Ordnance Survey®

www.ordnancesurvey.co.uk

Code-Point with polygons™

user guide

Awarded for excellence

Preface

This user guide contains all the information you need to make effective use of Code-Point with polygons™. It is designed to help you understand the information contained in the data, as well as providing detailed technical information and the data format specification.

This user guide has been checked and validated before issue and every endeavour made to ensure that the contents are accurate. If you find an error or omission, or otherwise wish to make a suggestion as to how this user guide can be improved, please contact us at the address shown under Contact details.

The contents of this user guide will be updated by the release of replacement chapters.

Contact details

Sales Information will be pleased to deal with your enquiries:

Sales Information Helpdesk
Phone: 08456 050 505
Fax: 023 8079 22615
Email: enquiries@ordsvy.gov.uk

or write to:

Sales Information
Ordnance Survey
Romsey Road
SOUTHAMPTON
United Kingdom
SO16 4GU

Visit the Ordnance Survey web site at: www.ordnancesurvey.co.uk

Product performance

If you have any problems with or identify any errors in the data, or this user guide, please contact us at the address shown under *Contact details*.

Liability

A full description of the terms and conditions of supply and use of Code-Point with polygonsis detailed in the signed *customer contract* held by you or your organisation.

Ordnance Survey® makes every effort to ensure that data supplied is free from errors and omissions. We will remedy, as soon as reasonably practicable, errors and omissions that the customer notifies to us in writing. It is the customer's responsibility to ensure that data ordered is suitable for the intended purpose. We will not be liable to the customer or any other party for any loss, damage, inconvenience or expense resulting from the use of, or reliance upon, the data.

Delivery of Code-Point with Polygons

You are advised to copy the supplied data to a back-up medium.

Written notification of any deficiency in the data or damage to the goods must be given to us within 28 days of receipt of your Code-Point with polygons data.

Data copyright

All Ordnance Survey digital map data are Crown copyright. Full details of the conditions under which our digital map data may be processed and/or manipulated or copied by a customer – whether or not for use on PCs or workstations or for making hard copies – are contained in Ordnance Survey leaflet: *Copyright 1 – Business use of our digital and conventional mapping*. This is available from Copyright Licensing, please see overleaf. You should check the terms and conditions with us before using the data. It is also the responsibility of the holder of the digital map data to ensure that any plotted or printed output contains the required copyright acknowledgement.

User guide copyright

Trademarks

Contents

Chapter 1 Introduction

Using this user guide

This is a temporary user guide for what is a new product that is still, to a certain degree, subject to iterative development. Some of the details of the product specification may be redefined in the light of customer reaction, resulting in the modification of this guide. In view of its interim nature, this guide concentrates on the polygons that are provided in the product. For information regarding the Code-Point data, reference should be made to the previously published *Code-Point® user guide*.

Code-Point with polygons™ is a dataset that provides:

- two sets of 121 Code-Point text files that gives Ordnance Survey National Grid (Ngref) coordinates for the postcode units on the United Kingdom, supplied in NTF and CSV formats;
- a set of 120 ASCII transfer files that, when viewed using appropriate geographical information system (GIS) software, provide a set of boundaries for the postcode units in Great Britain. These boundaries are nested within Geoplan® postcode sector boundaries and constrained to geographical features such as roads, railways and rivers; and
- further associated data, metadata and information files.

Chapter 2 Overview of Code-Point with polygons

Data overview

Basic principles

The Code-Point polygons are derived from ADDRESS-POINT, the Ordnance Survey dataset that provides Ngref coordinates for each postal delivery address in Great Britain. The *Thiessen* process tessellates these points, then the address boundaries inside each postcode are dissolved away, leaving boundaries for the postcode units.

The polygons are then additionally refined by clipping them to major roads, railways, rivers, etc and to the postcode sector boundaries produced by Royal Mail in association with Geoplan, a division of Yellow Marketing Information Ltd.

