



## **BikeLife - Cycle Infrastructure Categorisation Methodology (WP1)**

### **Purpose of this Document**

*The purpose of this document is to inform the user on how BikeLife's Cycle Infrastructure Categorisation Methodology was achieved and could potentially be carried forward.*

## Contents

1	Overview.....	3
2	Data and Software.....	4
2.1	Datasets.....	4
2.2	Software.....	4
2.2.1	QGIS (3.8.3) .....	4
2.2.2	PostgreSQL .....	4
3	Processes.....	5
3.1	Data Collation.....	5
3.1.1	Identify Features to be Migrated .....	5
3.1.2	Migrate Feature Geometry .....	5
3.1.3	Migrate Feature Attributes .....	7
3.2	Applying BikeLife Categorisation .....	7
3.2.1	Rules for Categorisation.....	7
3.2.2	Categorisation Processes .....	8
3.3	Example of Typical Process .....	9
4	Issues encountered .....	12
5	Future recommendations.....	13
	Appendix 1 - Geometry Edits.....	14

---

# 1 Overview

---

For this work package Compass has created a single dataset containing all of the cycle infrastructure features for the Greater Dublin Area using the input datasets set out in [Section 2.1](#) of this document. With the data collated into a single dataset each line segment has been categorised according to the specifications set out in the document *Bike Life: Guidance for cities S1 Total Length of cycle routes of various types*.

---

## 2 Data and Software

---

### 2.1 Datasets

The following datasets were assessed as part of this work package

- GDA\_Base\_Cycle\_Network (GDA)
- NTASchemes (NTA)
- JPLinks\_Metrop (JPL)

GDA was identified as the primary dataset by the NTA. Any extra cycle infrastructure found in NTA or JPL had been merged into GDA.

No geometry edits were made to any features apart from those specified in [Appendix 1 – Geometry Edits](#)

### 2.2 Software

#### 2.2.1 QGIS [\(3.8.3\)](#)

QGIS was chosen as the desktop GIS environment to be used throughout the project as it allowed for the use of the below plugins:

- [go2streetview](#) allows for the constant tracking between the linear data reference and Google's Street View service and was used extensively during the data categorisation process.
- [HCMGIS](#) provides Google satellite imagery as base maps.

Both plugins are installed directly in QGIS via the *Plugins* menu.

#### 2.2.2 PostgreSQL

PostgreSQL was chosen as the main datastore for this task, both for its interoperability and the spatial extension PostGIS.

The file geodatabase originally received from the NTA was imported to PostgreSQL using the GDAL module [ogr2ogr](#).

The final data was exported from PostgreSQL to file geodatabase through ARCSDE.

---

## 3 Processes

---

The work package was broken down into two distinct phases:

1. Data collation: the process of analysing the received datasets in order to identify cycle infrastructure and merge into one common base dataset.
2. Applying BikeLife Attribution

### 3.1 Data Collation

Collecting the three datasets into a common base dataset was comprised of three sequential tasks:

1. The identification of cycle infrastructure to be migrated to GDA
2. The copying of line geometries and *datasource* value to GDA
3. Using a spatial join to transfer a copied line geometry's relevant attributes to GDA

#### 3.1.1 Identify Features to be Migrated

A grid of 2 km<sup>2</sup> cells was created at the extent of the base dataset. Each grid square was checked in sequence to identify features from the JPL and NTA datasets which represent cycle network infrastructure. This method allows the data technician to keep track of coverage as the data is processed.

Two methods were used to identify and transfer features:

- Layer styling to identify if a segment has been:
  - checked, identified and transferred
  - checked, identified and rejected
  - not yet checked
- JLP data was filtered to only show features that represent cycle infrastructure based on received JPL metadata document.

#### 3.1.2 Migrate Feature Geometry

Once a feature is identified it is selected, copied & pasted into GDA, and saved. This process inserts the identified geometry from the source dataset to the GDA dataset.

Four new attribute fields were added to the GDA dataset and are described in Table 1 below. These are essential not only for keeping track of what data has been copied and where the originated from, but also for creating the spatial join required for the further transfer of cycle infrastructure related attributes.

The values for *datasource* are added manually each time a new set of line segments are copied to the GDA base dataset. When geometries are copied and pasted in QGIS the copied line segments are automatically selected in the destination dataset. Use the attribute table Field Calculator to update the selected fields with the relevant value for *datasource* (Figure 1), making sure the box to 'Only Update Selected Features' is checked (this should be the case by default).

