and for every consumer. Contracts would need renegotiation and in some situations, changes to on vehicle equipment would be required so investment cycles need to be considered. SIRI-VM's current request-response polling pattern does not scale gracefully to that rate so a change to subscription models would likely be required.

Our recommendation is to tighten the existing requirement over time: to allow long term contracts to be renegotiated; and invest immediate effort into matching quality and journey identification. If higher frequency is genuinely required, change the delivery mechanism first to a subscription, push or streaming pattern and pilot it with willing operators and authorities to ensure that appropriate business cases can be made before mandating it. RTIG would be pleased to work with DfT on evidencing this, and we think a well-designed pilot would help.
Vehicle-level facilities and specifications (USB chargers, Wi-Fi, audio-visual equipment, wheelchair spaces, priority seats, CCTV, vehicle type and capacity, fuel type, fuel efficiency, emissions, mileage by fuel type) we support in principle, with a structural recommendation.

Do not carry these attributes in the timetable feed. Establish a vehicle or fleet registry keyed on a persistent vehicle identifier, published once and referenced from the location feed. These attributes change on refurbishment and fleet transfer, not on

timetable change; repeating them in every TransXChange file guarantees they will be inconsistent, stale, or both. Use the existing NeTEx facility and accessibility vocabulary rather than new UK enumerations.

We would also ask DfT to check what is already held before collecting. Vehicle and licensing data sits with DVSA; accessibility compliance data exists under the PSVAR regime; emissions and fuel data exists in ZEBRA and BSOG reporting. Collecting the same facts again through a different channel produces an additional burden.
We support information about the type of school service being operated. This is a small, stable, well-defined attribute that resolves a genuine and long-standing ambiguity for authorities and data consumers, and it is cheap to provide. We would ask for care in publishing anything that identifies the specific schools served by closed services, or that could identify the children travelling.

We support route and patronage data for demand-responsive transport – but with care. Demand-responsive services have no fixed route, and the data model must reflect that rather than forcing DRT into a fixed-route shape. NeTEx Part 5 (alternative modes) and the flexible service constructs in TransXChange are the right starting points. Patronage data on very low-volume DRT services runs the risk of identifying individuals, and thresholds should be applied.

Passenger journeys and trips at individual service level, including concessionary travel is technically feasible; see below regarding publication. The technical prerequisite is a common definition of a passenger journey (boardings against linked trips); consistent treatment of through-ticketing; multi-operator products and interavailability; and a stated methodology for allocating journeys to services. This will almost certainly require consistent unique journey numbers to be created. Without those, service-level figures will not be comparable between operators and should not be presented as though they were.

Commercial items: RTIG takes no position

On the publication of operating costs, staff numbers, government support split by funding stream, and service-level patronage, our members are divided. We set out both views rather than adopting one.

Members who support publication argue that these services are substantially publicly funded and the public is entitled to see what that money buys; that authorities need this information to assess value for money, design networks and identify socially necessary services; that comparability across operators would improve commissioning; and that existing aggregate publication is too coarse to support any of this.

Members who oppose publication argue that cost and patronage data at service level is genuinely commercially sensitive in a market with live tendering and franchising competitions; that publishing it would inform competitors' bidding and could raise the cost of provision to the public purse; that patronage data is a core commercial asset developed at the operator's own expense; and that the figures are capable of serious misinterpretation without operational context that only the operator holds.

RTIG's contribution is on how, not whether:

- Distinguish clearly between “provide to DfT” and “publish”. These are separate decisions, the Act's powers permit them to be taken separately, and combining them removes a middle ground. See Question 28.
- If cost data is collected, publish the cost allocation methodology first, and require it to be followed. Operators allocate overheads, depot costs and vehicle capital very differently. Figures produced on different bases are not comparable, and publishing them as though they were, will produce misleading conclusions and consume a great deal of everyone's time in dispute.
- Consider aggregation thresholds and time lags as a middle path — publishing at a level of aggregation and with a delay that supports accountability without exposing live commercial position.
- Reuse existing returns. Much of this is already reported through BSOG and BSIP processes and the annual bus statistics survey. Requiring it again in a different shape adds additional burden for no additional information.
- Whatever is decided, decide it explicitly in regulations rather than allowing it to be settled by system design or by default.

