

Guidelines for BS 7666:2006

Section 5. Creating a street gazetteer

Version 1

May 2007

This Section provides detailed guidance on how to create a street gazetteer. It explains aspects of BS 7666 Part 1: Specification for a street gazetteer, and should be read in conjunction with that document. However it does not cover Public Rights of Way. It is aimed at those compiling such gazetteers, particularly in local authorities. The guidelines are general and not aimed at any particular implementation such as the National Street Gazetteer, for which specific guidelines will apply. A Glossary of Terms, list of abbreviations and references, and an explanation of the UML data modelling convention used in the Standard are given in Section 1 of these Guidelines.

Contents

	<u>Page</u>
Preface	2
1. Introduction	3
2. Details of changes from previous edition and implications for implementation	3
3. Explanations of requirements	5
4. Recommended approach	10
5. Common misunderstandings and errors	17
6. Conformance issues	18

Preface

These Guidelines are intended for use with BS 7666: 2006 *Spatial datasets for geographical referencing*. They complement the Standard with more detailed explanation of the content and a general approach to creation of gazetteers. They are not specific to any particular implementation, for which it is expected that specific data specifications and capture and maintenance rules will be produced.

The Guidelines are aimed at:

- gazetteer owners – those with ultimate responsibility for the gazetteer;
- gazetteer custodians – those responsible for the creation, maintenance and quality of gazetteers;
- suppliers of gazetteer software;
- those developing and implementing gazetteer systems
- providers of gazetteer data;
- others who are responsible for aspects of the quality management of gazetteers.

The Guidelines are currently in six Sections:

1. Introduction to BS 7666;
2. How to create a gazetteer of a new type of geographic object;
3. Quality assessment and reporting;
4. How to create a national gazetteer;
5. How to create a street gazetteer (this section);
6. How to create a land and property gazetteer.

No guidelines for public rights of way which form an informative annex to Part 1 of the Standard, or for Part 5 Gazetteer of delivery points are planned at present.

These Guidelines have been produced under the auspices of BSI IST/36 geographic information who are responsible for BS 7666. They were written by Rob Walker and Les Rackham working under the guidance of a Steering Group comprising representatives of major stakeholders in the Standard. The work is sponsored by the Department for Communities and Local Government (DCLG), Ordnance Survey, Office for National Statistics and Mayrise Ltd.

This publication may be reproduced free of charge in any format or medium provided that it is reproduced accurately and not used in any misleading context or in a derogatory manner. The material must be acknowledged and the publication cited when being reproduced as part of another publication or service.

1. Introduction

A street gazetteer is a list of streets in an area, together with information about their location. There is no simple definition of a street, and it is defined differently in different contexts. In BS 7666 it is defined as a “way or thoroughfare providing a right of passage on foot, by cycle or by motor vehicle”. This is intended to be all-encompassing, allowing flexibility in what may be included in the gazetteer. Essentially the gazetteer contains records of street references. These references may be any of the following:

- a designated street name – provided by the Street Naming and Numbering Authority;
- a street description – usually used for streets that do not have names, such as country lanes and service roads;
- a street number – a Department for Transport or other Highway Authority route number¹;
- an unofficial street name – likely to be a locally used name that has not been adopted by the Street Naming and Numbering Authority.

2. Details of changes from previous edition and implications for implementation

The main changes to Part 1 from the previous (2000) edition of the Standard are as follows:

- Addition of a requirement to provide metadata:
 - This is important, especially the scope statement, as it aids understanding of the dataset. The requirements are described in **3.4** of this Section of the Guidelines.
- Addition of a facility for recording descriptive identifiers in multiple languages:
 - This will aid those creating multi-lingual gazetteers. It does not affect mono-lingual gazetteers. This is described in **3.3**.
- Addition of a facility for classification of streets:
 - For many purposes, it is desirable to classify the streets, for example according to their vehicle carrying capabilities. This addition is optional, and its usage is a decision for the implementation.
- Extension of facilities for cross-referencing to other datasets:
 - These have been added throughout the Standard. They are optional in all cases. Cross-referencing to other datasets is very useful, particularly to integrate existing datasets. How it is carried out is a matter for the implementation.