Table 1 - Data Fields Added to GDA during data collation phase

Filed Name	Data Type	Notes
oid	serial	Set as the table primary key. This autoincrement datatype is recommended for editing of postGIS data tables in the QGIS desktop environment.
datasource	varchar	The origin dataset for each line segment.
datasourceid	integer	Unique id values transferred from origin dataset. JPL : <i>id</i> NTA : <i>gid</i> GDA : <i>objected_1</i>
overlapcode	integer	Sequential integer assigned to each segment where JPL or NTA overlapped with GDA but also providing some extra coverage or detail. Later split and deleted (see <a href="#">Appendix 1 – Geometry Edits</a> )

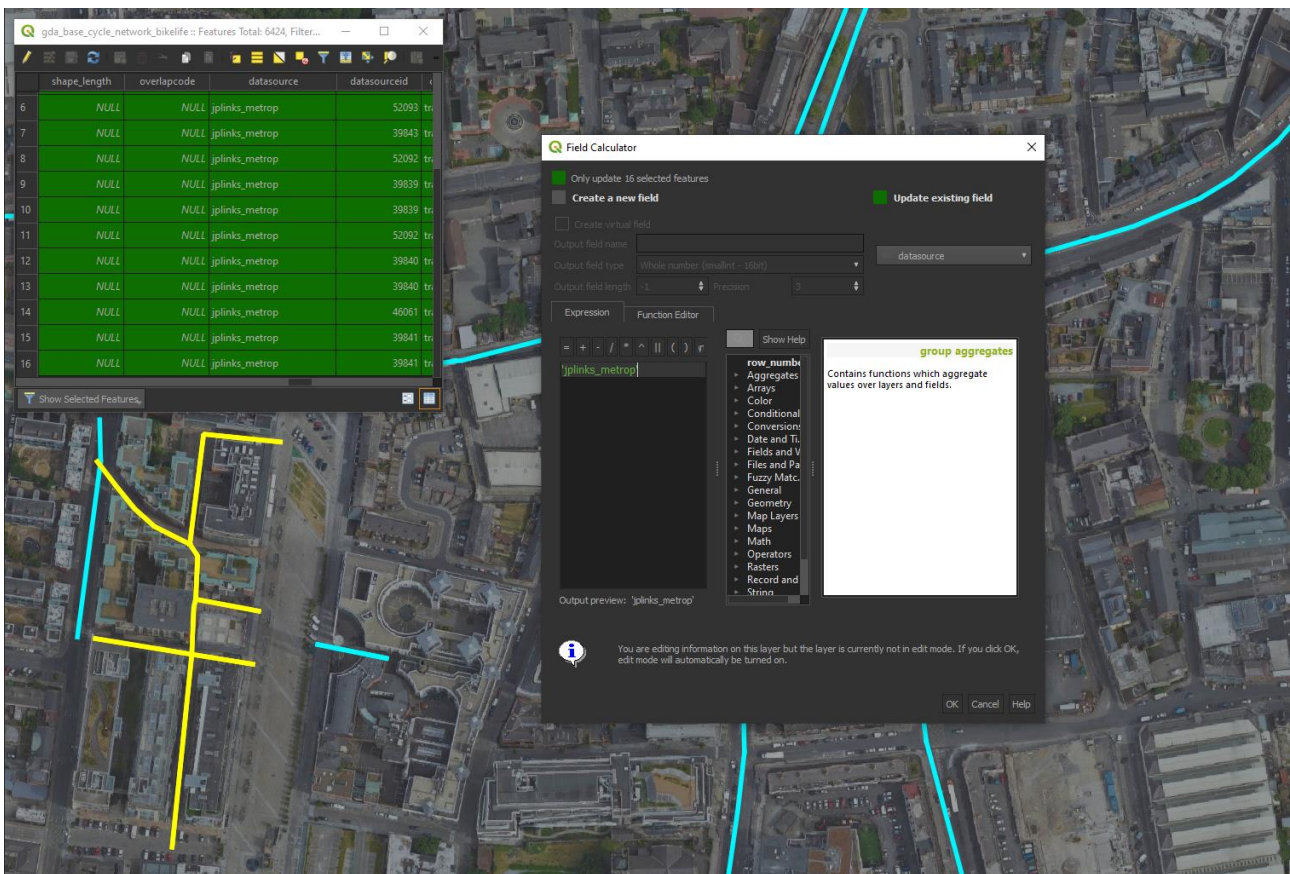


Figure 1 - Adding datasource attribute to copied geometries in QGIS

### 3.1.3 Migrate Feature Attributes

Data is transferred by:

1. Applying a buffer to the newly copied geometries in GDA
2. Selecting features from the source data that are contained by the buffer and have the relevant value for *datasource* applied in the previous step.

