

Traveline

National Operator Codes Database Development

Consultation March 2015

Created by: Kieran Holmes
Version: V1.0
Date: 17th March 2015

Introduction

The initial National Operators Codes database (NOC) was developed by Transport Direct in 2010. That project produced the first version of NOC, the development of which was achieved with input from a number of sources including technical specialists, Traveline and the Transport Direct team. It was recognised that this initial version of NOC would be further developed over time and was produced only as an Excel spreadsheet. A number of the recommendations for further development and improvement of that initial Excel Spreadsheet which were made in 2010 are still relevant today and have been included in this current consultation document

The aim of this original version of NOC was to provide a unique identifier for every bus operator in England, Wales & Scotland and for this code to be applied to all data managed by Traveline for these operators. It was recognised that a unique ID code would allow for the unambiguous identification of operators and their services and would assist with the identification of duplicates, for example between different datasets provided by different organisations e.g. adjacent Traveline regions. At the time it was felt that eventually the scope should be extended to include other operators carrying fare paying passengers (ferry, rail, coach etc.). Another driver for the NOC system is to have a 4 char code that can be used in ATCO.cif where it is the only identifier for an operator apart from name.

In creating an initial list of operator codes, the version of the NOC database that was created in 2010 served a very useful purpose in supporting the continued development of regional and national journey planning systems. Since 2010 NOC has been maintained and extended and various systems have drawn on this data in some form, including Transport Direct's (now defunct) journey planning portal, the National Coach Services Database (NCSD), Electronic Bus Services Registration (EBSR) and the Traveline National Dataset,

Over time the original Excel spreadsheet has been extended to include additional fields and there have also been requests for its contents to be further extended to include additional fields of useful information. However, there are concerns that the maintenance and update of operator codes in the current version of NOC has not been carried out on a comprehensive basis. The NOC Excel spreadsheet currently contains 2,925 rows and 44 columns, has no query tools and therefore is unwieldy to use. To provide an indication of the additional fields that have been added to NOC, the original version of the NOC database contained 30 columns of information.

This review of NOC therefore has a number of aims:-

To assess the accuracy of the data contained in the current version

To review the current content of NOC and determine if any fields currently recorded are superfluous to requirements

To consider what additional information should be recorded in NOC. (In making such suggestions consideration should also be given to the processes required for the creation

National Operator Codes Database Development

and update of this data and the ability of data editors to keep the information up to date)

To redesign NOC based on the conclusions of tasks 2 and 3

To redevelop NOC as a relational database in order to simplify queries on the data

To make this new data available as a web-based tool

To provide a user-friendly front-end for data input and data maintenance

To make the NOC database available as a download under the Open Data provisions

This consultation therefore gives consultees the opportunity to:-

Update and (where necessary) make corrections to the data currently held on NOC

Comment on the usefulness of the contents of the existing version of NOC

Propose any simplification of the existing contents

Propose additional fields of new information that might usefully be included in the new version

Suggest the content of standard outputs of a relational database that would be made available to users

proposed database development

The NOC database was originally developed in 2010 and since then has been extended in terms of both the number of operators that it contains and the data fields that it includes. This review provides the opportunity to improve NOC, not only in terms of its content, but also in its management and the provision of data to users.

The following list considers some of the possible changes that might be made. This is by no means exhaustive but is based on feedback from users and a general review of its management and its purpose.

Possible Database Enhancements

Some of the issues that might be addressed in this review include:-

Provision of the database in a normalised form (multi-table). The current version of NOC is still held as an Excel Spreadsheet in the format that was originally developed in 2010. It is unwieldy to use and holding the data in a relational database would allow for the production of tables that would simplify its contents and use

Inclusion of information on operator depots/garages and contact details/email addresses etc. Traveline often requires further information from operators on which depot a particular bus or service may be operated from in order to send operational queries, a complaint, or lost property request to the right office.

