Response to the Office of Rail Regulation's Consultation on Real-Time Travel Information

Alex Hewson

Independent Application Developer

mock@mocko.org.uk

25th February 2013

Confidentiality

Introduction

As an independent application developer I have been very much affected by the rules surrounding access to Darwin and their administration.

In this response I set out what I perceive to be the greatest problems with the Code of Practice and associated license, the artificial constraints they place on the transport data industry and an argument that Train Information Services Limited (TISL) are not an appropriate body to steward the only reliable source of real-time travel information for Britain's railways.

I finish with a vision of what applications might appear if access to the data in Darwin were truly open and how it would result in a railway more flexible for Train Operating Companies and more popular with its passengers.

Background

To avoid cluttering my answers to the consultation points I'll start with the background to my relationship with TISL and their implementation of licensing for Darwin's data.

Until November 2010 I was a user of TISL's Darwin API. I built a small, free-to-use web app which helped users plan a journey when they were blessed with a choice of stations. The pages were deliberately spartan so as to load quickly on mobile devices and the search was bookmarkable, meaning users need click only a single button on their phone to be presented with the current list of journey options.

When TISL announced that licensing for their data feed was to become mandatory I contacted them to ask for one. A handful of emails were exchanged with Derek Parlour but his responses were brusque and rather negative. He told me there was very little chance I would be given a license for a non-commercial application, so in desperation I published a blog post[1] about how TISL's actions were affecting the developer community. As media attention into the issue grew I started a web page to track other developers requests for licenses in the hope of bringing a little transparency to what was then a very opaque process. After TISL's Chief Executive Chris Scoggins chastised me for publicising a few lines of Derek Parlour's email I wrote further posts and appeared on Radio 4's "You and Yours" programme to discuss what was going on.

Around the time this was going on Mr Scoggins was making claims in the press[2] that his organisation was committed to open data. However, the less positive media attention when I blogged about my experiences seems to have caused great offence at TISL. I appealed against their refusal of a license and was told by Mr Scoggins that since I'd sullied TISL's name I did not fit their image of a "good" partner and would not be permitted to use the Darwin feed. I had been punished for speaking out against them.

---

1 http://bit.ly/15dkw1O

2 http://bit.ly/YREmdG

Faced with an immovable barrier at TISL I gave up on my app.

In December 2012 the ORR staged a transparency conference and I was glad to attend.

Following a bruising Q&A session with the conference's attendees Chris Scoggins announced TISL would hold a "Developer Engagement Day". Lasting 3 hours this would be an attempt to foster better communication with the developer community. It sounded good on the invitation but little has changed as a result. He and Derek presented on the challenges they face running TISL and shared some nontechnical information about Darwin. They repeatedly made the point that they'd like to avoid developers airing problems publicly – something which makes Chris's subsequent ignoring my emails until I became exasperated and blogged about it all the more disappointing.

At the end of the Developer Engagement Day I asked if I might now be granted a license. They made positive noises and sent me a copy of the license but when I tried to negotiate on its transparency limitations, stopped taking to me for three weeks. In fact TISL only started responding to my emails again when – after three weeks of silence from TISL - I published the latest happenings in another blog post[3]. And much of *that* response was just an admonishment from Mr Scoggins for publicising his reticence.

Despite my very genuine attempts to negotiate on the transparency issue TISL won't budge. I cannot have a license unless I agree to the "don't talk about the business relationship without permission" clause, something which as a blogger as well as a developer I am not eager to do.

---

3   http://bit.ly/XQba5M

Consultation Question 1

1.16 Consultation question 1: We are looking for stakeholder comments on TISL's proposed changes to its Code and where changes have not been made, comments on TISL's reasoning.

On 11th February 2013 National Rail Enquiries released its new Code of Practice. A copy of the license was appended which means at long last we are able to publicly discuss it. Perhaps this wasn't as transparent a gesture as it appears: the license was bound to come out during this consultation anyway.

Since the license comes appended to the CoP I have included it in the scope of my response. I have broadened my analysis of some items (particularly push data) in Friday's blog post[4].

VAGUE LANGUAGE & NO PROVISION FOR PUBLIC INTEREST

Throughout the Code of Practice vague terms are used which give TISL a huge amount of leeway in its implementation. This vagueness means they can use a great range of reasons to allow or deny licenses while still adhering to the rules.