With this method any of the data fields in the source datasets are available for transfer to the GDA.

This can be achieved by issuing an update SQL on the data or alternatively, via common tools available in most desktop GIS applications.

For example, to transfer an id from JPL as *datasourceid* in GDA using SQL:

```
SET
GDA.datasourceid = JPL.id
WHERE
GDA.datasource = 'jplinks_metrop'
AND
GDA.bufferpolygon
CONTAINS
JPL.linesegment
```

## 3.2 Applying BikeLife Categorisation

### 3.2.1 Rules for Categorisation

A workshop was held on 10<sup>th</sup> October 2019 at the NTA offices between the NTA and Compass. At this workshop the NTA outlined a series of rules to act as guidelines for categorisation. These are:

- Exclusion of potential cycle features where:
  - Pedestrianised Roads (e.g. Henry Street)
  - Presence of steps in route
  - Parks/entrances with gates
  - Private land with gates
  - Off beaten track/looped routes
- Semicolon classifications are only used where 2 or more types of cycle infrastructure are on the one side of the road. The first class used should be the furthest from the road e.g. ClassificationType1 = "Segregated;BusLane" while ClassificationType2 = "SharedUse;BusLane"
- Kissing gates are open for cycling and are not seen as a hindrance to cycling

### 3.2.2 Categorisation Processes

Using the go2streetview plugin outlined in [Section 2.2](#), each line segment is checked against Goggle Street View. Using this process, we can almost always identify the specific BikeLife category to apply to a segment. Segments where we were not able to classify based on Street View were noted and reviewed at a workshop with the NTA held on 10/10/19.

In order to efficiently tag each line segment a series of Boolean fields have been added to the dataset.

Table 2 - Data Fields Added to GDA during data categorisation phase

Field Name	Data Type	Notes
onside	boolean	if true we have found cycle infrastructure on one side of the road only
twoway	boolean	if true this line represents one segment of a two-way system. Used in the calculation of route length
oneplus	boolean	if true, more than one type exists on one side of the road
ndci	boolean	no discernible cycle infrastructure, if true no clear and obvious cycle infrastructure is present
cyclelane	boolean	If true, segment is of type cyclelane
buslane	boolean	If true, segment is of type buslane
segregatedcyclelane	boolean	If true, segment is of type segregatedcyclelane
shareduse	boolean	If true, segment is of type shareduse
trafficfree	boolean	If true, segment is of type trafficfree
signedroute	boolean	If true, segment is of type signedroute
checked	boolean	If true, segment has been viewed and tagged
cdo_categories	vchar	Intermediate field used to visualise the data and aid the categorisation process
classificationtype1	vchar	BikeLife category 1
classificationtype2	vchar	BikeLife category 2

With this in place the data editor can simply click the relevant fields, this approach greatly speeds up the process and almost completely removes human error.

A series of SQL statements are used to process the Boolean fields into the required BikeLife columns:

- classificationtype1
- classificationtype2

For Example:



*IF onside AND buslane are TRUE  
THEN classificationtype1 = 'buslane' and classificationtype2 = NULL*

*IF onside is FALSE AND buslane is TRUE THEN  
classificationtype1 = 'buslane' and classificationtype2 = 'buslane'*

### 3.3 Example of Typical Process

Below is a series of steps that represent a typical categorisation process and how a data technician might approach the analysis and categorisation of a road segment. Here the steps are laid out as sequential however in reality step 4 can happen before or inbetween steps 2 and 3. In the vast majority of cases, the technician would have a clear indication of how to classify each segment of this road by step 4.

