National Rail Enquiries

# Developer Engagement Day

## Requests Report

16 June 2015

**Purpose:**

The purpose of this report is to provide an open account of discussions held at the Developer Engagement Day and the requests made to NRE by the developer community. NRE support the concept of transparency and practice its principles through the provision of open data, frequent engagement with developers, and through the publication of material from developer engagement events.

**Introduction:**

National Rail Enquiries held a developer engagement day on 16 June 2015 at the Museum of London. The event was attended by circa 100 developers comprised of a mixture of independent developers, consultants, commercial companies and transport executives as well as representatives from Network Rail and DfT.

**Agenda:**

The agenda opened with an introduction to the Industry transparency agenda, covered by RDG Comms Director, Edward Welsh. This was followed by a session on the Customer Information Strategy, covered by Head of Customer Information, Jason Durk. Subsequent items included NRE's open data roadmap, Darwin licensing, and retail information, covered by Client Relationship Manager Lindsay Bleakley. Following, was a session on industry fares and routing guide data, covered by data architects Darryl Tai and Rasheed Amzart. The event finished on a technical Q&A with NRE's Online Systems Manager Andrew Tolley and an open floor session for developers to ask questions and make requests relating to the provision of data.

**Event Aims:**

The event aims to provide the developer community with an opportunity to engage with NRE regarding the items covered in the agenda. It also gives developers a platform on which to ask technical questions about NRE data and to make requests and/or suggestions on how NRE can improve its service offering. For NRE, the event provides an understanding of developer requirements, supporting us in making informed decisions, both technically and on a policy level, regarding how to make data available for maximum possible benefit to the developer, the industry and ultimately the customer.

**Outcomes:**

NRE collected general feedback along with a number of requests from developers for potential improvements to the provision of NRE services. Developer requests were used to formulate a questionnaire which has been circulated to all NRE developer clients. The questionnaire will assist NRE in prioritising future actions based on the quantitative and qualitative feedback received. In line with the principles of transparency, feedback will be published on NRE developer engagement channels.

National Rail Enquiries

**Developer Engagement Day**

Requests Report

8 July 2015

| Reference | Developer Requests | Score[1] |
|---|---|---|
| 2015-6-16_#1a | Can NRE enable Darwin to consume validated data from 3rd parties? | 35 |
| 2015-6-16_#1b | Would NRE pay for validated data sets that would enhance Darwin? | -9 |
| 2015-6-16_#2a | Can NRE remove the clause preventing users of Darwin from altering the forecasted arrival & departure predictions? | 19 |
| 2015-6-16_#2b | Will NRE allow users of Darwin to display information that could contradict Darwin if the source of that information is clearly identified? | - |
| 2015-6-16_#2c | Can NRE clarify if licensees are permitted to show historical averages timeliness of services as a service reliability forecast? | - |
| 2015-6-16_#3a | Can Darwin provide real-time departure/arrival/location information for Rail Replacement Buses (RRBs)? | 78 |
| 2015-6-16_#3b | Can Darwin intake ad-hoc RRB schedules, as and when they happen, rather than just the ones submitted into the industry timetable (CIF)? | - |
| 2015-6-16_#3c | Can lat/lon data exist for RRB departure points be made available? | - |
| 2015-6-16_#4a | Can NRE increase the message queue "time to live" in the event of disconnection for the queue to at least 10 or 15 minutes? | 39 |
| 2015-6-16_#5a | Can Darwin show GPS data in the Push Port, when available? | 107 |
| 2015-6-16_#5b | Will the output from Darwin show all GPS data processed by Darwin, or just the pings that Darwin selects as the most accurate location report? | - |
| 2015-6-16_#5c | Can NRE expose all the GPS data in a separate feed? | 82 |
| 2015-6-16_#5d | Can lat/lon data for TDs be made available? | 85 |
| 2015-6-16_#6a | Can NRE offer a support service with SLAs for users of Darwin feeds? | 50 |
| 2015-6-16_#7a | Can NRE allow automated requests to the Real Time Journey Planner Webservice rather than user driven requests only? | 43 |
| 2015-6-16_#8a | Can ATOC make the advance fares data open & free of charge? | 107 |
| 2015-6-16_#8b | Can ATOC make the routing guide data open & free of charge? | 87 |
| 2015-6-16_#9a | Can NRE offer an alternative attribution logo that doesn't include the "Powered By" text? | -1 |