It has obviously been written with the interests of the TOCs as paramount rather than those of the public. Given TISL's funding comes from TOCs this should not be surprising: it is primarily to serve them that ATOC exists and not the public.

Some examples are:

- *"The application or proposed service does not risk a material adverse impact on*

---

4   http://bit.ly/XQba5M

TOCs, whether financially, strategically, operationally or in regards to their reputation or the reputation of the industry as a whole." - what if an app would highlight an underperforming TOC? This could be as simple as a punctuality league table using data from Darwin: however good the developer's intentions somebody would have to be at the bottom and publicising this would affect their reputation. Even though doing so would be in the public interest the code's terms forbid it and legitimate analysis of the TOC's performance would not be allowed. Also no rules are laid down about how TISL will judge the impact of an application.

- *"The applicant is a reputable company or person."* - who decides what "reputable" means and by what criteria?

- *"Applicants who use TISL services or data without permission or breach their licence terms may either be disqualified for consideration for a licence or have the same withdrawn."* - not only can my license be declined for being against the interests of TISL or the TOCs, but I can be disqualified from all future license applications.

APPEALS PROCESSSs

Ostensibly the biggest change to the code is the offer of external arbitration for the appeals process. This is purported to solve the original conflict of interest inherent in TISL's Chief Executive being the last line of appeal.

The appeals process is a red herring. All TISL are really committing to do is follow the code they've given themselves. If a license refusal goes to appeal the arbitrator can only judge on whether the code has been followed – they cannot declare it unreasonable or order changes. The terms of the code & license are so broad that TISL can easily find

ways to decline an application whilst still following it to the letter.

SMARTPHONE APPS AND PUSH DATA

"Push data" is a term for a live feed of information about train services pushed directly to the developer's applications.   For a smartphone application to alert a user with interesting, timely updates it is vital to the developer yet the terms under which push data may be made available are skewed heavily against them.

The most serious issue is the requirement for security and liability insurance which is levied against developers. Earlier in the CoP TISL says "much of the data that feeds into the service is obtained without warranty from sources external to TISL". Yet while *they* give no warranty about the quality of data TISL expect application developers to give a strong (and expensive) warranty about their usage of the data. This is grossly unfair.

The CoP says:

> *Licensees should be aware that they may be required to gives indemnities and security for damages that any errors they make may cause. Because of this licensees applying for push services should ensure they have the financial capacity to give these indemnities and to put up the security.*

...which effectively prevents any independent developer from taking part in the market. Also the terms are so vague ("errors" is undefined, and there is no upper bound to financial penalties[5] and size of security required) as to give TISL complete discretion over

---

5   In a recent conversation TISL told a developer they would need to provide <u>unlimited</u> financial indemnity to the TOCs and would have to show they have the capacity to compensate the TOCs up to at least seven figures for loss of revenue or reputation caused by errors.

what requirements they place on each applicant.

I should be very interested to see evidence of the concerns which led to this clause. Are we to believe some TOC has told TISL the writs will start flying if a website publishes incorrect data and costs them a few passengers?

APPLICATION PROCESS

The wording of the application process betrays TISL's real attitude to openness. It is up to the developer to prove their application is compliant with the CoP and earn access to Darwin. TISL may take up to 21 days to decide whether the application is worthy which in a world of mash-ups and rapid application prototyping is an unforgivable delay.

Applicants *"will need to submit a separate application for each permitted use"*. This is a stone-aged way of looking at innovation. Developers will constantly be dreaming up new ways to leverage the data they have access to, but they must reapply and wait 21 days whenever they want to try out their next idea.

In case there was any doubt about TISL's intention to prevent any uncontrolled innovation it's laid out in the license terms:

> 3.2. The Customer shall not:
> ...
> 3.2.3. save as provided in the Schedule make any additions, alterations or deletions to the Service and the data held therein without the written authority of TISL.

CHARGING STRUCTURE

The CoP states "the use of real time services in mobile apps ranges from £1.00 to £1.50

per app downloaded". TISL gives its Android app away so with the strength of their brand it would be very hard for any external developer – commercial or an individual – to compete. At the very least TISL ought to charge the same £1.00 to £1.50 for its own application to level the playing field.

"DISREPUTE" CLAUSE SILENCES CRITICISM

The most insidious parts of the licensing agreement are the gagging clauses. You may lose your license for publicly speaking about your relationship with TISL:

> 15.3. TISL shall be entitled to terminate the Agreement on notice without liability at any time if:
>
> 15.3.1. the Customer brings TISL, the TOCs or the passenger rail industry into disrepute;
>
> …
>
> 18. No public announcement, communication or circular (other than to the extent required by law) concerning the subject matter of this Agreement shall be made or despatched by the Customer without the prior written consent of TISL.