Items we would question

- Fuel efficiency and emissions of vehicle. We support the intent but ask DfT to be precise about what is meant: a type-approval figure, a fleet average, or measured in-service consumption. Type-approval figures are largely derivable from vehicle records and add little; measured in-service data is far more useful and far more burdensome. State which, and why, and check the overlap with the zero-emission provisions in sections 37 and 38 and with existing ZEBRA and BSOG reporting.
- Mileage of vehicles by fuel type. This is largely derivable from the location data and fleet registry that DfT would already hold under these proposals. Prefer derivation to collection wherever the department already has the inputs.

Question 28: Is there any data that you think DfT should require provision of (for example, to effectively enact its role as custodian of the bus sector), but not publish in the public domain?

Yes — and in our view this distinction is the most useful tool available to DfT for resolving the tension in Question 27 without having to decide it as a binary. Candidates for "provide but not publish," or for publication only in aggregate or with a lag:

- Service-level operating costs and staff numbers, for the reasons set out at Question 27.
- Patronage at a granularity that reveals competitive position, particularly on corridors with competing operators or where a tender or franchise competition is live.
- Vehicle registration marks in association with live location, for the data protection reasons at Question 27, provide the registration mark to DfT, publish a pseudonymous identifier.
- Operator addresses and contact details that are residential, which is common for sole traders and small operators and arises directly from the registered particulars.
- Data relating to school services that could identify the pupils travelling or the specific schools served by closed services.
- Network specifications still within a live procurement or franchising competition, until award.

We would suggest DfT to apply the following principles:

1. Publish by default; withhold by exception, with each exception stated in regulations and justified - rather than decided case by case in operation.
2. Where data is withheld, publish an aggregate, redacted or lagged version if one can be made safe, rather than publishing nothing.
3. Publish the fact that the data is held, and a documented access route through which authorities, researchers and other legitimate users can obtain it. A managed access route is frequently sufficient where full open publication is not appropriate, and it preserves most of the analytical value.
4. Complete and publish a data protection impact assessment covering the full proposed dataset before regulations are made.

Question 29: If data such as that listed in previous questions was required for additional reporting, at what frequency do you think DfT should require operators to provide updates?

Tier the frequency to the volatility of the data. A single blanket frequency will be wrong for almost everything on the list.

Data	Suggested frequency
Registration, variation and cancellation particulars	Event-driven, at the point of application, as now
Timetables and fares	On change, published in advance of the effective date — as now, but enforced
Real-time vehicle location	Real-time (see Question 27 on rate)
Cancellations and diversions	Real-time for same-day events; on decision for planned changes
Vehicle and fleet attributes	On change, with a periodic reconciliation (we suggest quarterly) to catch drift
Vehicle-to-journey allocation	Real-time, through the location feed
Passenger journeys and patronage	Quarterly at most, aligned to existing statutory returns
Operating costs, staff numbers, funding	Annually, aligned to the annual bus statistics survey and BSOG/BSIP reporting cycles

Two principles underpin these suggestions.

1) never require a periodic return of something that can be derived from data already flowing continuously. Mileage by fuel type is the clearest example.
2) align every periodic requirement to an existing reporting cycle, so that operators and authorities maintain one compliance calendar rather than several.

Question 30: If DfT were to pursue new regulations in this space, it would likely seek to require any additional data to be provided via the existing BODS. Do you foresee any potential benefits or challenges in requiring additional data to be provided and distributed via that service? Do you have a preferred method or format for certain data items to be provided by?

Benefits

- One place, one account, one support route and one body of guidance for operators.
- Existing supplier integrations and existing operator familiarity, which materially reduces the cost of change.
- Existing validation, quality reporting and open publication machinery, rather than building it again.
- The new data joins to the timetable and location data that is already there, which is where most of its value lies.

Challenges

- Not everything belongs in BODS. It is designed for operational, passenger-facing data published openly. Annual financial and staffing returns are a different kind of data with different sensitivity, a different frequency and different consumers. Forcing them through an open publication platform risks both over-publication and a poor fit, and would be a difficult decision to reverse.
- Scaling. Vehicle-level attributes across the national fleet, and continuous cancellation data, represent a step change in volume and update rate.
- Broadening the mandate broadens the risk of failure: more of the sector becomes dependent on a single service, and the consequences of an outage grow.
- Change management capacity. The sector can absorb a limited rate of specification change, and this programme will already consume much of it.