¹ This may be a PROW (Public Right of Way) number or cycle track number where appropriate.

- Removal of the tolerance attribute and the replacement of the spatial locator by a pair of extremity points:
 - This is a simplification. The tolerance attribute was designed to enable snapping together of points in building up a topographic network of the streets. It defined a circle of influence whereby points that were near to each other could be considered to be the same. However, it was little used. The spatial locator, which positioned the street in a coordinate space is consequently replaced by a simple pair of extremity points.
- Removal of the requirement for the identifier of an elementary street unit to be the coordinates of a reference point, and its replacement with a general identifier:
 - Previously, the identifier of an elementary street unit was a coordinate. This was unnecessarily complicated, and has been replaced by a general identifier. The existing identifier can still be used, but the change allows a simpler approach for new street data.
- Addition of informative annex on Public Rights of Way:
 - Previously, Part 4: *Specification for recording data for public rights of way* was a previous part of BS 7666. This is closely linked to Part 1, and the content is now included as an informative Annex, as an example of how additional data about a street might be recorded.
- Other minor changes:
 - Clarification of state as logical state of street (see **3.2.6**);
 - Facility to include a textual description for an elementary street unit;
 - Removal of fixed formats for attributes;
 - Introduction of levels of conformance for gazetteers, formalising high-level options for level of detail recorded;
 - Revision of definitions to improve understanding.

3. Explanations of requirements

3.1 Gazetteer Scope

The scope is a description of the type of objects to be included in the gazetteer. In the case of a street gazetteer, it is the types of street. The scope statement defines the bounds of what should be included and what should not be included. For example a gazetteer might include all streets accessible to motor vehicles, but not those on which motor vehicles are prohibited or unable to pass. The scope statement should also include a statement of what level of detail the gazetteer contains, i.e.:

- Level 1: only streets;
- Level 2: streets and elementary street units with no intermediate points;
- Level 3: a full description of the elementary street units.

For further details see Annex A of BS 7666-1.

It is important to have a clear scope statement, because of the loose definition of a street, to ensure that compilers know what to include, and users know what to expect. It is also important that where gazetteers from different sources are combined, they have similar scopes, otherwise an inconsistent gazetteer will result.

Example of Gazetteer scope for a street gazetteer:

Level 1 Street Gazetteer for the county of Dorsetshire, including all streets with designated street names, all streets with county Highway Authority numbers, Department for Transport numbered roads (class A and B), Motorways and any other streets used for accessing and addressing properties.

The statement of the gazetteer scope is recorded in the gazetteer metadata (see **3.4**).

3.2 Street records

3.2.1 Attributes

Details of the attributes of streets are defined in Clause 6 of BS 7666-1. Here some of the issues relating to them are described.

3.2.2 Unique street reference number (USRN)

The USRN uniquely identifies the street within the gazetteer, and is constant over the life of the street. It should not be changed when attributes of the street change. Where street gazetteers are to be merged to cover a wider area, there needs to be a central authority to allocate street number ranges to ensure uniqueness. This might be achieved by the addition of pre-fixes.

3.2.3 Descriptive identifier

The descriptive identifier is a spatial reference in the form of a description of the location of the street, for example 'Green Street, Cam, Dursley, Gloucestershire'. Sufficient elements must be present to make the descriptive identifier unique within the territory of use of the gazetteer.

A street may have more than one descriptive identifier. Where the gazetteer is multi-lingual, the descriptive identifier will include a code to identify the particular language. Ordering of descriptive identifiers is not important. Note that the language is the nominal language. This may contain names in other languages, for example Welsh contains some English words and vice-versa.

All names should be given in full. Abbreviations and punctuation should not be used unless they appear in the designated name (e.g. ‘Earl’s Court Road’). Only single spaces should be used, and the use of leading spaces should be avoided.