Disrepute is a broad term and TISL can make its own judgements about what that means. If a commercial developer is dependent upon a Darwin feed for her livelihood he will not dare to speak about TISL or the TOCs in less than glowing terms for fear of losing access. How far it goes is undefined. In an extreme example, if one listed all late-running Chiltern Trains services in red blinking text to bring attention to them could that be seen as damaging a TOC's reputation?

In case TISL assert that they would be unlikely to invoke this clause let us remember Mr Scoggins' stated reason for denying my license appeal back in 2011 – that in his opinion my blog posts had been intended to give TISL a "poor reputation". It is a real threat.

SUMMARY

In giving ground on a couple of minor points and organising a Developer Engagement Day TISL probably believed they would quieten their critics. But the license terms are still

heavily stacked against developers and designed to exert control over anyone who enters the real-time train information industry. It hinders competition with their smartphone app and silence critics with a threat of losing access to the feed.

In its present form it is not compatible with the goals of openness or transparency.

Consultation Question 2

1.17 Consultation question 2: We are looking for stakeholder comments on the extent to which Network Rail's data feed represents a viable alternative to Darwin and the uses that these feeds can be put to.

The feeds available from Network Rail[6] are of interest to developers and can be used for a broad variety of applications. It is positive that NR have chosen to make them available and continue to assess other feeds and datasets for release.

When talking about the usability of feeds we should consider the license governing their use. Network Rail have adopted a very permissive license for the usage of their feeds: they clearly want the data to be consumed and accept that many innovative and unforeseen applications will appear. They have created no barriers to entry for anyone who wishes to participate in the market, be they a private individual or an established company.

Sadly for the most popular purpose – live journey planning - the feeds are not yet a viable alternative to Darwin. This is not through any fault of Network Rail: they simply don't have all the data. While Network Rail have a good snapshot of what's happening on the network at any given time their feeds don't incorporate short-term cancellations, delays or re-routing by TOC's and are thus not reliable enough for giving live updates to travellers. A live journey planning app based solely on these feeds would end up with some angry users.

Given sufficient budget, buy-in from TOCs and the will to do so Network Rail might implement a new feed that mixes the data they already have with additional data from the Train Operating Companies. This "raw" data could be enhanced by making predictions

---

6   listed at http://www.networkrail.co.uk/data-feeds/

about train timing in the same way Darwin does. For example a train passes a waypoint 5 minutes late but has 2 minutes slack in its timetable before reaching the destination: Network Rail can therefore predict it will be 3 minutes late. This would make a truly open alternative to Darwin and in many respects (public accountability, approachability, track record of openness) their stewardship would be infinitely preferable. TISL might even save the £2m it spends every year developing Darwin by adopting this new, shared source.

Finally an unasked-for opinion on Network Rail's data strategy...

The open data policies Network Rail have implemented are balanced and act in the public interest. The few rules about usage are sensible and unambiguous. They explicitly state that "You are free to" … "exploit the Information commercially" - a commendable attitude given they might have tried to recover costs through a two-tier system where commercial use was charged for. This was an intelligent decision: administering a system of paid licenses would probably have cost more than it was worth.

Network Rail's[7] commendable approach to open data should be the yardstick by which we measure the rest of the industry and sets an exemplary benchmark for others to to follow.

---

7 It would be remiss not to mention also Tfl, who have been sharing data even longer and whose licenses are similarly encouraging to innovation. Network Rail's positive approach is not an isolated case.

Consultation Question 3

1.18 Consultation question 3: We are interested to hear consultees' views on the evidence that we present in Chapter 5 on the number of new licences and apps., and on any reasons why they consider this growth might overstate the health of this market. In particular we welcome stakeholder views on: (a) The medium-term sustainability (to the extent that this is possible to predict in a fast-moving technology market) of the relatively large number of apps that are currently on the market, including on the feasibility of paid and ad-funded or free-to-download apps coexisting; and (b) The likelihood of a significantly better range of applications and functionality being made available under a more open data standard.

TISL's figures on growth only cover cases they know about. They have no figure for the number of applications which existed before an API key was mandatory to access Darwin and were quietly snuffed out.