Preferred methods and formats

- Registration particulars: Initially TransXChange, on a profile aligned with the BODS PTI profile: submitted once and shared with the registration store. Migrate to NeTEx as soon as is feasible.
- Timetables: TransXChange 2.4 on the PTI profile now, with a stated direction of travel (see Question 32).
- Fares: the NeTEx UK profile. The priority here is quality and uptake, not additional requirements. Update to reflect experiences in developing European fares profiles.
- Real-time location: SIRI-VM, with GTFS-RT as a core derived output rather than a secondary conversion.
- Cancellations and diversions: SIRI-SX and estimated-timetable semantics, converged with the Create Transport Disruptions Service, with GTFS-RT service alerts and trip updates derived from them.

- Vehicle and fleet attributes: a separate registry with a simple documented API and bulk CSV/JSON, using NeTEx vocabularies for facilities and accessibility. Not embedded in timetable files.
- Patronage, costs and staffing: a structured return against a published schema, submitted through an appropriate departmental collection route rather than through the open data publication path.

Question 31: If DfT was seeking to enhance BODS in this way, it would have the opportunity to update the existing data reporting regime, including the existing technical specification.
For example, it could increase the frequency with which live automatic vehicle location data should be provided, to provide a smoother display of service movements on a map, or improve bus priority infrastructure, or require additional information to make linking of timetable, fare and location data easier to achieve.
Are there any opportunities you think could come about as a result of these changes, or any challenges that could be addressed through them?

This is the right moment to fix things that cannot be fixed incrementally, and we would encourage DfT to be ambitious about the following - rather than treating them as secondary to the new data items.

- Identifier integrity and matching. Make consistency between SIRI-VM and TransXChange identifiers a validated and enforced requirement. Require VehicleJourneyRef, BlockRef and a persistent vehicle identifier. Publish match-rate statistics per feed and act on persistent failure. In our judgement, this single change would do more for passenger information quality than everything else in this consultation combined.
- Persistent, stable identifiers across the whole estate. A registered service, a line, a journey and a stop should each carry an identifier that survives republication and is the same in every dataset. Identifier churn between publications is a recurring and entirely avoidable cause of downstream breakage.
- Timeliness obligations. Require data to be published on a date, which is a a defined minimum period before the effective date, and measure and publish compliance. Late publication is a passenger-facing failure but is currently invisible.
- Cross-feed consistency validation. Validate the location feed against the published timetable and the registration record, not each in isolation. Most of the failures that matter are failures of agreement between feeds, and no single-feed validator can detect them.
- Bus priority. The question rightly raises this, and it deserves a specific workstream. Bus priority at signals needs low latency and high integrity, and benefits from schedule adherence and occupancy data — a materially different requirement from passenger information, and perhaps better served by a defined interface than by every highway authority consuming the national feed. Several RTIG members have direct experience here, particularly internationally - where significant developments are taking place and we would welcome the opportunity to contribute.
- Linking timetable, fares and location. The practical blocker to fares data being usable alongside timetables is inconsistent identifiers between the

NeTEx fares data and the TransXChange timetable. Validate that join explicitly.

- Historical archive. Publish and retain a versioned archive of all BODS data. A great deal of the analysis members need to do - including much that authorities require for BSIP evaluation and for identifying socially necessary services, is impossible today simply because the past is not available.

The challenges to manage are the sector's capacity to absorb change; the fact that the cost falls on operators and suppliers while the benefit accrues to passengers and authorities; and the risk of a large specification change landing at the same time as franchising transitions in several city regions. We would suggest DfT sequence deliberately, publish the sequence, and commit to changing the specification with a known frequency, for example no more than twice per year.

Question 32: Through BODS DfT requires both UK-specific and particular international data standards to be used when data is provided. Noting that the international standards have evolved since the origins of BODS, are there any particular standards you would like DfT to consider if existing requirements are to be updated or additional data is to be required?

RTIG's position is to evolve deliberately: align with the current CEN standards family where there is a demonstrable benefit, and avoid change for novelty. Specifically:

- SIRI. BODS currently mandates SIRI-VM 2.0 and 2.0Q. The SIRI family has moved on since, and later versions bring capability directly relevant to what this consultation proposes: particularly around estimated timetables, situation exchange and subscription-based delivery. Given the interest in higher-frequency location data and in real-time cancellations, a review is timely. We would prioritise the delivery mechanism subscription and streaming, rather than polling over the version number itself.
- NeTEx. DfT already uses NeTEx for fares. Two extensions are worth considering now: the NeTEx facility and accessibility vocabularies for the vehicle-level attributes proposed at Question 27, in preference to new UK enumerations; and NeTEx Part 5 (alternative modes), which was developed for flexible and shared mobility and is a considerably better fit than fixed-route constructs for demand-responsive transport.
- TransXChange. TXC 2.4 with the PTI profile works and is deeply embedded, and a change needs planning. The DfT should clearly state a direction of travel. NeTEx has become the common format for timetable exchange across much of Europe, and continued use of a national format carries a long-term cost in supplier tooling, in the availability of international products, and in interoperability. We suggest DfT publish a roadmap position even if the position is “TransXChange for the next X years” so that suppliers and authorities can plan. Any migration should be evolutionary, with a published mapping and a long parallel-running period.
- GTFS and GTFS-RT. These are the de facto international standards for consumption and are what most application developers actually use. DfT currently produces them as secondary conversions. We recommend making them first-class outputs which are validated and documented: with published conversion rules and known limitations, and covered by the same service commitments as the source formats. This costs little and materially widens the value of the data.
- NaPTAN and NPTG remain the right national reference data. Their quality and currency is a dependency for everything proposed here, and continued investment in them should be treated as part of this programme - rather than as a separate matter. Migration to NeTEx structures should be planned over a period of time; there is much in common and it may be possible for the DfT to initially convert in the centre to assist with transition.

- OJP is worth keeping in view for distributed journey planning, particularly in the context of the" Better Connected" strategy and multi-modal integration.
- Street Manager and DATEX II for roadworks, already ingested into disruptions data — retain and extend rather than duplicate.
- Established accessibility vocabularies rather than new ones, so that UK data remains comparable and usable by international tooling and by assistive technology.

Two general recommendations.

1) Wherever a suitable standard exists, adopt profiles of international standards rather than UK-only formats: the supplier market is international, and UK-specific formats raise costs for every participant.
2) publish the schema and the validator as the normative artefact, with prose guidance as commentary. Ambiguity between prose and schema is a recurring source of inconsistent implementation, and it is entirely avoidable.

Question 33: For operators of bus services who currently report data via BODS – how much lead-in time do you anticipate you will require to accommodate any changes in data provision requirements?

Properly a question for individual operators. From the sector view, the binding constraint is usually the supply chain rather than the operator: most operators depend on scheduling, ticketing and telematics suppliers to implement specification changes, and those suppliers serve many customers from a single development roadmap. The operator's own lead time is largely a function of where they sit in that queue.

RTIG's recommendation:

- At least 12 months from publication of the initial final, frozen technical specification to the mandatory compliance date - with the specification not changing during that period.
- A sandbox and validator available from the start of that period, not shortly before the end of it.
- A parallel-running window in which both the existing and the new requirements are accepted.
- For any change requiring new on-vehicle hardware - and a move to sub-five-second location reporting could imply that for some fleets, 24 to 36 months is realistic because the constraint is vehicle fitment and refurbishment cycles rather than software release cycles.
- A commitment to no more than two breaking specification changes per year, announced on a predictable cycle.

Question 34: For operators of bus services who currently report data via BODS – what costs do you anticipate you would incur should additional data be required?

This is properly a question for individual operators, and we encouraged members to answer it with their own figures. The cost categories our members identify are:

- Supplier charges: scheduling, ticketing, telematics and real-time information system upgrades, usually passed through as licence increases or project cost.
- New or upgraded on-vehicle hardware where existing equipment cannot meet a new requirement. This is the largest and least avoidable category, and it is the principal reason why we counsel caution on sub-five-second location reporting.
- Mobile data volumes, which scale directly with location reporting frequency, across every vehicle, continuously.
- Internal staff time for data preparation, validation, correction and ongoing maintenance. This is recurring rather than one-off and is consistently under-estimated in impact assessments.
- New data collection where the operator does not currently hold the data at all, for example, vehicle-level facilities, measured in-service fuel efficiency, service-level cost allocation. This may require new processes and new systems, not just new reporting.
- Finance and management time to develop and apply a cost allocation methodology, if service-level cost reporting is required.
- Assurance and compliance overhead, including responding to data quality challenges.

These costs fall disproportionately on small operators, who face broadly the same fixed compliance cost spread across far fewer vehicles and far fewer registrations. We would strongly encourage DfT to publish an impact assessment before making regulations. The consultation as published contains no cost-benefit analysis, and members have found that a significant obstacle to forming a considered view on the proposals.

Question 35: For operators of bus services who currently report data via BODS – what tools or support would assist your business to meet any new requirements?