There is actually no such thing as a *correct* unit postcode boundary, a postcode unit only being the delivery point, or collection of delivery points, that constitute the unit postcode. The boundary is therefore a notional one, the position of which is arbitrary. What has been created, however, is a set of boundaries that follow a consistent logic and portray the notional footprint of each postcode unit, enclosing every delivery address for which positional data of sufficient quality is available and following the major physical features that could reasonably be regarded as part of the postcode boundary.

The polygons are thus of particularly high quality in that every address, provided its match in ADDRESS-POINT is of sufficient status, will be shown in the correct postcode polygon. Equally, the polygons will only contain the addresses appropriate to the postcode represented.

This high quality standard allows the polygons to be used for a wider range of applications than the products previously available. This will include analysis of geographically based information or statistics by postcode and the pictorial display of information that has been analysed or sorted by postcode.

The polygons are supplied with the Code-Point data set, which provides a complementary set of point references for each postcode. From the second release (version 2001.2.0), the two datasets will both be created from the same edition of ADDRESS-POINT, ensuring that their data is synchronous. It is also worth noting that other products, and data attached to the products, derived from the Gridlink® database of the same release date, will also be synchronous.
Note: The first release of Code-Point with polygons will **not** incorporate synchronised Code-Point and polygon datasets.

What you need to use Code-Point with polygons

Hardware

To operate efficiently, Code-Point with polygons requires only a reasonably modern and powerful PC or similar, ideally with a Pentium® II /200 processor (or similar) and a minimum of 32 Mb RAM. If national cover of the product is to be used, a hard disk of at least 1 Gb capacity will be required.

Software

Most proprietary GIS packages will suffice, for example, MapInfo, ArcView or Autodesk® products.

Code-Point with polygons supply options

There is one supply extent provided within the product. That is *national cover*. That is defined as Great Britain for the polygon set and United Kingdom for the Code-Point point data. Although the polygon coverage does not include Northern Ireland, the point data for that province (BT postcode area) has been left in the data for simplicity of production.

Transfer format(s)

Code-Point with polygons is available in the following formats, the preferred choice of which will be influenced by the software used:

- ArcInfo Import/Export format (E00)
- Mapinfo Interchange format MIF/MID

Media, data compression

Code-Point with polygons is supplied on CD-ROM. The Code-Point and polygon files (only) contain substantial amounts of cartographic and attribute information, which in both E00 and MIF/MID formats necessitate that file compression be used to distribute the files.

File sizes

The following table outlines the total file sizes for both formats in compressed or uncompressed states.

Compressed	Uncompressed	Total including associated and meta-data files
Code-Point (CSV & NTF total)	Code-Point (CSV & NTF total)	
Polygons in MID/MIF format 582 Mb	Polygons in MID/MIF format 2.21Gb	
Polygons in E00 format 1.1 Gb	Polygons in E00 format 4.15 Gb	

Update

Code-Point with polygons is still undergoing a degree of specification development and refinement. It is anticipated that the first release in June 2001 will be followed by an update in September 2001, which will be followed by updates at six-monthly intervals thereafter.

Code-Point with polygons version numbering

Each edition of Code-Point with polygons will have a version number showing the year, the release number for that year and the version (if appropriate) of that release, for example:

2001.1.0 is June 2001, first release.

2001.2.0 is September 2001, second release.

Within the Code-Point with polygons product, the Code-Point data and the polygons will have their own version numbering, based on the same principles as above, for example, Code-Point with polygons vs 2001.1.0 will contain a set of Code-Point vs 2000.2.0. and a set of polygons vs 2001.1.0. To maximise compatibility, the first release of polygon data has been accompanied by the issue of Code-Point data that is closest in content.

It is appreciated that the version numbering may appear somewhat odd during 2001, a result of the time it has taken to develop the product. It is expected that in 2002, the version numbering will be more logical and will set the pattern for future releases:

Code-Point with polygons release date and version number	Code-Point version	Polygon data version
June 2001 – vs 2001.1.0	2000.2.0	2001.1.0
September 2001 – vs 2001.2.0	2001.3.0	2001.2.0
March 2002 – vs 2002.1.0	2002.1.0	2002.1.0
September 2002 – vs 2002.2.0	2002.3.0	2002.2.0

Product Development

Code-Point with polygons is a new product, designed to meet the requirements of users as expressed through market consultation exercises. It is envisaged that the product will evolve to some extent as users become familiar with it and refine their views.

