NCSD Cif File Structure and Output Files

1. Table of Contents

2. Overview

2.1 Distribution

All publicly available NCSD data can be downloaded from the NCSD ftp site. Details of this site are available on request.

2.2 ATCO.cif

The NCSD export is compatible with the ATCO CIF 5.10 specification, which can be downloaded from:

http://www.travelinedata.org.uk/CIF/atco-cif-spec.pdf

2.3 Reference Data

Reference data in the NCSD adheres to the following general principles:

- Coding of data to NaPTAN and National Gazetteer standards
- Ensuring uniqueness of operator codes from data sourced from multiple providers
- Addition of stop interchange times, including those which make stops unsuitable for interchange purposes. This data will only be added where supplied by the Operator and where it does not prejudice other operators.
- Cluster information provided by an operator will be used where the information relates to stops in the NCSD
- Additional clusters and walk links will be added to the NCSD to form a connected network.
- Modification of stop names where the existing stop name is unsuitable for use with a national rather than local mode of travel.

2.4 Service Data

Service data in the NCSD adheres to the following general principles:

- Service data is only retained where it applies to valid NaPTAN coded stops
- Addition of association records for guaranteed connections records
- Creation of journey identifiers for use with association records

In both cases, this data is only added where provided by the operator.

Routing guide information from individual operators is encoded within the main data file where possible.

3.   Currently included data

The table below summarises the ATCO.cif record types being exported from NCSD.

Record Header	Notes	Operator(s)
QS	Operator codes are as supplied, any local codes are matched against National Operator Codes(NOC) and the NOC codes are used in the export file	All
QE	As supplied by operator, or added by NCSD through pre-processing files	All
QN	As supplied by operator, or added by NCSD through pre-processing files	All
QO / QI / QT	All service data is converted to the NaPTAN codes All are set to T1 timing points, unless specified by supplier NaPTAN: Non-NaPTAN stop codes will be ignored.	All
QL	Name generated by NCSD using NaPTAN fields to make a meaningful text description within 48 character limit	All
QB	9000 stop coordinates are replaced with the coordinates of the equivalent NaPTAN stop Other coordinates taken from NaPTAN	NX All
QC	9000 physical stop clustered with NaPTAN equivalent Any other NaPTAN stops within 500m also included in same cluster	All
QP	Using the latest version of NOC these are created for all services	All
QG / QJ	As supplied. 999 minute interchange times are supplied by NX to prevent interchanges at defined stops	NX
QW	Included where supplied. NCSD generates fixed links BUT adds distance to end of record	All
QD	Use QD records of supplied in source files or create new QD records	All
QX / QY	Converted to the NaPTAN code of the association	NX

4.   NCSD Output

4.1   QS - Journey Header Record

4.1.1   Midnight

The Journey Header record (QS) will show the actual dates of operation. Vehicles running over midnight may affect the dates in these records. For example a journey from 22:00 to 01:30 would have different dates in the QS record.

All service timings (QO, QI, QT) will be between 00:01 and 23:59.

4.2   QE – Journey Date Running Record

These will be output as supplied.

4.3   QN - Journey Note Record

Notes records may be added to the data where it is felt that extra information needs to be conveyed for particular journeys. This will be achieved using QN records.

Notes records supplied will be output without any modification

These records should be displayed to the end user.

4.4   QO QI QT – Origin / Intermediate / Destination Stop Records

4.4.1   Renumbering of stops in services (QS, QO, QI, QT records)

The following examples show a before and after example of changes to service records. Changes are in bold in the after example.

```
ZSNFKO210 210
QSNFK  LH    2004060120040829000011  210       FSD
I
QO900063086A  0945   T1
QI900063057C  09550955P   T1
QI900057217   10451045B   T1
QI900057286   11001100B   T1
QI900057257C  11101110P   T1
QI900048023   12251225B   T1
QI900033343   13101310B   T1
QI900033023A  13301335B   T1
QI900033013C  13501350B   T1
QI900019023   14051405B   T1
QT900019063   1425   T1
```

```
QSNFK  210LHI2004083120041224000011  210       FSD
I
QO4400GTWSTHX70945   T1
QI4400CY0390  09550955P   T1
QI490000104N  10451045B   T1
QI4900080124  11001100B   T1
QI490015598S  11101110P   T1
QI340000072BS 12251225B   T1
QI3800C638703 13101310B   T1
QI4300210301  13301335B   T1
QI4300093993  13501350B   T1
QI43000875320 14051405B   T1
QT43000700101 1425   T1
```