3.2.4 Record type

The record type identifies the type of street reference. The allowable values are:

1. **Designated street name:** as allocated by the Street Naming and Numbering Authority;
2. **Street description:** a concise description of a street that does not have a name, e.g. ‘King’s Parade - service road behind numbers 101-119’;
3. **Street number:** a number allocated by a highway authority, either the local Highway Authority or Department for Transport (or Scottish equivalent);
4. **Unofficial street name:** a name given by other than the Street Naming and Numbering Authority, for example a private street that has a name.

3.2.5 Street classification

A street classification code has been added to the 2006 edition of BS 7666-1. However this is optional. A recommended list of codes is given in Annex C of BS 7666-1. This does not have to be used. Where a different code list is used, it should be specified in the gazetteer metadata (see 3.4).

3.2.6 State

The state is the logical state of the street, for example whether it is under construction, in use or stopped up. This is an optional attribute. A recommended list of codes is given in Annex C of BS 7666-1. This does not have to be used. Where a different code list is used, it should be specified in the gazetteer metadata (see 3.4).

3.2.7 Geographic extent

The geographic extent of the street is described by a pair of extremity points. These are two points at the ends of the street, each represented by a coordinate pair. Some general rules for extremity points are as follows:

- a) If the street has more than two ends, then the two points should be the two ends that are farthest apart.
- b) If the street has only one identifiable end, then the second extremity point should be the identifiable point (e.g. a corner or junction) that is farthest from the identifiable end of the street.

- c) If the street does not have ends (i.e. it is circular), then any identifiable point (e.g. a junction) should be used, and the same coordinated point may be taken for both extremity points.
- d) Where properties in the street are numbered, the order of the extremity points should follow the order of progression of property numbers.
- e) Coordinates should be measured to an accuracy of not less than 10 metres (where a metric grid is used) relative to the other recorded coordinates.
- f) If the street is less than 10 metres long, then the two extremity points may have the same value.

Note that streets can overlap. Examples of extremity points are given in Figure 1.

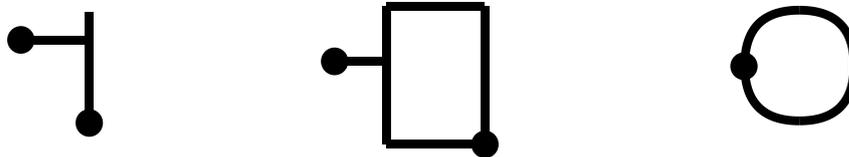


Figure 1. Extremity points for a street.

3.2.8 Dates

Dates should be recorded at an appropriate level of resolution. Normally this will be a day, but where this is not known, it may be only a month (e.g. '2006-08') or a year (e.g. '1900'). The dates should be recorded consistently either in the basic format (YYYYMMDD) or extended format (YYYY-MM-DD, YYYY-MM or YYYY), where YYYY is the year, MM the month and DD the day. The two formats should not be mixed, and for each implementation it will have to be decided which format is to be used.

Care needs to be taken to distinguish between actual dates, when something happened either in the real world or to a source of information, and capture and update dates, when changes are made to the data.

Actual dates:

- **start date:** the date when the street came into existence;
- **end date:** the date when the street ceased to exist;
- **current state date:** the date when the street came into its current state;
- **current date:** the date at which the gazetteer is considered to be current.

Capture and update dates:

- **entry date:** the date when the street record was entered into the gazetteer;
- **update date:** the date when the street record was last updated;
- **metadata date:** the date when the metadata was last updated.

Where dates are not known exactly, a notional date at which the date criterion was known to be correct should be used. This applies in particular to start date. Where the update date is not known or the record has not been updated, the update date should be same as the entry date.

3.2.9 Elementary street unit records

Details of the attributes of elementary street units are defined in Clause 7 of BS 7666-1. Much of the guidance for streets also applies to elementary street units.

3.3 Multi-lingual gazetteers

BS 7666 now contains explicit provision for multi-lingual entries (In previous editions of the standard, multi-lingual entries could be recorded through use of the ‘alias’ facility). The main requirement for this facility is in Wales, where many local authorities have bi-lingual gazetteers.