[1] The 'Score' is the value attributed to each question to demonstrate importance based on feedback from the 75 developers who answered the questionnaire. Scoring is calculated for the response to each questions as follows: Strongly Agree = +2, Agree = +1, Neutral = 0, Disagree = -1, Strongly Disagree =-2.

| Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
|---|---|---|---|---|
| -150 to -75 | -75 to -20 | -20 to 20 | 20 to 75 | 75 to 150 |

| **Ref: 2015-6-16_#1** | **Adding Data Into Darwin** |
|---|---|
| Question(s) | a) Can NRE enable Darwin to consume validated data from 3rd parties?<br>b) Would NRE pay for quality validated data sets that would enhance Darwin? |
| Context | Discussion was held around the accuracy of Darwin data. Developers stated that, on occasion, they have extra information about train location that isn't currently consumed or processed by Darwin. Under the current licence conditions, a user of Darwin information cannot use this additional data to alter forecasted arrival or departure times produced by Darwin.<br>Developers were concerned that this licence condition was stifling innovation related to improving the accuracy of information to customers. There was also concern that this was prohibiting commercial opportunities for suppliers to create a more accurate version of Darwin. |
| NRE Response | Darwin has been made openly availably to underpin the Customer Information Strategy Vision of creating a single consistent message to the customer. It would be counter-intuitive to this wider industry remit, to licence Darwin data in a way that enabled users to alter the forecasted arrival and departure times and create multiple 'versions' of Darwin.<br>Users of Darwin data can currently inform NRE of errors and/or omissions in the data by contacting the NRE Service Desk to resolve this issue at source. |
| Actions | 1) NRE to respond to the request for enabling a mechanism through which 3rd parties could add data to Darwin.<br>2) NRE to respond to the question about whether remuneration for validated data sets will be considered. |

| **Ref: 2015-6-16_#2** | **Darwin Licensing** |
|---|---|
| Question(s) | a) Will NRE remove the clause preventing users of Darwin from altering the forecasted arrival & departure predictions?<br>b) Will NRE allow users of Darwin to display additional information that could contradict Darwin if the source of that information is clearly identified? (eg. Can users show data GPS from passengers, even if that data indicates a train is closer or further away than Darwin predicts)<br>c) Can NRE clarify if licensees are permitted to show historical averages timeliness of services as a reliability indicator forecast (in a journey planner for example)? If so, up until what time window (eg. departure time, 2hrs before departure etc?) |
| Context | Developers voiced concern that the Darwin licence clause requiring users to ensure "consistency with Darwin forecasts" was stifling innovation related to improving the accuracy of information to customers. The clause prohibits suppliers from using 'added value' data-sets to create more accurate predictions than Darwin, thus limiting related commercial opportunities. The impact here being that without commercial incentive, there will be no investment in innovation.<br>Some users want to use Darwin information to analyse historical arrival and departure times so as to provide customers with an historical average timeliness of train services. The intention would be to provide this information during a journey plan. It was noted that the licence condition on Darwin forecasts doesn't make it clear whether this is a permitted use of data or not. |
| NRE Response | Darwin has been made openly availably to underpin the Customer Information Strategy Vision of creating a single consistent message to the customer. It would be counter-intuitive to this wider industry remit if NRE were to licence Darwin data in a way that enabled users to alter the forecasted arrival and departure times and create multiple 'versions' of Darwin. |
| Actions | 1) NRE to provide further formal response to requests 2a & 2b on the Darwin licence clause about "forecasted arrival & departure times".<br>2) NRE to clarify guidance as per request 2c. |