1. Select line segments to be assigned a BikeLife category. Most often a junction to junction selection works best however it might be that the best selection covers multiple junctions (arterial roads) or single line segments (city centre roads) (Figure 2).
2. Zoom into the junction at one end (Figure 3).
3. Zoom into the junction at the other end (Figure 4).
4. Check the junction on Street View for some signage or road markings as indication of cycle infrastructure (Figure 5).
5. If none exist, repeat this process at each junction along the road in both directions (Figure 6).
6. Once a clear picture of the selected segments should be categorised is formed, open the data attribute table and check the relevant Boolean fields for each segment (Figure 7)

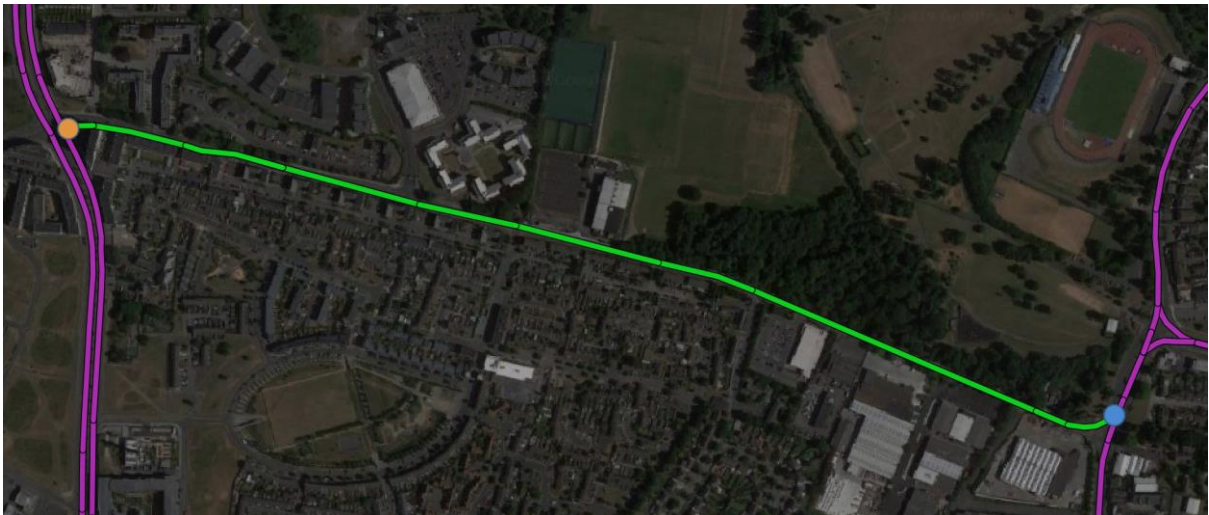


Figure 2 – A suitable section of data is selected