Code-Point with polygons product structure

On the product CD-ROM(s), the root directory contains two folders: *Info* and *Data*.

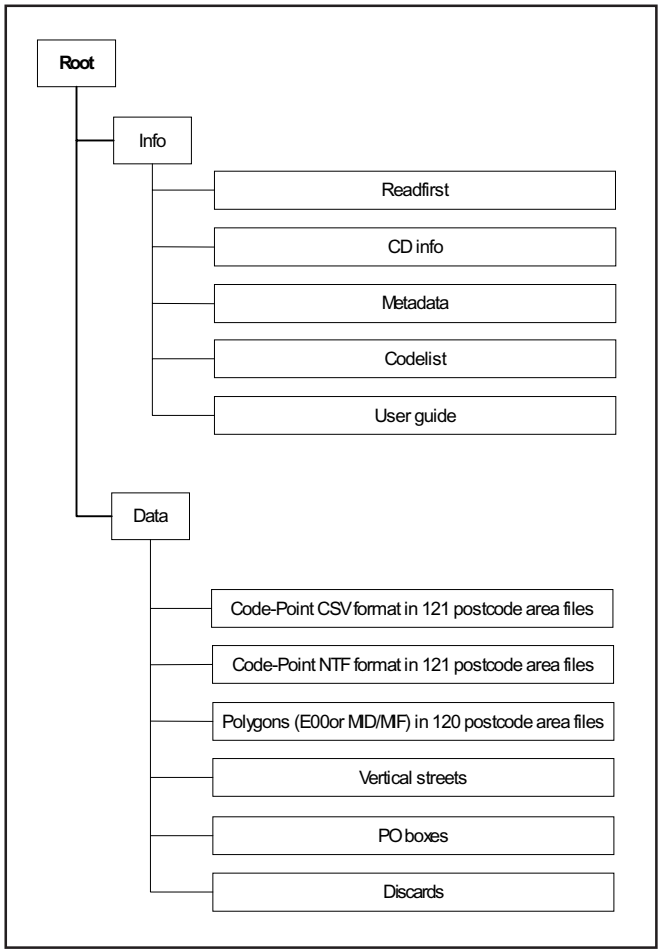
The *Info* folder contains the following files:

- Readfirst – a file summarising copyright and licensing information that must be read and understood before any data files are opened.
- CD Info – a file summarising the content and file structure of the CD-ROM.
- Metadata – lists the number of postcode units, sectors and districts in each postcode area and the date of the most recent version of PAF that has been incorporated into the data
- Codelist – a list of the ONS county, district and ward codes and their full text equivalents.
- User guide – this user guide.

The Data folder contains the following sub-folders:

- *Code-Point* – containing Code-Point point data in two sets of 121 postcode area files (both CSV and NTF).
- *Polygons* – containing polygon data in 120 postcode area files (either MIF/MID or E00).
- *Vertical Streets* – a look-up table of vertical street reference codes and the postcodes contained in them.
- *PO_boxes* – a look-up list of the postcodes that have not been included in the polygon creation process because they are PO boxes.
- *Discards* – a look-up list of the postcodes that have not been included in the polygon creation process because there are no ADDRESS-POINT records of sufficient positional quality PQ classification.