Stop numbers have been converted from National Express numbers to NaPTAN numbers. The relationship between National Express stop numbers and NaPTAN numbers is preserved in the reference data.

The journey identifier has been changed from the National Express flight identifier to a composite of service number (210) (characters 1-3), flight identifier (LH) (characters 4-5) and direction (I) (character 6). This, in conjunction with the operator code, gives an identifier suitable (210LHI) for use with association records for guaranteed connections.

4.5   QL – Location Record

4.5.1   Stop Names

Refer to section 6 for a description of the derivation of stop names.

4.6   QB – Additional Location Record

4.6.1   Location Records (QL, QB records)

Additional location (QL and QB) records have been added to the data so that the file contains both National Express numbered and named stops and the corresponding NaPTAN numbered and named stops.

```
QLN900018283    Malvern Link:
QBN900018283    3785373024775973Worcs               Malvern Link
QLN900018283A   opp Co-op Supermarket
QBN900018283A   3785373024775973Worcs               Malvern Link
QLN900018283C   o/s Co-op Supermarket
QBN900018283C   3785373024775973Worcs               Malvern Link


QLN900019023    Dudley: Bus Stn
QBN900019023    3946405329038402Dudley              Dudley
```

```
QLN900018283    Malvern Link:                                  7
QBN900018283    378603  0247822 Worcs               Malvern Link
QLN2000G009802  Lower Howsell: Malvern Link, op. Pioneer Store   BE0027780
QBN2000G009802  378604  0247829 Worcs               Malvern Link
QLN900018283A   opp Co-op Supermarket                          8
QBN900018283A   378604  0247829 Worcs               Malvern Link
QLN2000G009803  Lower Howsell: Malvern Link, Pioneer Store       BE0027780
QBN2000G009803  378602  0247815 Worcs               Malvern Link
QLN900018283C   o/s Co-op Supermarket                          8
QBN900018283C   378602  0247815 Worcs               Malvern Link

QLN43000875320  Dudley: DUDLEY BUS STATION, Stands A>S            BE0057924
QBN43000875320  394697  0290384 Dudley
QLN900019023    Dudley: Bus Stn                                7
QBN900019023    394697  0290384 Dudley
```

National Express supply group codes (9000 + 5 digit national express number without a suffix) and stop codes (9000 + 5 digit national express number + suffix letter) in their data. The stop codes are used in the service data if there are separate stops for each direction of a service and group codes are used where there is a single stop covering both directions.

In the above example Malvern Link has both a group code (900018283) and stop codes (900018283A and 900018283C), whereas Dudley just has a single code to denote both the group and the stop (900019023). The Group code in Malvern Link is not matched to a NaPTAN stop and there is no corresponding NaPTAN record, but the stop records have corresponding NaPTAN records (2000G009802 and 2000G009803). The record in Dudley is matched to a single NaPTAN stop (43000875320).

The QB Town field is populated from the town as specified in the NaPTAN record. Where this is missing, this field will be derived from the National Gazetteer locality record.

The QB District field is populated with the NPTG AdminAreaID (a number in the range 1 to 142).

There are a number of places (including Victoria Coach Station) where several different National Express stops have been allocated to the same NaPTAN stop. This will be addressed in future releases of the data as the National Express to NaPTAN translation table is refined.

In the case of VCS, separate NaPTANs exist for the arrivals and departures hall and surrounding areas. NCSD had previously created a process to identify the correct NaPTAN, though this is temporarily disabled.