| **Ref: 2015-6-16_#3** | **Real-Time Bus Data in Darwin** |
|---|---|
| Question(s) | a) Can Darwin provide real-time departure/arrival/location information for Rail Replacement Buses (RRBs)?<br>b) Can Darwin intake ad-hoc RRB schedules, as and when they happen, rather than just the ones submitted into the industry timetable (CIF)?<br>c) Can lat/lon data exist for RRB departure points be made available? |
| Context | Discussion was held around rail replacement bus (RRB) information in Darwin. The perception of some developers was that existing RRB information in Darwin is static and can be incomplete. Developers enquired about why Darwin doesn't currently provide real-time RRB information.<br>There was a lack of clarity in the room about the origin of RRB information, with many people assuming that bus companies provide the RRB services and thus should also provide the information. Multiple comments were made that NRE simply need to integrate with real-time bus information APIs like Traveline.<br>Another point discussed was the lack of data about RRB departure locations at stations. Developers were interested to know if data existed for this and if so, could it be made available. |
| NRE Response | RRBs are not run by bus companies, they are not timetabled by bus companies and they are not available through real-time bus APIs like Traveline. RRB services replicate train services, stopping at train stations rather than public bus stops. The schedules for RRBs are generated by TOCs, who submit the information to Network Rail who, in turn, include the RRB schedules in the industry timetable (CIF file) that feeds downstream systems.<br>RRB information in Darwin isn't currently real-time. From a technical perspective, anything is possible, however Darwin does not currently have the ability to take in location data from RRBs, nor the logic to process that data to provide arrival or departure predictions. |
| Actions | 1) NRE to feedback on plans / feasibility of displaying real-time RRB data in Darwin.<br>2) NRE to feedback on existing process for adding RRB info into Darwin and plans to improve this.<br>3) NRE to investigate the possibility of getting lat/lon data for RRB bus stops. |

| **Ref: 2015-6-16_#4** | **Message Queue Time Limits** |
|---|---|
| Question(s) | a) Can NRE increase the message queue "time to live" in the event of disconnection for the queue to at least 10 or 15 minutes? |
| Context | With the existing Push Port architecture, when a user disconnects from their message queue, messages continue to be received by that queue for a maximum of 5 minutes. The queue will only hold the last 5 minutes worth of messages before messages before they drop of the end and are lost.<br><br>In disaster recovery mode it leaves Push Port users with very little time to identify and resolve system issues, and for a new machine to be booted to connect to the queue so messages can be received again. This is a problem, as if there is a cancelation (for example) during the time when a user has been disconnected from the queue for more than 5 minutes, the user would miss that piece of critical data when reconnecting to the queue.<br><br>NRE explained that all users have access to the FTP site, where all the messages are saved in 5 minute blocks. Users can access all lost messages through the FTP site. Developers noted this was a cumbersome and difficult approach. Extending the message queue time limits would make a big difference. Most developers noted that an increase of 5 -10 minutes would make a big difference. One developer noted that an increase to 1 hour would be preferable. |
| NRE Response | NRE will consider increasing the message queue time limit to 15 minutes. We won't make it 1 hour at this stage, as this would require us to ramp up storage capacity significantly, meaning increased costs for NRE. The user would also have to consume an hours' worth of data in the event of an hour long disconnection. Doing so would result in a long delay when reconnecting to the queue while the users system processes the data.<br><br>Should we make the change to 15 minutes, we will assess the impact to NRE and the user. If there is a need to extend this further, we will consider doing so. |
| Action(s) | 1) NRE to investigate feasibility of increasing the message queue "time to live" to 15 minutes? |