Structure of the Code-Point with polygons CD-ROM



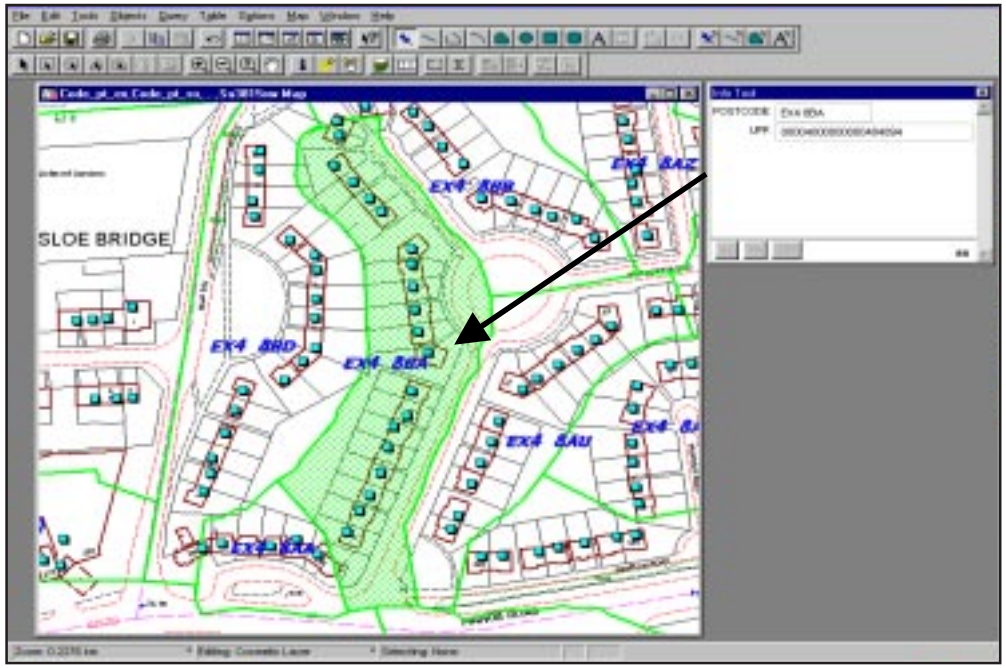
Text

The text files in the *Info* folder files are in text (.txt) format with the exception of the user guide, which is an Adobe PDF file.

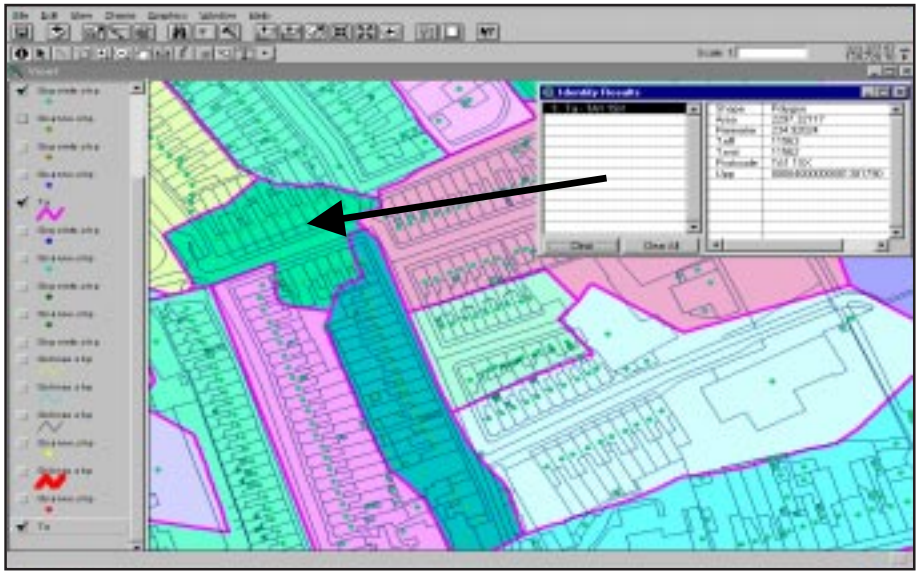
Using Code-Point with polygons

The uses of Code-Point georeferenced postcode data are covered in the user guide specifically issued for use with that product. This user guide will focus on the use of the unit postcode polygons.

The polygons provide a boundary around the addresses within unit postcodes and so can be used to relate any occurrence or point within that boundary to postcode concerned. In this way, analysis of data on a postcode basis can be achieved, even when the occurrence is not at a postcoded address.