National Express group numbered records have been given a gazetteer code of 7 and stop numbered records have been given a gazetteer code of 8.

Coordinates have been changed on a National Express numbered record to match those on the corresponding NaPTAN numbered record. Coordinates of National Express group records have been taken as the mean of the associated stop records if there was no associated NaPTAN record.

4.7 QC – Cluster Record

The new NaPTAN numbers have been added to the cluster records, retaining the original National Express cluster number

```
QCN900018283    Malvern Link                                  900018283
QCN900018283    Malvern Link                                  900018283A
QCN900018283    Malvern Link                                  900018283C
QCN900019023    Dudley Bus Stn                                900019023
```

```
QCN900018283    Malvern Link                                  900018283
QCN900018283    Malvern Link                                  900018283A
QCN900018283    Malvern Link                                  2000G009802
QCN900018283    Malvern Link                                  900018283C
QCN900018283    Malvern Link                                  2000G009803
QCN900019023    Dudley Bus Stn                                900019023
QCN900019023    Dudley Bus Stn                                43000875320
```

4.8 QP – Operator Record

4.8.1 Retained Operator Codes

Services for the following National Express operator codes will be retained.

Operator Code	Trading Name
AB	National Express Airport
FK	National Express Airport
IN	Eurolines – Irish Night
IS	Eurolines
JL	National Express Airport
NX	National Express
SD	Eurolines – Shipside Day

SH	National Express Shuttle
WR	Woking Railair

This list will be reviewed from time to time to take account of National Express operational changes.

Other operator data provided by National Express will be analysed, and where a better data source exists, the NX data will be replaced with this alternative data. Where no other data source exists, the NX data will be retained.

4.9 QG QJ – Location and Cluster Interchange Records

The ATCO.cif format does not allow interchange times to be related to specific operator(s) so the values provided by one operator affect all operators at a given location.

Interchange data received by NCSD from National Express includes some extremely high interchange times, effectively prohibiting the interchange of vehicles at the specified locations. Increasingly, as more operators are added to the NCSD, these high interchange times are prohibiting acceptable, inter-operator interchanges.

NCSD has therefore decided to replace any interchange time as defined by one operator (or group of operators) where another operator uses the same location with a default value of 15 minutes.

The original value of any interchange data modified in this way will be written to a separate file – nx-interchangetimes.cif - and exported to the NCSD ftp site as part of the weekly process, so that downstream systems can chose to reintroduce this data if they wish.

The first example shows the QJ records give cluster interchange times as supplied by National Express and the new QG records set the stop interchange times of stops within the cluster to the same value.

```
QJN900018283    999
QJN900019023    999
```

```
QJN900018283    999
QGN900018283    900018283    999
QGN900018283A   900018283A   999
QGN2000G009802  2000G009802  999
QGN900018283C   900018283C   999
QGN2000G009803  2000G009803  999
QJN900019023    999
QGN900019023    900019023    999
QGN43000875320  43000875320  999
```

The second example shows where National Express allow an interchange provided there is the specified time between arrival and departure of their services.

```
QJN900017353    25
```

```
QJN900017353    25
QGN900017353    900017353    25
QGN900017353A   900017353A   25
```

```
QGN16000GL1271 16000GL1271 25
QGN900017353B  900017353B  25
QGN16000GL1272 16000GL1272 25
```

In both of the preceding examples, all the data would be written to the nx-interchangetimes.cif.

Note that the interchange time considers walk element in its makeup.

4.10 QW – Cluster Walk Link Record

As more operators are added to the National Coach Services Database, it has become apparent that in many locations, the stops used by different operators are sufficiently far apart that the stops can neither be clustered together, nor a sensible length walk link offered to connect the two locations. For example in Birmingham, Digbeth (National Express) and Snow Hill (Megabus) are more than a mile apart across the City Centre. It is the ambition of the NCSD to provide a connected, coach-only network, and not to hold data relating to local services.

