# Specification overview

**DNF release 1 product data: An overview of the specification of DNF data supplied in bulk and as change-only update, for storage and management locally by the user.**

## DNF release 1

## Responsibility for this document

Ordnance Survey DNF Business Manager is responsible for the content of this document.

## Change history

| Version | Date | Summary of change |
|---|---|---|
| 1.0 | May 2001 | First issue |

Details of the change history of this document are held on file OS 100/93/33.

## Content

This document consists of 16 pages

## Distribution

The data file for this document is archived by Documentation Services as: d00506.doc

## Approval for issue

Andrew Trigg, Ordnance Survey DNF Business Manager

## Trademarks

# Contents

# 1 Introduction

## 1.1 Purpose

This document is an overview of the DNF detailed topographic data product. You are recommended to read this document first, referring to the other DNF specification documents for further details as appropriate, see section 1.3.

## 1.2 Scope

This document:

- Provides description of data supplied to maintain the user's local data holdings **only**. Similar data may also be available online in real-time, removing the need for some users to hold DNF data locally – this is outside the scope of this document.
- Covers DNF detailed topographic data consisting of discrete vector features, only. It does not cover other DNF products that may be available, such as feature gazetteers, height model, imagery layer, and so on.
- Outlines the specification of the data delivered by DNF services. It does not give a full description of the services that will give access to this data, although some aspects are mentioned.

## 1.3 Related documents

Detailed specifications and information on various aspects of the product are given in the following documents available from the DNF pages of our web site.

| Document title | Type |
|---|---|
| *Classification and attributes of DNF features* | PDF file |
| *DNF themes* | PDF file |
| *Life cycles of DNF features* | PDF file |
| *DNF geometry and topology* | PDF file |
| *DNF data in GML* | PDF file |
| *DNF glossary* | PDF file |
| *OSDNFFeatures* | XML schema |
| *OSGeometryTopology* | XML schema |
| *OSQueryResult* | XML schema |
| *OSSimpleTypes* | XML schema |

# 2 Product overview

Detailed digital data derived from large-scale topographic surveys is part of the Ordnance Survey Digital National Framework (DNF). This is the first topographic data product to be released in the DNF programme. The DNF provides a topographic referencing framework for Great Britain.

The purpose of the data is to support a wide range of applications dependent either directly or indirectly on geographical data, including:

- geographical analysis;
- geographical referencing;
- feature management and feature attribution (data association); and
- cartographic representation.

National coverage of this product contains approximately half a billion features. DNF features are representations of many types of topographic real-world objects in Great Britain including buildings, roads, track, paths, railways, rivers, lakes, ponds, and structures such as oil storage tanks and pylons. The data also includes non-topographic features such as administrative and electoral boundaries. These objects are represented digitally as point, line, area, symbol and text features, each with attributes that give more information about the feature.

DNF data is designed for use as an intelligent digital map within GIS and database systems. In DNF data there is a closer correspondence between data features and discrete real-world objects than has been the case in previous generations of Ordnance Survey products.

The life cycle of each feature is matched, as far as is possible, to that of the real-world object(s) that it represents. The main exception to this principle is the life cycles of line features, which are constrained by topological structuring rules and so cannot always follow the life cycle of the real-world objects they represent.

Every DNF feature has a topographic identifier (TOID). This is a 16-digit integer that uniquely identifies that feature. Each feature also has a version number which is incremented each time there is change of any kind to the feature in the Ordnance Survey database.

Each DNF feature has several descriptive attributes, including one or more descriptive group, and optional descriptive terms. These attribute values are reflected in the 5-digit feature code of the feature.

DNF data also includes inferred links, which are non-topographic line features inserted to improve the subdivision of area features, in situations where physical boundaries do not serve this purpose. Inferred links represent inferences about the real world based only on the topographic map information, not on any other sources of fact. Inferred links are normally used to:

- Divide roads into sections at junctions.
- Separate individual garden plots in residential areas where no dividing fence, hedge or wall has been captured in the topographic data.
- Divide roads from car parks and hardstanding areas.
- Close fields which have gaps in their boundaries.
- Define the edge of the extent of a named area (for example, the edge of a playing field).

The insertion of inferred links for these purposes is not complete in the first-release data. A consequence of this is that DNF area feature boundaries do not always follow apparent physical property boundaries or other land parcel boundaries.

DNF data provides a seamless representation of the whole of Great Britain. All DNF products and services are based on this seamless data concept. Because there is no ‘map tile’ or similar data unit, the basic unit of data in the DNF is the feature.