Where data is already analysed by postcode, a pictorial representation, for example, using a colour scheme, can be obtained that will differentiate between postcodes, or groups of postcodes:



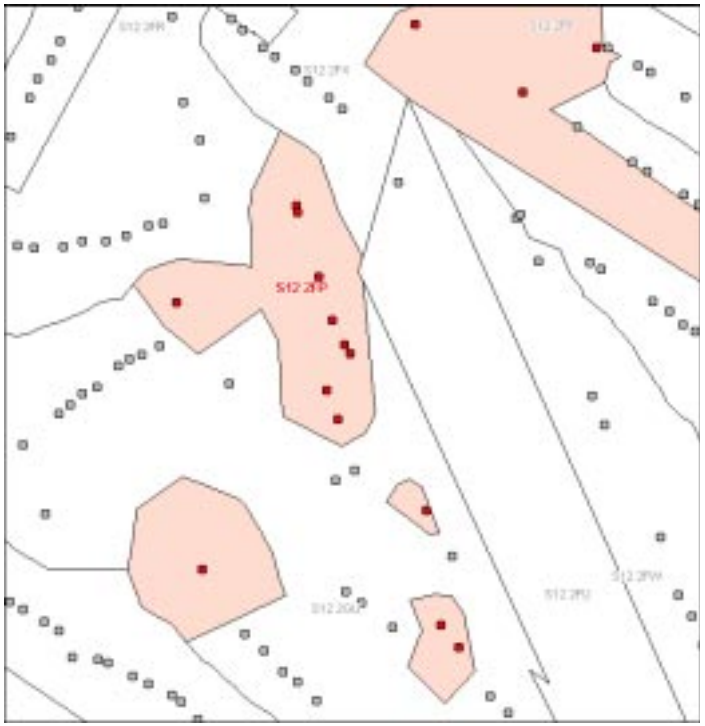
Chapter 3 Code-Point with polygons explained

The standard Code-Point product is described in detail in the *Code-Point user guide*. This chapter will concentrate on the polygons provided within the product.

Description of the polygon set

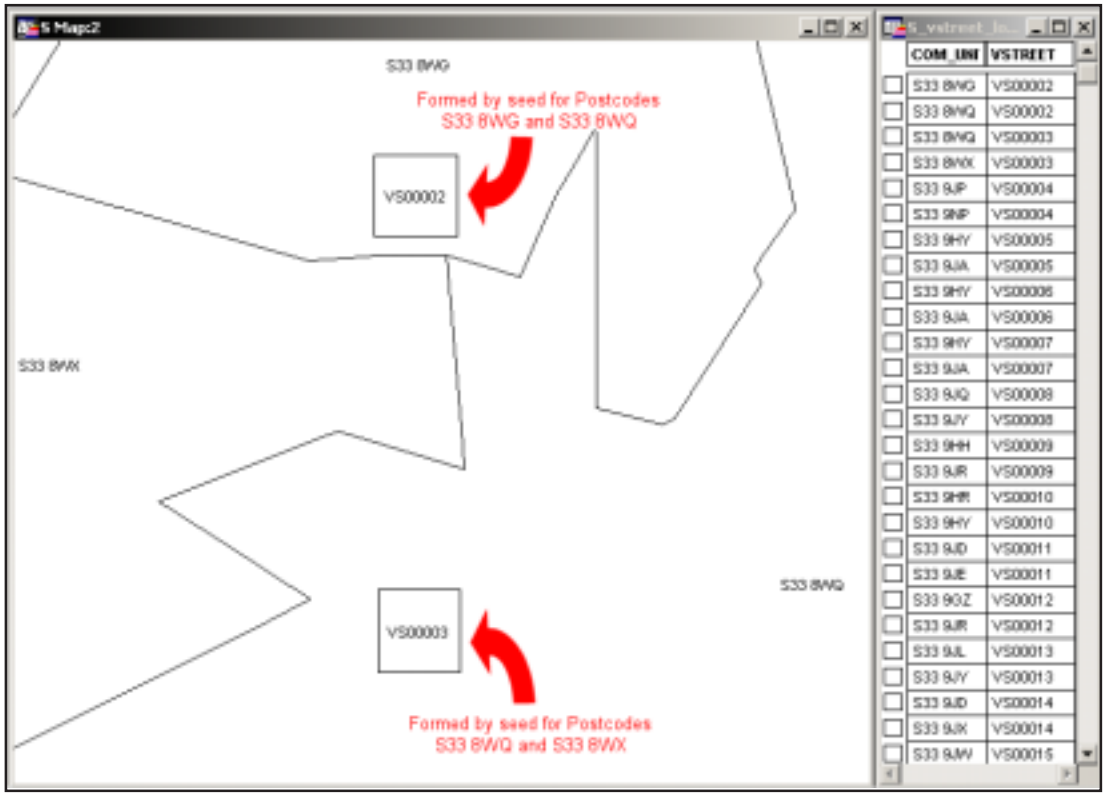
- The set is produced by the tessellation of ADDRESS-POINT coordinates for individual Royal Mail delivery addresses.
- Where the postcode unit boundaries form a section of postcode sector boundaries, the unit postcode boundary nested within Royal Mail's postcode sector boundary maps.
- Only addresses, in ADDRESS-POINT, having a positional quality value of 2 or 3 are used to create the polygons file.
- PO Boxes are not included, but their postcodes are supplied in a look-up file in the *Data* folder on the CD-ROM.
- A supplementary file is also provided, in which postcodes for which there are only addresses in ADDRESS-POINT with a PQ value of 1, 4 or 5 are represented and given a point grid reference.