| **Ref: 2015-6-16_#5** | **GPS Data** |
|---|---|
| Question(s) | a) Can Darwin show GPS data in the Push Port, when available?<br>b) Will the output from Darwin show all GPS data (pings) processed by Darwin, or just the pings that Darwin selects as the most accurate location report?<br>c) Can NRE expose all the GPS data in a separate feed?<br>d) Can the lat/lon of TDs be made available? |
| Context | Discussion was held about the Customer Information Strategy plan to develop a GPS gateway into Darwin, allowing WiFi/GPS enabled train sets to report their location directly into Darwin. The the room was positive about the plans. Discussion was then held around the format in which GPS data would be presented to users and whether users would be able to identify the GPS data within the Push Port.<br>Developers indicated that they would like the GPS data being made available as a feed separate to Darwin. Developers also discussed the possibility of crowd sourcing GPS data and using that data to improve Darwin predictions. Question was also raised about the lat/lon data for TD feeds. Developers expressed a desire to have access to the data if it was available. |
| NRE Response | NRE noted that the architecture for the GPS gateway into Darwin is still conceptual. The data format has not yet been determined, nor have we ascertained the frequency at which the GPS signal will be sent or consumed by Darwin. We will consider the developer feedback when designing the output.<br>In reference to crowd sourcing GPS data to improve forecasts made by Darwin, NRE also noted that the core aim of the strategy is to improve data at source so that all users of the information. Should developers wish to crowd source GPS data and match that to services, the NRE preference would be that the GPS data was consumed by Darwin so all users and all customers benefit. |
| Action(s) | 1) NRE to consider developer feedback when designing the GPS data output.<br>2) NRE to confirm if GPS data can be provided in a separate feed.<br>3) NRE to update developers on GPS at the next developer day.<br>4) NRE to find out if lat/lon data for TDs can they be made available. |

| **Ref: 2015-6-16_#6** | **Service Support for Darwin Feeds** |
|---|---|
| Question(s) | a) Can NRE offer a support service with SLAs for users of Darwin feeds? |
| Context | Developers were challenged to innovate using NRE data. Developers then raised that point that order to commit to using the data and investing in innovation, they would need a 24hr supported service with an uptime guarantee and SLAs.<br><br>Developers indicated that they would be prepared to pay for support, including that investors would not invest if there was no guarantee of continual supply and the availability of support for the service. Effort and investment in innovation would be limited without the option to pay for support.<br><br>The overriding message was that if users are generating revenue from Darwin then they would be happy to pay for a robust SLA. Developers generally agree that NRE will incur costs in developing the service, so should be paid. However, a free option with no SLAs should still be made available for those unable to pay. |
| NRE Response | NRE took the feedback on board and agreed to take it back for consideration. |
| Action(s) | a) NRE to consider offering paid for support with SLAs for users of Darwin feeds? |

| **Ref: 2015-6-16_#7** | **RTJP Webservice Usage** |
|---|---|
| Question(s) | a) Can NRE allow automated requests to be made to the Real Time Journey Planner Webservice rather than user driven requests only? |
| Context | Developers were challenged to innovate using NRE data but developers raised the point that usage restrictions to the RTJP webservice prevent beign innovative with this feed. Developers stated that the source of the query should be of no interest to NRE as long as the volume of queries does not cause service instability, including that arbitrary restrictions such as this devalue the idea of 'open data'.<br><br>Developers suggested that without this clause they would be able to push relevant information to users before they actively request it. This is a current trend with products like "Google Now" which offer things like delivery and traffic updates without a user explicitly requesting them.<br><br>Other developers suggested potential uses such as the ability to monitor a user's planned journeys, provide disruption alerts along with suggested alternative routes. This can't be done if the user is required to query the journey planner manually.<br><br>Developers questioned why the rule exists when they are paying for the queries. However, others commented that unless there are great IT systems behind this, it could easily lead to overload from automated requests. There would seem to be a higher risk profile with this. |
| NRE Response | NRE explained that the journey planning engine is owned by SilverRail and the webservice which queries the journey planning engine is owned by NRE. The charges that are passed on to developers are a pass through of the charges that are incurred by NRE.<br><br>NRE included that the RTJP Webservice is hosted by an NRE system called the Online Journey Planner (OJP), which powers all NRE journey planning products. The restriction around automated usage is in place to prevent excessive load to NRE servers, which could in turn crash NRE's own channels.<br><br>The business case for NRE to migrate the RTJP Webservice onto dedicated resilient servers is not there because we don't own the journey planning engine. Rather we are customers of SilverRail.<br><br>NRE also agreed to take the feedback on board and review the request for further comment. |
| Action(s) | 1) NRE are to respond to this request taking account of feedback |