- A guided data-creation route for operators without scheduling software. In our view, this is the highest-value single intervention available in this programme.
- A sandbox and public validator, available well before the compliance date.
- Reference implementations and worked examples for each data type, covering the awkward cases including short workings, journeys crossing midnight, part-cancellations, demand-responsive services, cross-boundary services.
- Supplier conformance certification - so that an operator can buy with confidence that a product meets the requirement. RTIG would be willing to work with DfT on the design of such a scheme.
- A published change-management process with a predictable annual cycle and committed notice periods.
- Plain-English, role-based guidance (see Question 37) and validation messages that tell the submitter what to do, not only what is wrong.
- A properly resourced helpdesk with transport domain knowledge rather than generic service-desk triage.
- Transition funding for small operators, and consideration of a proportionate regime for the smallest; the compliance cost of a requirement should bear some relationship to the number of passengers who benefit from it.
- Peer support and shared learning. RTIG convenes some of this for BODS today and could do so for whatever replaces it.

Question 36: For existing users of BODS data – what tools or support would assist your ability to make effective use of any new data?

- Stable, documented, versioned APIs with a published deprecation policy and advance notice of change.
- Bulk downloads and daily change files alongside the APIs, so that consumers can maintain local copies without polling.
- A complete historical archive with point-in-time access.
- Published data dictionaries, schemas and conversion rules - particularly for derived outputs such as GTFS, where the conversion decisions materially affect what consumers see.
- Data quality and completeness metrics published as data, so that consumers can assess fitness for purpose programmatically and exclude known-bad feeds rather than silently passing their errors on to passengers.
- The registration-to-publication reconciliation, exposed as data rather than only as a dashboard.
- Reference client libraries and worked examples in common languages.
- A public issue tracker, and a route to report a data error that actually reaches the publisher and is visible to the reporter.
- A sandbox with realistic sample data, including the awkward cases.
- Unambiguous licensing, with no doubt about downstream commercial reuse.

Guidance

Question 37: Would you find guidance on using the new database useful? If so, what would you like it to cover?

Yes, and members regard this as important rather than incidental. Guidance is where the long tail of small operators and the smaller authorities will succeed or fail. What it should cover:

- Role-based routes. Separate guidance for the small commercial operator; the group operator; the LTA with a delegated registration function; the franchising authority; the agent acting on an operator's behalf, and the data consumer. A single document serving all of them will serve none of them well.
- End-to-end worked examples of the whole journey: register, vary, cancel. Including a first registration, a timetable-only variation, a short-notice application and a cancellation.
- Explicit treatment of the awkward cases, which is where guidance most often fails: services crossing authority boundaries; school and works services; demand-responsive and flexible services; seasonal services;, part-cancellations; journeys crossing midnight; and services operating across the England–Wales and England–Scotland borders.
- The statutory framework in plain English: notice periods; who must be notified and when; what the traffic commissioner decides; what happens if an application is refused; and how EP delegated arrangements differ from the standard process.
- Data specifications - with the schema and validator as the normative artefact and the prose as commentary, plus a clear statement of which fields are mandatory, which are conditional, and on what.
- A migration guide for those moving from paper and for those moving from EBSR, written for each starting point.
- Quality guidance: what good looks like, the common errors, how to fix them, and how to interpret each validation message.
- Guidance for consumers on what the data means, how it was produced and what its known limitations are.
- Maintained as a live product with a visible changelog and version history, published on GOV.UK in HTML rather than as a set of PDFs.

Two closing observations.

1) guidance is not a substitute for good design: every item that requires extensive explanation is a candidate for redesign, and we would encourage DfT to treat a growing guidance burden as a design issue.
2) guidance should be co-produced with the people who will use it, including a small-operator review before publication. RTIG would be glad to help.

Response Summary

RTIG welcomes this consultation and supports its direction. The proposals could resolve problems the sector has worked around for many years: in particular, the absence of a single authoritative record of local bus services, and the inability to reconcile that record with the data passengers actually receive.

Our central message is that the value lies less in the new data items than in the connections between what already exists. A single submission, a single identifier spine, and a published reconciliation between the registered position and the published data would deliver more for passengers than any of the individual new requirements, and would make those new requirements worth having when they come.

We would welcome the opportunity to work with the department on the technical specification, on the design of the transition, and on the supplier conformance and guidance arrangements. We would also be pleased to convene members; authorities, operators and suppliers together to test design proposals as they develop.