Due to the nature of postcode geography, the polygons representing some unit postcodes are unavoidably split. Every effort has been made to ensure the absolute minimum of postcodes is represented by multiple polygons. Furthermore these *split* polygons representing a single postcode remain a single object with one set of attributes. When this occurs, any issues arising will depend upon the mapping package used and the nature of its use. Typically, users may see only one of the group polygons labelled, however it is worth checking the labelling options in your mapping package. If however the user was to query the polygon(s) or display associated data (using the postcode as a link field) all the split parts of a polygon would correctly represent the associated attributes. An example of such a polygon is shown on the next page, with *seed* ADDRESS-POINT locations for that postcode shown in red:



- Wherever it is possible without moving the boundary to the wrong side of an address, the unit postcode boundaries are constrained to motorways, railways, rivers and canals, A roads and B roads, in that order of priority.
- Each polygon is assigned a unique identifier. The identifier will be a 16-digit series. These identifiers are not re-used should a polygon be deleted.
- An attribute is attached to each polygon, which defines the number of addresses used in the creation of the polygon (from version 2001.2.0).
- The polygon dataset contains non-overlapping polygon coverage of Great Britain, constrained by the extent-of-realm coastline from Ordnance Survey's Boundary-Line® data. Should any addresses fall outside the coastline they will be included in the *discards* file.
- There are no polygons or areas enclosed by polygons without a postcode allocated, or overlaps in the polygons.
- The dataset is accompanied by a *readme* file that provides information regarding the currency of the data used, including the date of the version of ADDRESS-POINT (and the PAF release therein) and other Ordnance Survey datasets used, and the date of the Royal Mail sector boundaries produced in association with Geoplan that have been used.
- If the position of a sector boundary is shown approximately parallel to a physical feature such as a road or railway, at a distance of up to 50 metres, so that it was clear that Royal Mail had intended the sector boundary to follow that feature, the boundary is clipped to that feature, provided the boundary does not then fall on the wrong side of any address.
- The data is divided into 120 postcode area files, each file named with a one- or two-letter postcode area code.
- Where two or more postcodes are associated with a single building seed, a single distinctive square polygon will represent all the postcodes attached to the seed. These polygons have a special series of identifiers, all commencing with the letter *V*. A separate look-up table lists the postcodes and their 16 digit unique identifier that are represented by each special polygon. Where these distinctive polygons are crowded closely together, they are reduced in size to prevent overlaps hiding some of the polygons.

Example showing two vertical streets and, on the right, an extract from the *Vertical Streets* look-up table:



Chapter 4 Quality statement/specification

Lineage

To the Royal Mail Postal Address File (PAF) address records for Great Britain, Ordnance Survey add National Grid coordinates, mainly to 0.1 m resolution, and other information to produce the product called ADDRESS-POINT. This unique product thereby provides highly accurate positioning of all postal addresses in Great Britain.