| **Ref: 2015-6-16_#8** | **RSP Data – Fares, Timetable, Routing Guide** |
|---|---|
| Question(s) | a) Can ATOC make the advance fares data open & free of charge?<br>b) Can ATOC make the routing guide data open & free of charge? |
| Context | RSP presented information about the industry fares and routing guide data. The primary issue in the room was that the information was both complex and expensive to access. The key argument was that the data is so complex that it will take a long time to interpret. If a user then has to pay many thousands of pounds to use the data, it makes the business case for using is very disadvantageous.<br>By a show of hands, a majority percentage of attendees from the developer community agreed that they would like the fares and routing guide data to be made open and free of charge. Attendees were invited to respond to a questionnaire giving their thoughts on the matter. A cross section of those responses are shown below:<br>- *"I feel very strongly that this should be the case. I can't see why the daily feeds have such a large license fee."*<br>- *"This data forms a central point for potential innovation to clarify a Byzantine rail fare structure for customers."*<br>- *"the customer will benefit since this will making comparing prices easier and thus increase competition between providers."*<br>- *"Third parties have generally improved information flow on timetables and real time data, there is no reason to expect that this would be different for fares."*<br>- *"It is just obvious, choosing the most cost-effective travel option is a problem for many, and I am not sure that the current offering is doing a great job."*<br>- *"This and passenger volumes is the biggest missing piece for commercial modelling and prediction. Really understanding how the railway is used and how to advice customers on getting the best experience is critical to success."*<br>- *"It is critically important that all fares data is available for anyone to inspect and use - to trust that the industry is fair and customers can find the best value ticket for their journey, if so inclined."*<br>- *"This is absolutely essential and should always have been available in the public domain at no cost. This could lead to services that would hugely benefit the consumer and TOCs alike."*<br>- *"Aventix is already provided in 'closed' formats - allowing the data to be used in an open format would allow for additional journey planning/ticketing platforms to be integrated. Restrictions and ticket validity information would be a big benefit as that information is often hard to get in full at a point of sale."* |
| NRE Response | It was agreed to raise the matter within ATOC and to provide a formal response. |
| Action(s) | 1) NRE to respond to these requests, taking account of feedback |

| **Ref: 2015-6-16_#9** | **NRE Attribution & Branding** |
|---|---|
| Question(s) | a) Can NRE offer an alternative attribution logo that doesn't include the "Powered By" text? |
| Context | Discussion was raised about the attribution logo that developer are required to use on their products. Some developers felt that the logo was misleading customer to believe that their product was a national rail product and/or contained only national rail data.<br><br>Other developers suggested that the logo was a certification of the quality of the data and they had no problem with the logo. Feedback obtained through the questionnaire is as follows:<br><br>- *"A modern clean look isn't achievable with the National Rail Enquiries requirements for branding"*<br>- *""Powered" is too much of a strong word, it sound like we were reusing code developed by you, not just your data. It would make my application look like a white label application of yours on top of which I just put some branding."*<br>- *"The requirement to display this logo should be relaxed. A single line of text (as per the licence for the Network Rail feeds) as an attribution is sufficient in my opinion."*<br>- *"I have no issue with the logo, but allowing the use of a smaller logo, or a text link may be more applicable to some people/projects."*<br>- *"It's a pain having to show this great big logo everywhere, and if every public transport data provider insisted on this there'd be no room left on the screen for everything else in multi-modal applications. I'd prefer to be able to just attribute in text form, ideally on the "About" page rather than on every page the data appears."*<br>- *"My site makes use of several feeds from elsewhere all of which only require text attribution, or nothing at all. It should be permissible to just credit the data source to National Rail Enquires without use of logo."* |
| NRE Response | The existing conditions state that the user is not obligated to use the logo in instances where the logo negatively impacts the user experience. They also state that where multiple data feeds are being used, the logo can be displayed on an attributions page.<br><br>It was also agreed that NRE would review the request, in light of the feedback, and provide a formal response. |
| Action(s) | 1) NRE to respond to this requests, taking account of feedback |