Incorporate generalised information on ticketing and, potentially, about operators participating in common ticketing schemes

Supplementary Data: The following are additional fields that users may wish to hold on an operator. These would require ongoing data management which may not currently be desired by all Traveline regions:

- Postal address (for customer contact)
- Licence holders name(s)
- Company contact name
- Contact email address
- Contact telephone number
- Contact fax number
- Customer service telephone number (complaints, lost property etc.)
- PSV Licence expiry date

National Operator Codes Database Development

PSV Licence status (Valid, Refused, Surrendered, Continuation not sought, Revoked, Withdrawn, Application in progress)

PSV Licence classification (Listed within the TransXChange user guide - StandardNational, StandardInternational, Restricted, SpecialRestricted, CommunityBusPermit)

Local Authority area(s) in which the service operates

EBSR user (i.e. does this operator provide registrations electronically)

Website address details for operators

Provision of a web based data management tool. This tool could provide data creation and edit access for Traveline regional data managers and a database download feature for registered users. Features of the tool would need to include:

The system would need to provisionally allocate of a code which would be confirmed at a later date. This would include a reject edit feature for an administrator.

A full audit trail of edits – when, what (which fields have been edited) and by whom

A fuzzy logic search feature to identify if a ‘new’ operator is already present within the database – this would search all name fields for matches.

A ‘no delete’ or reuse of operator codes rule

Certain fields would need to be mandatory and it may advisable to add others. Currently the mandatory fields in the existing Excel spreadsheet include:-

National Operator Codes Database Development

NOCCode

OperatorPublicName

TravelineOwner

AuditDate

AuditEditor

AuditComment

The development of a new database also provides the opportunity for users to review the content of the existing Excel spreadsheet and consider if fields included in the current data are all still relevant.

In suggesting modifications to the NOC database every effort must be made to avoid any developments that would lead to extensive re-working of, or significant modifications to, existing systems. This will not only cause issues for data providers, but is also likely to have an impact on the current data maintenance processes and delay the creation of a new database.

The DVSA provides a fortnightly download of Open Data which contains all bus registrations based on Traffic Areas <http://data.gov.uk/dataset/local-bus-service-registration>. This download provides information on all registered local bus services and in some fields duplicates the information contained in NOC. There is currently no direct link available between the contents of NOC and the contents of this download. If it were considered desirable that these two tables should be linked, this review may provide the opportunity to also consider that option.

The current NOC also contains information about other public transport operators e.g. rail, coach, air, ferries, light rail/metro, community transport etc. Some of this data is managed centrally; others are managed by the relevant Traveline Region. This review also provides the opportunity to consider if the inclusion of this data is up to date and still relevant.

Fields Currently Recorded in NOC

The fields currently held in NOC are outlined below for reference and explanation. These explain the rationale behind their inclusion in the current version of the system.

Field Name	Description
NationalOperatorCode	<p>Mandatory. National operator code (the key field) as provided for as a field within TXC.</p> <p>Codes for recognised airlines.</p> <p>Codes for national rail companies. Prefixed with an '=' symbol to avoid duplication with airline codes.</p>
OperatorPublicName	<p>Mandatory. The name of the operator as marketed to the public – modifications to this can only be made upon agreement with the Owner Traveline Region</p>
Operator ReferenceName	<p>A name used to help distinguish this operator from other similarly named operators for data providers and managers. Typically a more expansive name or hometown of operations would be added.</p> <p>This field would not normally be presented for public use.</p>
VOSA_PSVLicenseName	<p>The registered company name (as per the first PSV licence in the database). Included to help aid searching for an operator within the dataset.</p>
VOSALicenceNumber1	<p>Unique 9 alphanumeric code assigned by VOSA for Bus Operators licence (and, using separate and unrelated numbering systems, for CT and Taxi and DRT operators)</p>
VOSALicenceNumber2	<p>For additional licence held by this operator</p>
VOSALicenceNumber3	<p>For additional licence held by this operator</p>
VOSALicenceNumber4	<p>For additional licence held by this operator</p>
VOSALicenceNumber5	<p>For additional licence held by this operator</p>
VOSALicenceNumber6	<p>For additional licence held by this operator</p>
VOSALicenceNumber7	<p>For additional licence held by this operator</p>
VOSALicenceNumber8	<p>For additional licence held by this operator</p>
TravelineOwner	<p>Mandatory. A short code, one of SW, WA, YO, WM, SC, NE, NW, LO, EA, EM, SE, NI or Admin to identify who has edit rights over the record.</p> <p>The owner is typically defined by the operating centres on the VOSA</p>