Features are grouped into themes such as Buildings, Roads tracks and paths, and Land, to enable easier data selection. This provides increased flexibility for customers who require common subsets of features only in their application. A single feature can be a member of any number of themes. In data supply, each feature is included only once, even if it is a member of more than one requested theme. A line feature which bounds area features is included in the theme(s) of those area features, in addition to any themes it belongs to based on its own classification.

Representative points (also known as centroid points, seed points and area label points) for each area feature are **not** included in DNF data. DNF data does **not** currently include complex features (features made up of other features). Therefore there are no higher-level or hierarchical data structures that group together, for example, the various features that make up a certain named road, or the features which belong to a named object such as a hospital or residential district.

This data is designed to be kept up-to-date via our change-only update service. The update supply provides only those features that have changed. This makes the maintenance of local data holdings easier, because change-only update will usually constitute no more than a few percent of the total data volume.

Both initial data supply and change-only update are provided as text files in GML format. See the DNF specification document *DNF data in GML* for full information.

# 3 Representing the real world

The Digital National Framework detailed data product is a digital representation of real-world topography and associated information in the form of vector features with attributes. The complexity of the real world is simplified using rules of aggregation, selection and generalisation with a hierarchic classification imposed. A distinction is made between man-made and natural objects.

The digital model equivalent of a real-world object is a feature. We use the term feature to mean the model representation of a real-world object, never the real-world object itself. Where possible and desirable, each DNF feature represents a discrete and identifiable real-world object. In some situations, arbitrary division of real-world objects is desirable – for instance, breaking large road networks into more manageable units. In other cases, several real-world objects are generalised as a single feature.

A DNF feature may be:

- a topographic point feature representing a small object, such as a post or a letter box.
- a topographic line feature representing the centre line of a linear object, such as a fence or hedge.
- a topographic line feature representing the boundary of an area object, such as the outside face of a building.
- a topographic line feature representing the inferred division between two objects, such as an internal house division.
- a topographic line feature representing the shape of the terrain, such as the top or bottom of a slope.
- a boundary line feature representing an administrative boundary, such as a district boundary.
- an area feature representing an object bounded by a closed chain of line features, such as a building, road section, field, or pond.
- a cartographic text feature containing a label giving additional information, such as the distinctive name of a topographic feature.
- a cartographic symbol feature giving additional information, such as the direction of water flow.

# 4 Classification of features

DNF features are classified by their feature type, descriptive group, and descriptive term. For each combination of these in use, a 5-digit feature code is allocated. Therefore, feature codes provide an alternative to the feature description attributes, but do not add any information to them.

DNF themes are **not** part of the DNF feature classification. Themes exist to allow users to easily select a fixed set of features which are often required as a group. New themes can be created to facilitate data selection by particular groups of users. DNF Themes are defined in the DNF specification document *DNF themes*.

For full details of the information presented in this section, see the DNF specification document *Classification and attributes of DNF features*.

## 4.1 Feature type

There are seven types of DNF feature. Each feature type has an associated set of feature attributes. The feature types and their attribute sets are:

- Topographic area — All DNF area features.
- Topographic line — All DNF line features, except administrative boundary features. This includes features representing objects which are not physically present, such as inferred area feature boundaries.
- Boundary line — Administrative boundary line features.
- Topographic point — All DNF point features. These include features representing objects which are not physically present, such as spot heights.
- Cartographic symbol — All cartographic symbol features.
- Cartographic text — All cartographic text features.
- Departed feature — In change-only update data supply only, this type is used to indicate features which have been deleted or have otherwise departed from the requested data since a given date.

Most features with a topographic attribute set represent objects which are physically present in the topography. However, these attribute sets are also used to represent some types of objects which may not be physically apparent, such as inferred polygon closing links and spot heights.

## 4.2 Feature description attributes

DNF topographic features have at least one descriptive group attribute, such as building, terrain and height or general feature. Some features also have one or more descriptive term attributes, which further classify the feature. For instance, a feature with descriptive group Structure could have descriptive term Pylon.

Other descriptive attributes include: make, which indicates if the feature is natural or man-made; physical level, which indicates whether the feature is at, above, or below ground level; and physical presence, which indicates how the feature is present in the topography.

## 4.3 Feature codes

Each DNF feature is assigned a five-digit numeric feature code based on (i) its feature type; (ii) its highest priority descriptive group attribute, and (iii) its descriptive term attribute(s), if any. For instance, all topographic line features which have descriptive group Rail and descriptive term Buffer are assigned feature code 10160.