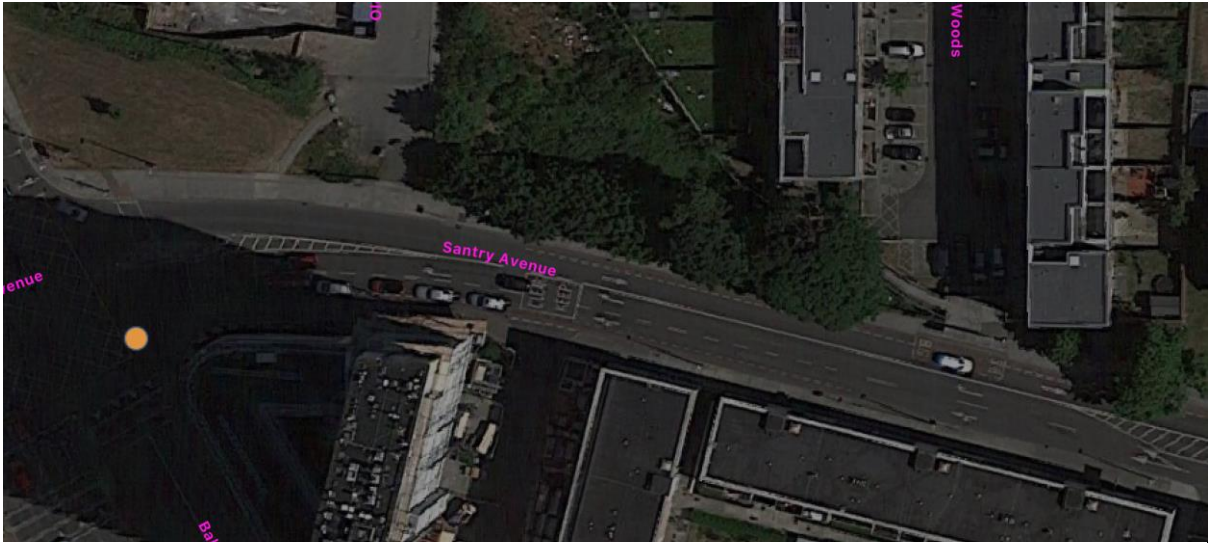


Figure 3 – Zoomed into the junction at the west end. There is a good indication that this is a Cyclane running west to east and a Buslane running east/west converging to a Cyclane just before this junction.



Figure 4 - Zoomed into the junction at the east end. There is no clear indication of cycle infrastructure.



Figure 5 - Check junction on Street View for some signage or road markings as indication of cycle infrastructure.





Figure 6 - Checking Street View at each junction in both directions

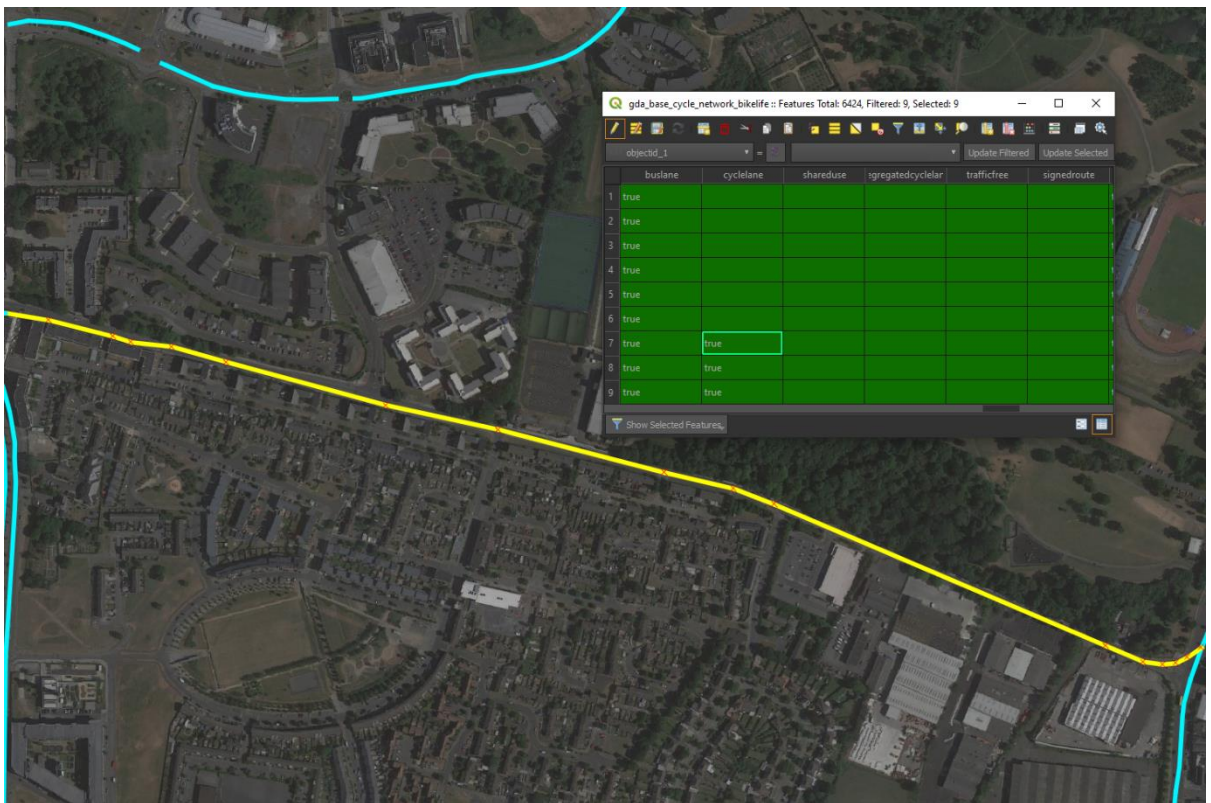


Figure 7 - Updating Boolean fields in QGIS

---

## 4 Issues encountered

---

### **Issue: Changeability of cycle infrastructure**

The Changeability of cycle infrastructure along a road or street can greatly increase the time it takes too categorise.