National Operator Codes Database Development

	<p>licence, or if these cover a multi-region area then the registered address of the licence holder.</p> <p>Air, rail and a small number of national coach operators (e.g. National Express coach) are owned by Admin. This is because these need to be consistent with other data sources.</p>
VehicleMode	<i>A useful reference to capture the primary mode of transport for each operator for use in data analysis Air, Bus, Coach, Metro, Ferry, Train, Tram, Underground, DRT and Other.</i>
Parent	<p><i>Parent Company. Used to identify an intermediate hierarchical level within an operators business.</i></p> <p><i>E.g. Lancashire United would have a Parent Company of Blazefield Holdings and an Ultimate Parent Company of Transdev.</i></p>
Grandparent	<i>Additional hierarchical layer within an organisation. Only filled in after Parent and UltimateParent are populate, if required</i>
UltimateParent	<p><i>Used to identify for analysis purposes all companies within a larger bus group.</i></p> <p><i>There are also a few instances of part ownership by major groups, for the purposes of this process a rule of minimum 51% ownership is applied for it to be considered part of a parent company.</i></p> <p><i>(Dependant on the use made of this data, consideration may be given to rationalisation of these three fields)</i></p>
EBSRAgent	<i>The name of an agent who prepares and submits EBSR files for the operator concerned. E.g. a council providing this service to local bus operators</i>
LO	<p><i>The current Traveline Regional Code being used by the London Region</i></p> <p><i>This field may eventually be phased out as data suppliers fully adopt the national code.</i></p>
SW	<i>As above for South West</i>
WM	<i>As above for West Midlands</i>
WA	<i>As above for Wales</i>
YO	<i>As above for Yorkshire</i>
NW	<i>As above for North West</i>
NE	<i>As above for North East</i>
SC	<i>As above for Scotland</i>
SE	<i>As above for South East</i>
EA	<i>As above for East Anglia</i>
EM	<i>As above for East Midlands</i>

National Operator Codes Database Development

NI	<i>As above for Northern Ireland</i>
NX	<i>Codes used for National Express</i>
Megabus	<i>Required by some data users to identify these services</i>
New Bharat	<i>Required by some data users to identify these services</i>
Terravision	<i>Required by some data users to identify these services</i>
Thandi	<i>Required by some data users to identify these services</i>
EasyBus	<i>Required by some data users to identify these services</i>
Yorks RT	<i>Test field to hold different codes used in the Yorks real time system.</i>
	<i>It was expected that this would be needed by other regions but no others added as yet</i>
Travel Enq	<i>Travel Enquiry phone number</i>
Comment	<i>A free text note for useful remarks/comments about this record – e.g. for suspected duplicates, flagging an operator with an expired license etc.</i>
AuditEditDate	Mandatory. <i>Date of update</i>
AuditEditor	Mandatory: <i>Person who made change</i>
AuditComment	Mandatory. <i>Details of changes made</i>
Duplicate	<i>To identify instances of duplicate codes</i>
Date Ceased	<i>Date when operator ceased to operate</i>
Cessation Comment	<i>Explanation of cessation</i>

The Consultation Process

Traveline will consult with interested parties via the PTIC Group. Members of PTIC are welcome to seek the views of other related users. Current users of Traveline Open Data are also being invited to respond.

Responses should be emailed to admin@traveline.info. That address can also be used for any queries.