A priority order of descriptive groups is defined, which gives priority to one group in the case of features with more than one descriptive group. The highest priority descriptive group of a feature is used to assign the feature code.

# 5 Geometry and topology of features

The following is a brief summary of the DNF geometry and topology model. For full details see the DNF specification document *DNF geometry and topology*.

## 5.1 Geometric and topological attributes

The geometry of a DNF feature is one of its attributes. Each feature type has an attribute set which includes a geometric attribute of a particular type.

The geometric attribute of a topographic point, cartographic text or cartographic symbol feature is a point, which is a horizontal location specified as a pair of coordinates in the British National Grid spatial reference system. The geometric attribute of a topographic line or boundary line feature is a polyline, which is an ordered set of points. The geometric attribute of a topographic area feature is a polygon, which is a region defined by an outer boundary and zero or more inner boundaries ('holes' in the polygon).

The data is available in two forms, which differ only in the way the geometry of polygons is encoded. The two forms are called topological polygon data and independent polygon data.

In topological polygon data, each (inner or outer) boundary of a polygon is encoded as an ordered set of references by TOID to line features. The boundary is a closed ring formed from the polyline attributes of those line features.

In independent polygon data, each (inner or outer) boundary of a polygon is explicitly stated as a list of coordinate pairs. Hence each polygon is independent because it does not rely on references to other features for its coordinate geometry. However, this leads to duplication of coordinate information where features have common boundaries.

## 5.2 Rendering of text and symbol features

Each cartographic text feature has the following rendering attributes:

- an Anchor Point – the coordinated point to which the text is attached. This is located close to the topographic feature which the text names or describes;
- an Anchor Position to indicate its alignment relative to the Anchor Point;
- an optional orientation angle; and
- a font style and text size.

Cartographic symbol features have a point attribute and an orientation attribute.

# 6 Feature attributes

The term feature attributes means all the data items within the feature. For instance, the geometry of the feature is one attribute of that feature. *NOTE: The term attribute is also used in XML data formats such as GML, with a quite different meaning. Most (but not all) DNF feature attributes are encoded as GML properties. In DNF specification documents the terms 'feature attributes' and 'XML attributes' are used whenever the two meanings might be confused.*

The principal DNF feature attributes are as follows. Not all these attributes are present in all attribute sets, and this list is not complete. See the DNF specification document *Classification and attributes of DNF features*.

- **TOID** – a unique 16-digit identifier which is the 'handle' of the feature whether topographic (eg a building) or non-topographic (eg an administrative boundary).
- **Version** – an integer which is updated each time there is any change to the feature; this allows users to be sure which version of a feature they are dealing with.
- **Version date** – the date on which this version of the feature became current. A feature does not indicate if it is no longer the current version, because superseded feature versions are not stored in the OS databases.
- **Feature geometry attribute(s)** – an attribute of type point, polyline, or polygon, depending on the feature type.
- **Positional accuracy** – the positional accuracy of the feature geometry relative to the spatial reference system, stated at the 95% confidence level.
- **Descriptive group** – one or more of 21 feature classification categories.
- **Descriptive term** (optional) – gives further information on the classification of the feature. A feature may have zero, one, or more than one descriptive term.
- **Physical level** – the level at which the feature lies: underground, obscured below normal level, at normal level, or overhead.
- **Physical presence** (optional) – can indicate a variety of states related to the physical presence of the feature.
- **Make** (optional) - indicates whether the land surface is natural or manmade;
- **Feature code** – a 5-digit integer which classifies the feature. It is derived from the other attributes of the feature and does not add any additional information.
- **Theme -** the one or more themes to which this feature belongs.
- **Change history**– a list of successive changes to the feature. Each entry in the list consists of a reason for change and a change date.
- **Reference to feature** – a TOID reference in selected cartographic symbol features to the relevant topographic feature. Also used in some heighted topographic features to refer to unheighted features.
- **Text rendering** – information on the cartographic depiction of text features.

# 7 Feature life cycles

The rules that set out the circumstances in which features are retained, created or deleted during data maintenance by Ordnance Survey, are defined in the DNF Specification document *Life cycles of DNF features*.

## 7.1 Feature TOIDs and version numbers

The identity of a topographic feature is defined by its TOID. It is the TOID which preserves the identify of a feature through successive changes. We speak of the life cycle of features, not of the life cycles of TOIDs.