It is recognised that journey planning systems might have local public transport information to connect such coach services, or might be able to estimate the time and / or distance of these connections. The NCSD output will identify the stops to be connected that might otherwise fall outside of the normal limits.

Where the NCSD identifies a fixed link to join stops used by distinct operators, that the ATCO.cif QW record be used, but that BOTH a time and distance be exported. A standard QW record identifies the cluster at each end of the fixed link and the minimum time to traverse the link. Each QW record exported by the NCSD will only contain data about a single fixed link and that the four digits after the time field will be the estimated distance, in metres. The time field will be calculated based on a walking speed of 66 m/min (15min/km). The distance will be calculated as the crow flies distance between the stops multiplied by 1.4.

The exported QW record format is:

Record Identity	2 (1)	A	QW - Cluster Walk Link
Transaction Type	1 (3)	A	N = New
			D = Delete
			R = Revise
Origin Cluster	12 (4)	A	Short code form of cluster
Destination Cluster	12 (16)	A	Short code form of cluster
Interchange time	3 (28)	I	Minimum travel time from origin cluster to destination cluster
Origin Cluster	12 (31)	A	Short code form of cluster
Destination Cluster	12 (43)	A	Short code form of cluster
Interchange time	3 (55)	I	Minimum travel time from origin cluster to destination cluster
Origin Cluster	12 (58)	A	Short code form of cluster
Destination Cluster	12 (70)	A	Short code form of cluster
Interchange time	3 (82)	I	Minimum travel time from origin cluster to destination cluster
Origin Cluster	12 (85)	A	Short code form of cluster
Destination Cluster	12 (97)	A	Short code form of cluster
Interchange time	3 (109)	I	Minimum travel time from origin cluster to destination cluster
Interchange distance	4 (112)	I	Estimated distance from origin to destination, in

			metres.

The walk time created between the Coach Exchange Points using the "crow flies" method may not be long enough to circumnavigate natural barriers. A method to amend or remove these walk links from the output has been implemented.

Please also note, that any QW records received in operator or traveline supplied data (though there are currently none) would be exported without modification.

4.11   QD – Route Description Record

The National Express source data provides the QD record to represent route description and is fed into the NCSD export process accordingly. Other operators have no QD records.

Where QD records are not provided for a service, new QD records are then created for these services by using the start and end stops of the services.

4.12   QX QY – Route and Journey Association Records

4.12.1 Route Association Records (QX records)

Currently no QX records are supplied by the imported data and consequently they are not output.

4.12.2 Journey Association Records (QY records)

Association records have been added to the National Coach export file to show guaranteed connections that break National Express's normal minimum interchange times. These records use the new journey identifiers.

```
QYNNX   328LAONX   502LGO20050307200603051111113600SOC30962G
QYNSH   070KEINX   380LBI2005030720060305111111145003022    G
QYNNX   320JBOFK   201MEI20050307200603051000005710AWA10296G
QYNSH   420OUINX   531KAO2005030720060305111111143002103001 G
```

5. National Express Issues

The NCSD receives a weekly update of National Express data in ATCO CIF format. Major changes between the National Express and the NCSD versions of the data are:-

- Translation of National Express stop numbers to NaPTAN numbers on services and the addition of stop information with NaPTAN numbers (both National Express and NaPTAN codes are included in the file).

- National Express easting and northing coordinates have been standardised. National Express coordinates have been replaced by coordinates taken from NaPTAN

- Creation of ATCO CIF cluster information between an original National Express physical stop code and the equivalent NaPTAN stop code. This is designed to show the relationship between National Express stop number and NaPTAN numbers and is not designed to show information in NaPTAN grouping records.

- Addition of stop interchange times, including those which make stops unsuitable for interchange purposes

- Addition of association records for guaranteed connections where the connection breaks National Express's normal interchange times.

5.1 NX Business Rules

ATCO.cif is capable of conveying the following business rules:

5.1.1 Location Availability

On a particular day a location is closed and vehicle events should be ignored. This would be delivered in the NCSD as Journey Variations.