#### **Mitigation:**

More time given to categorisation

### **Issue: Nonsense Data**

Rare, but sometimes the data provided did not accurately represent the network as observed in satellite imagery.

#### **Mitigation:**

Such data was collated and addressed directly by the NTA at the October workshop.

### **Issue: Street View and Satellite View out of Sync**

There can be varying and conflicting views between street view and satellite views.

#### **Mitigation:**

Compare Street View image date (available to view as option in QGIS plugin) with Google Satellite image date.

### **Issue: Conflicting Signage**

Where there is signage for one type of category and road marking for another

#### **Mitigation:**

Direction from NTA: Signage is not seen as appropriate indication of category. This is rarely used as classifier. Use road markings as main classifier.

### **Issue: Junction Categorisation**

Junctions can provide a complex mix of categories and will often have a different cycle infrastructure to the majority of the segments leading to the junctions. The close inspection of each junction is very time consuming. The data provided is not usually digitised in a way that would be conducive to tagging junctions effectively.

#### **Mitigation:**

Detailed categorisation of junctions into BikeLife categories will not form part of this study. Recommendation to segment at junctions and re-classify at future date.

### **Issue: Trafficfree roads**

There is often no StreetView for traffic free roads and so categorisation is difficult.

#### **Mitigation:**

Such data was collated and addressed directly by the NTA at the October workshop.

### **Issue: Duplicate Data**

Some line segments are duplicated in the original dataset as received by Compass. An accurate calculation of total lengths is impossible.

#### **Mitigation:**

On instruction from NTA duplicates are deleted from the dataset by Compass.

### **Issue: No discernible cycle infrastructure**

Where data existed in the GDA base dataset but no infrastructure could be determined at the same location via Google Street View

#### **Mitigation:**

Such instances were collated by Compass and addressed directly by the NTA over email.

---

## 5 Future recommendations

---

Compass have collated a non-exhaustive list of data recommendations. These include:

- Continuation of the use of Boolean fields ([Table 2](#)) used throughout the process to remove unwanted data errors relating to manually entered text data.
- Conduct geometry cleaning exercises on cycle infrastructure geometries:
  - Merge/connect geometries from each data source
  - Clean/update geometries so that it reflects the real world cases
  - Create geometries to represent each side of a two way road
  - Split and join geometries at junctions
  - Add geometries for junctions where they are missing



---

## Appendix 1 - Geometry Edits

---

Geometry edits were applied to data from the NTA dataset where it overlapped with data already present in the main GDA dataset. In the cases outlined here the NTA was found to have a more comprehensive network coverage than the GDA. In these 4 cases Compass carried out some simple edits to the NTA geometries prior to geometry transfer and attribution.

The images below show exactly how each NTA segment was edited for inclusion in the main dataset. In general this process was comprised a few simple steps:

1. NTA is split at points representing the start and end points of the section overlapping with GDA
2. NTA is split where the cycle infrastructure changes from CycleLane to SignedRoute.
3. The newly split segments that do not form an overlap with are added to the main dataset and categorised according to the processes laid out in [Section 3.2](#) above.

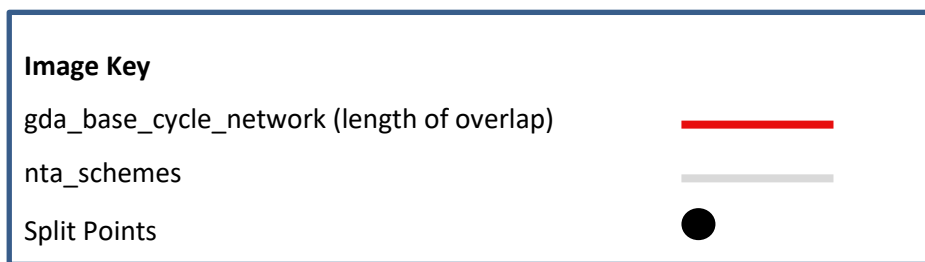


Figure 8 - MAIN ROAD TALLAGHT.



Figure 9 - GRIFFEEN ROAD / WILLSBROOK ROAD.





Figure 10 - LITTEPACE ROAD/HAZELBURY PARK



Figure 11 - LUCAN ROAD / CHAPELIZOD BYPASS