Every DNF feature is allocated a TOID. Once allocated to a feature a TOID will never be re-allocated to another feature. This TOID will remain with the feature throughout its life. When a feature is deleted, we retain a record that its TOID has been used and the feature no longer exists, but we do not store attributes of the feature.

The feature version number is incremented each time any attribute of the feature changes. The version date attribute states when this version became the current version of the feature. Therefore a reference to both TOID and version number identifies a feature with no ambiguity. Successive changes to a feature are recorded in the change history attribute.

## 7.2 Feature life cycle rules

The general principle is that a DNF topographic feature (with its unchanging TOID) will be retained for as long as possible while the real-world object it represents can still be reasonably considered to exist.

For instance, if a building is extended, the area features representing the changed building and surrounding land area are retained with their unchanging TOIDs, and their version numbers are incremented. If most of a large residential garden is sold off for development, the existing feature is retained to represent the new smaller garden. However, if a field is divided into three fields of approximately equal size, the original feature is deleted because no one of the new fields can claim to be the obvious successor to the original. All three fields are new features with new TOIDs.

The principle of matching the feature life cycle to real-world object life cycles applies only to area and point features, not to line features. Because the geometry of each DNF line feature is a single link between two topological nodes, there cannot be a one-to-one correspondence between DNF line features and linear real-world objects. If a new line feature is added which intersects another, the pre-existing feature is split into two by the topological structuring rules, although there has been no change to the real-world object(s) (or part of an object) it represents. For this reason, it is rarely advisable to link user data items to DNF line features by TOID references.

# 8 Themes

A DNF theme is a set of features which are can be collectively selected for supply by users. A feature may be a member of any number of themes. All features belong to at least one theme. Each theme is defined by rules based on feature attributes, most commonly the descriptive group attribute. A full description of each theme, and the rules for assigning DNF features to themes, are given in the specification document *DNF themes*.

The following table gives an overview of the current DNF themes. *NOTE: The example column gives real-world objects that may be represented in that theme*

| Theme | Description | Example |
|---|---|---|
| **Buildings** | Roofed, and usually walled, constructions. | Barns, private houses, factories, schools. |
| **Land** | Features defining and/or describing natural and manmade parcels of land not specifically within other themes. | Fields, forested areas, playing fields, cliffs, quarries, sand dunes. |
| **Roads, tracks and paths** | Features related to transport by vehicles, cycles or on foot. | Metalled roads (public and non-public), cycleways, footpaths. |
| **Heritage** | All heritage-related features. | Burial mounds, earthworks, ruins and battlefields. |
| **Height** | Non-real-world features that describe the elevation of particular locations. | Spot heights and bench marks, areas of cliff and slope. |
| **Rail** | Features related to transport by railway or tramway. | Track alignments, switch points, permanent way. |
| **Structures** | Man-made constructions, usually unroofed. | Chimneys, storage tanks, telephone boxes, electricity transmission lines, masts. |
| **Water** | Features that contain, delimit or are related to real-world objects containing water. | Rivers, ponds, reservoirs, canals, wells. |
| **Administrative Boundaries** | The limits of responsibility and representation for electoral and administrative purposes. | Parish boundaries, European parliament constituency boundaries. |

# 9 Coordinate reference system

The coordinate reference system used is the British National Grid, as in all other Ordnance Survey products. The National Grid transformation OSTN97 is available free of charge to convert precisely between this system of eastings and northings coordinates and the precise latitude and longitude coordinate system ETRS89 which is used by the Global Positioning System (GPS) – see the National GPS Network web site, www.gps.gov.uk, for details.

It is recommended that all precise surveys in Great Britain are carried out using the National GPS Network to produce ETRS89 coordinates, which can be converted using the OSTN97 transformation to British National Grid coordinates compatible with this DNF data product. Via these free OS services, DNF is fully compatible with precise GPS positioning.

# 10 File-based supply of DNF data

This section outlines the principles of DNF product supply.

An online service provided by Ordnance Survey allows customers to browse graphical displays of the data to which they have access, and accepts requests for data.