5.1.2 Interchange Points / Connection Times

Some stops are not to be used for interchanging. Interchange points are currently defined in ATCO.cif as cluster interchanges and NCSD will add stop interchanges too.

Note: Non-interchange points are defined by National Express with an interchange time of 999 minutes.

5.1.3 Pick up / Drop off

A standard feature in ATCO.cif. The default activity is Both.

5.1.4 Composite Flights (guaranteed connections)

National Express supply this data (QX and QY records) based on 9000 group codes, within the NX_atco.cif. (This was previously sent from National Express as a separate file, tb_comp_flt.csv, but is now superseded by the current NX_atco.CIF.) These are converted to the equivalent NaPTAN in the NCSD export.

5.1.5 Multi-leg only

Some services can only be used as part of a multi-leg itinerary. It has been agreed with National Express that this is a commercial rule that they have in their system which does not need to be replicated in external planners.

6. Stop Names

The following table describe the origin of stop names in each of the output files where stop name fields are included.

File Name	Stop Type	Derivation of Stop Name
NCSD ATCO.cif	National Express 9000 codes (QL records)	The majority of stop names are taken from the NX data provided. There are exceptions when the provided name is considered to be incomplete or misleading, in which case NCSD will assign a more appropriate name. These are on a case-by-case basis.
	NaPTAN stop codes (QL records)	The default is to use: Parent - Locality: Common Name, Identifier (note syntax and NaPTAN fields used) However, this rule does not generate a meaningful stop name in some cases and in others the resulting text would be over the CIF limit of 48 characters. In these cases NCSD will override the default rule using any other NaPTAN fields to create the stop name. This is on a case-by-case basis.
NaPTANLookUp.csv	All stops in file	This file contains an ATCOName field which is the modified NaPTAN name (see rule in NCSD ATCO.cif for NaPTAN stop records above). The file also contains the NaPTAN fields for: LocalityName, CommonName, Street, Identifier, Landmark, all taken directly from NaPTAN
NCSDStopNames.csv	All stops in file	ATCOName field is the modified NaPTAN name, see rule in NCSD ATCO.cif for NaPTAN stop records above.
NCSDStops.csv	All stops in file	All names fields are taken from NaPTAN
StopsInCluster.csv	All stops in file	Stop Name format as per NaPTANLookUp.csv

7. NPTG issues

7.1 Coach Exchanges

.Adding / Deleting Coach Exchanges

Additionally, NCSD will identify new locations where a CEP should be created. This will be based on the existence of NCSD services to significant locations not currently served by an existing CEP, and will be reviewed by the DfT. NCSD will advise Landmark of the details of new CEPs (the CEP codes will be in the range 900098xxx). Landmark will assign localities to the new CEPs and update NPTG accordingly.

Any CEP which has not been used for a period of time will be identified as part of the weekly build. Landmark will be asked to remove these CEPs. Currently this review takes place fortnightly and any CEP not used for one month will be identified. This frequency may be reviewed.

The analysis described in this section does not create any publicly visible files.

8. Exported data

8.1 Delivery

The NCSD is distributed via an ftp site.

Usernames and Passwords have been distributed to those who are authorised to use the data. Others wishing to do so should contact the NCSD team.

8.2 Timing of Weekly build

The data is built by 20:00 each Friday.

Extraordinary circumstances that affect the delivery or quality of data will be communicated by email to known users of the data.