ADDRESS-POINT is then used to create both Code-Point and the Code-Point unit postcode polygons.

Content

Code-Point with polygons contains:

- Code-Point georeferenced unit postcode data, with associated metadata such as address counts and quality indicators. Also provided are the health and administrative area codes related to each postcode. The coverage of this data is the whole of the United Kingdom and it is provided in both CSV and NTF formats.
- Also provided, in association with the Code-Point data, is a text file that provides the full text equivalents of the administrative area codes, and another that provides the numbers of unit postcodes in each postcode sector, district and area.
- Unit postcode polygons describing notional boundaries around the approximately 1.7 million unit postcodes in Great Britain. This data is supplied in either E00 or MID/MIF formats.
- Also provided, in association with the polygon data, are three CSV text files:
  - *Vertical streets* – a list of polygons, identified by a serial number that is prefixed by the letter *V*, that contain more than one postcode. This situation can occur in, for example, blocks of flats where there is more than one postcode within a single building.
  - PO Boxes – a list of the PO Box postcodes, none of which will have been used in the creation of the polygon set.
  - *discards* – a list of the postcodes for which polygons have not been included, because there is no data of sufficient quality to use in the polygon creation, or because their constituent addresses lie outside the extent-of-realm coastline.

- A text file named *Readfirst* that contains important information regarding copyright, and so on; another called *CD Info* that contains information regarding the structure and content of the CD; and a third called *user guide* that contains the current user guide for the product.

Completeness

The Code-Point product contains a point reference for every unit postcode in England, Scotland, Wales and Northern Ireland that is contained in Royal Mail's PAF product.

The polygon set contains a polygon for every postcode in England, Scotland and Wales that is contained in Royal Mail's PAF product, with the following exceptions:

- postcodes for which there is no data in ADDRESS-POINT of sufficient quality;
- postcodes for which there is no data in ADDRESS-POINT that lies within the extent-of-realm coastline;
- postcodes that relate to PO boxes; and
- postcodes that are *vertically-stacked*, that is, two or more postcodes within a single building that is represented by a single Land-Line building seed. In these situations, a single square polygon represents all the postcodes attributed to the single building seed.

Currency

Because of the dynamic nature of the source information any comprehensive, national list of postal can never practicably be 100% correct. The time taken to collect and incorporate additions, changes and deletions from the real world into PAF, then into ADDRESS-POINT and hence into Code-Point and Code-Point with polygons can only add to any lack of currency.

Having said that, PAF is issued monthly and incorporated within a week into the ADDRESS-POINT database. ADDRESS-POINT is then issued quarterly and the same data is used to create Code-Point within a further week. The polygon set is then created and released approximately one month later, from the same data. At the time of their issue, therefore, the polygons will be based on PAF data that is approximately two months old (v 2001.2.0 onwards).

Attribute accuracy

All polygons are attributed. Nothing is done to change the accuracy inherited from the PAF source data.

Logical consistency

The logic used to create both the Code-Point point and polygon data is consistent across the whole of each dataset.

Medium

Code-Point with polygons is supplied on CD-ROM, ISO 9660.

Formats

MID/MIF is the transfer format of the MapInfo Company and E00 is the transfer format of ESRI (UK) Ltd. The data conforms to the latest published specification.

Record breakdowns for the transfer of Code-Point (polygons) MIF/MID

Field Name	Type	Width	Description	Example
POSTCODE	Character	8	Full Postcode from Code-Point	HG1 1BA
UPP	Character	20	Unique Polygon Identifier	00004000000000590783

Record breakdowns for the transfer of Code-Point (polygons) in E00

Field Name	Type	Width	Description	Example
POSTCODE	Character	8	Full Postcode from Code-Point	HG1 1BA
UPP	Character	20	Unique Polygon Identifier	00004000000000590783

Note: E00 format automatically adds fields to show the values for: Surface Area, Perimeter and Region Identifiers (internal). For example, the E00 file for Postcode HG will also feature the following fields: AREA, PERIMETER, HG_REGION_ and HG_REGION_ID