The main parameters chosen by the user when requesting data are:

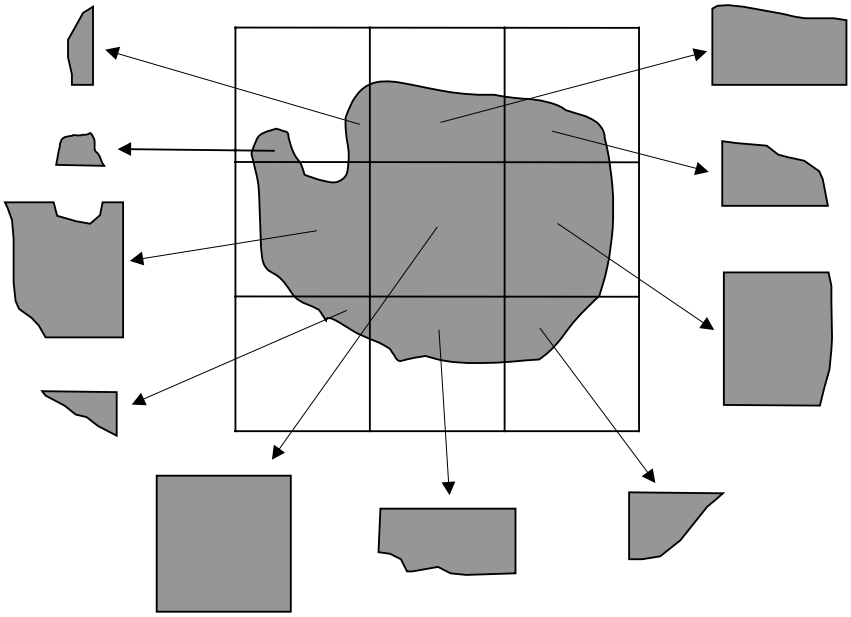
- The area of supply, specified as an arbitrary polygon.
- The data themes required. Features which are members of multiple selected themes are supplied only once.
- The delivery mechanism. Usually, initial supply will be on DVD or CD, and change-only update will be delivered via Internet FTP.
- The data format. There are two options: topological polygon GML data, or independent polygon GML data.
- The change-since date. This is required for change-only update requests only.

## 10.1 Initial supply of data

Because DNF data is seamless, GML files containing initial supply of large areas could be very large – several hundreds of gigabytes in uncompressed form. In order to restrict individual uncompressed GML files to a reasonable size, initial data supplies are optionally divided into chunks, each of which is supplied in a separate GML file.

This is achieved by overlaying a square grid on the data selection polygon, such that the grid is optimally aligned with the data selection polygon. The part of the data supply area which falls in each grid square is supplied in a separate GML file. Where a grid square lies entirely within the data selection polygon, the GML chunk covers the entire square. Where the data selection polygon intersects the grid square, one or more GML chunks are created, bounded partly by the grid square edges and partly by the data selection polygon. This is illustrated in the figure below.

**Figure:** *The division of a data supply polygon (large grey region) into chunks by overlaying a 'chunking grid'. In this case ten chunks are created. The central chunk is a complete grid square; the others are partly bounded by the data selection polygon. The upper left square shows the effect when the data selection polygon crosses a grid square twice – two or more separate chunks are created.*



In a GML data chunk, features are not clipped to the nominal supply boundary. Therefore, each initial supply chunk includes all current features which overlap the nominal chunk boundary. In topological polygon data, line features which are outside the nominal boundary, but which are part of the geometry of area features inside the supply polygon, are also supplied. This ensures that the full geometry of all area features which overlap the chunk boundary is available in each chunk.

A consequence of this is that some features are supplied in more than one chunk. Systems reading DNF data must recognise these duplicated features.

Chunk boundaries are imposed purely for the purpose of dividing a large supply into pieces of manageable size in a geographically meaningful way. Chunks cannot be treated as persistent units for data supply or management – the origin of the chunking grid may be different in a future supply.

All GML files are compressed using the 'gzip' standard.

## 10.2 Change-only update

In a change-only update request, the user specifies a change-since date. All features which have changed, been created, or departed since 00:00 hours on that day are supplied. By default, change-only update GML files are supplied by Internet FTP, not on CD or DVD.

New features and modified features are supplied together in the change-only update file. For departed features, only the TOID is given in the GML file.

Because change-only update refers to the start of a given day, not to the user's last transaction, the user system must be able to handle repeated change information. That is, it must recognise when a feature being presented as a new, modified or departed feature in a

change-only update process is already known to the user system.

Users can at any time request a change-only update with any change-since date. This gives the user considerable flexibility in managing the update process and rectifying mistakes or corruptions in a series of regular change-only updates. For instance, if a corruption is detected but the exact date or extent of the corruption is not known, it can be rectified by requesting a change-only update covering the whole time period and area that might be involved. It is not necessary to reload a sequence of archived update files.

**Figure:** *The principle of applying change-only update to a local holding of DNF data.*