They can also have no idea about the number of developers who have been actively discouraged. Such discouragement can come from:

- The complexity and daunting legalese of the license

- The liabilities set out in the license which the developer must shoulder

- The relative difficulty of the licensing process (i.e. apply in writing to Derek Parlour, spell out everything their application will do and commit to re-applying whenever they have a new idea)

- TISL's historical attitude to its critics

If more developers were encouraged to participate in the market we would see a greater flow of ideas and new talent. This is covered in more depth in my answer to the next question.

Consultation Question 4

1.19 Consultation question 4: We ask consultees for views on whether an open data approach, if adopted, would lead to change in the market for RTTI products and services and if so: (a) what this change might look like; and (b) whether it would be desirable.

If the data contained within Darwin were shared under a license as open as the one used by Network Rail it would be of significant benefit to the transport information industry and ultimately to the traveling public.

The results might include:

- Rapid innovation in distribution of service information. If push data was available for free, for everyone, an avalanche of smartphone applications would appear to help people work around disruptions to service in their area. I've appended an idea of my own at the end of this document.

- A rapid influx of developers into the sector who are presently being discouraged by onerous license terms. More people means more imagination and this will speed up innovation yet further. It should be born in mind that for someone with programming skills the barriers are minimal: we're not just talking about commercial development firms, anyone with a computer can try their ideas out in hours.

- Far more comprehensive, public assessment of the performance of TOCs. It would be easy to take a live stream of "push" data and publish up-to-the-second statistics on punctuality and cancelations. Some assessments of the data would be better than others (it's complicated and we know that oversimplifying can sometimes paint the wrong picture) but once it's out in the open passengers and the media will learn which sources to trust.

- Better, more timely information means greater tolerance of the market to delays, which will be of benefit to TOCs. Historically the timetable was the only source of "truth" when planning a railway journey and when a service deviated from this it would cause a lot of inconvenience. But what if people stop planning with their biannual timetable and instead do it half an hour before they travel with a smartphone app? They will have far more accurate information to act upon and long waits on the platform will be a thing of the past.

- Related to this – is the logical conclusion that TOCs would be able to notify huge swathes of passengers of alternation to services within seconds. A great example of this would be the communication of revised timetables when it snows. At present such changes often lead to chaos (often for the TOC's employees as much as their passengers!) but with push data fed into an ecosystem of ubiquitously deployed applications we'll be faced with an amazing opportunity to get everyone on the same page.

These opportunities are highly desirable to the developer community, the Train Operating Companies and to the public at large.

Appendix 1 – Design for a Smartphone Application Making Innovative Use of Push Data

This took five minutes to come up with and I include it as an example of the kind of rapid innovation possible when real-time service information is available freely to all. Perhaps the most important characteristic is that it could be developed by a single competent programmer wanting to "scratch an itch" and build something that'll make thousands of people's days a little easier.

RATIONALE

Many people have smartphones. Given GPS chips these devices know where their owner is and can periodically record the data for future use. What if an application could "learn" the start, end and normal time of my morning commute and automatically alert me of problems?

Under the present system of Darwin licensing this application could not be made. The developer would need to provide indemnity and security to TISL against any conceivable problem his application might cause. No small developer would be able to implement it and no large one would wish to shoulder the risk. An opportunity is lost and passengers suffer.

IMPLEMENTATION DETAILS

The app would:

• Be developed in under a month by a single programmer

- Distributed in the Android marketplace either for free or for a couple of pounds – low enough to be affordable to everyone.

- Use the GPS on the owner's phone to "learn" the journeys they regularly make and when they make them. For example "Alex travels from Ladywell to London Bridge at 8am most weekdays".

- Armed with this knowledge, connect to Darwin's push system an hour before the usual commencement of travel and look out for significant disruptions.

- Alert the user to disruption as soon as possible so before even leaving the house they have the opportunity to make alternative plans. In well-served areas this might mean using a different line; for non city-dwellers it might mean taking the car or working from home that day.

- [Optionally] automatically email your boss to say you'll be late.

POTENTIAL RESULTS

- The user worries far less about being late for work and doesn't need to leave the house 20 minutes earlier than necessary in case of disruptions. They get an extra 20 minutes in bed.

- Initially we might say that the greater visibility of disruption will cause the user to think less of their TOC. But I think the opposite will happen – an ability to re-plan their day means the cost of disruption is greatly reduced. The user will never again get a nasty surprise when they get to their local station and this greatly improves their relationship with the operator.