8.3 Files Uploaded to ftp site

The following files are uploaded to the ftp site:

Directory / Filename	Description	Update frequency
NCSD-ATCO.zip	Zip file containing the NCSD-ATCO.cif data	Weekly
WeeklyQualityCheck.txt	Summary of data issues in the current NCSD-ATCO.cif file	Weekly
IncludedServices.xls	Current services in the NCSD	Weekly
NaPTANLookUp.csv	Conversion of NX9000 codes to NaPTAN equivalent	Weekly
NCSDStops.csv	Extract of NaPTAN v1 for each stop in the NCSD build	Weekly
NCSDStopNames.csv	Name generated by NCSD Contains only the ATCOCode, ATCOName and National Express Physical StopID	Weekly
StopsInClusterReport	Stops and services for each cluster	Weekly
Archive yyyymmdd.zip	ZIP archive of data	Weekly
CoachReferences.csv	Contains the NPTG v2 data derived from the NCSD	Weekly
UnusedCEPs.csv	List of NPTG Coach Exchange Points which are not used in the	Weekly

	current NCSD build	
CurrentBarriers.csv	Cluster walk links to discard	Weekly
_NCSD ftp site - guide to contents	This document is a brief explanation of each of the files contained on the NCSD FTP page	Weekly
140811 NCSD Guidance v2.1	Further background information on NCSD	Weekly

8.3.1 NCSD-ATCO.cif file

The NCSD is exported as a single file "NCSD-ATCO.cif", containing both service and reference data.

8.3.2 NaPTAN extract

A further CSV file is uploaded to the ftp site. This file – NCSDStops.csv – contains the NaPTAN v1 fields for all the stops in the current build of the NCSD.

It also contains a pseudo-NaPTAN entry for 998 prefixed stops. These stops have been generated by the NCSD processes to cater for locations awaiting NaPTAN codes to be allocated. The 998 codes include

- o StopID
- o Easting / Northing
- o Stop Name
- o Locality ID / Name to which the stop is assigned.
- o All 998 stops are marked as
  - o BusStopType = "MKD,
  - o StopType = "BCT",
  - o BusRegistrationStatus = "OTH"
  - o RecordStatus = "ACT".

This file may include NaPTAN codes for stops which are marked as DELeted or PENding but used in that weeks' data build. Users of the data may wish to convert these types before importing them into their systems.

8.3.3 Interchange Times

Any QG or QJ records that were stripped out of the source data are added to an

"nx-interchangetimes.cif" file, which will contain only QG or QJ records.

8.3.4 Unused Coach Exchange Points

At the end of each weekly build, the current NPTG Coach Exchange Points table is compared to the CEPs in use in the NCSD data. Any NPTG CEPs not in use are written to this CSV. DfT, NCSD and Landmark may update NPTG on an occasional basis to reflect CEPs that are no longer used.

8.3.5 Weekly Archive

A single ZIP file containing all the files from that week's build:

8.3.6 Weekly Quality Check

Additionally, further tests are run on the data each week to check a number of parameters. The results of this script are recorded in the file "WeeklyQualityCheck.txt. Additional checks may be added.

Title of test	Reason for inclusion	User Action
Date of report	Are you looking at the latest data ?	Check the date is recent
NaPTAN details	Confirm that the NaPTAN data used in the build is current	Check the date is less than one week old and number of stops is within reasonable limits (380,000+)
Routes, Clusters, Vehicle Journeys	Indicators of NCSD contents	Check the values in this build are consistent with the benchmark values
Incorrect Stop Types	Indicators of NaPTAN stop codes used	All stops types should be "BCS" or "BCT" or "BCQ" or "FER". Others are flagged.
Time Running Backwards	Indicator of timetables with departure time prior to arrival time.	This should not occur. Suppliers to be asked to correct data at source.
DELeted or PENding NaPTAN codes	Indicators of NaPTAN stop codes used	All stops should be ACTive. Others are flagged.
Services added / removed since last week	A list of data source / operator name / service numbers added or removed since the last build	Information only. Data providers will be contacted to confirm services removed are intentional.
Duplicated Services	A list of details where the service number is supplied from more than one source	Information only. A check to ensure a service is not duplicated. Duplicates will be removed.
Summary of operators and services	The number of services by operator and data source.	Information only.
Detailed list of services	A list of each service number, operator and data source in the NCSD.	Information only.
Journeys with over 25 stops	A listing of routes for review as they have a large number of stops, which are included in the NCSD.	Information only.