In a multi-lingual gazetteer, a record may have multiple descriptive identifiers, each in a different language. A code identifies the particular language used. These codes are listed in Table C.3 in Annex C of BS 7666-1.

It is not possible to mix languages in a descriptive identifier. However a Welsh language descriptive identifier may contain elements in English, and vice-versa. Ordering of descriptive identifiers is not important.

The languages used in addition to English should be recorded in the metadata (see 3.4), together with any additional character sets used.

Where the gazetteer is mono-lingual, nothing extra need be done, and the language is assumed to be English.

3.4 Gazetteer Metadata**3.4.1 Requirements**

The recording of metadata is a new requirement in the revised edition of BS 7666-1. The purpose of this is to provide users and potential users with information about the gazetteer. This metadata should be supplied with the gazetteer.

3.4.2 Mandatory elements

The mandatory metadata elements are as follows:

- a) **Name:** of the gazetteer, e.g. ‘Borsetshire County Council Street Gazetteer’;
- b) **Scope:** see 3.1;
- c) **Territory of use:** the geographic domain of the gazetteer, e.g. ‘Borsetshire’;

- d) **Gazetteer owner:** with overall responsibility, e.g. ‘Borsetshire County Council Highways Department’;
- e) **Custodian:** the practitioner who compiles the gazetteer, e.g. ‘the Borsetshire Street Gazetteer Custodian’;
- f) **Coordinate system:** this will usually be the National Grid (NG) of Great Britain, or Irish Transverse Mercator (ITM);
- g) **Current date:** the date at which the gazetteer can be considered to be current in relation to the real-world (not the date of last update of the data).

These are described in more detail in Clause 5 of BS 7666-1.

3.4.3 Optional elements

The following metadata elements are optional. They should be recorded where they are applicable.

- a) **Coordinate axis units:** this will usually be ‘metres’;
- b) **Metadata date:** when the metadata was last updated (not the date at which the data is deemed to be current);
- c) **Street classification scheme:** where one is used, details should be given (see 3.2.5);
- d) **State coding scheme:** where one is used, details should be given (see 3.2.6);
- e) **Language(s):** any languages other than English that are used should be recorded, e.g. ‘CYM’;
- f) **Character set(s):** any character sets other than English that are used should be recorded, e.g. "UNICODE UTF-8 character set" to support the full Welsh language;
- g) **External cross-reference(s):** If external cross-references are included in the gazetteer, they should be recorded, e.g. ‘Ordnance Survey TOIDs’.

These are described in more detail in Clause 5 of BS 7666-1.

3.4.4 Creation and maintenance

It is recommended that the metadata is collected as part of the gazetteer creation process. This requires little additional effort, and may aid the understanding of the creation process. The metadata should be updated periodically to reflect changes to the gazetteer, as part of the data maintenance process. Additional metadata elements may be recorded as required.

3.5 Data quality

The Standard specifies a requirement for a data quality report. This is an assessment of the quality of the data in terms of the following:

- Lineage
- Currency
- Positional accuracy
- Attribute accuracy

- Completeness
- Logical consistency

These general aspects are discussed in detail in Section 3 of these Guidelines. Additional specific guidance for street gazetteers is given in **4.8** of this section of the Guidelines.

4. Recommended approach

4.1 Planning and initiation

4.1.1 Scope

Before starting a gazetteer implementation, it is necessary to define the scope of the gazetteer. The scope is described in **3.1**. It should state what is to be included in the gazetteer and what is not included, including the level of detail (whether elementary street units are to be included, and if so whether their course is to be recorded). Note that only objects that are streets should be contained in a street gazetteer.

4.1.2 Data specification

Any gazetteer implementation should have a data specification. BS 7666-1 defines a general structure for a street gazetteer, and allows several options. It does not define the detail of the content of the gazetteer. As part of any implementation, it is necessary to specify in detail the data to be included. This should include the following:

- **Description of the details of the implementation of the Standard:** definition of the data structures to be used, including referencing schemes (e.g. unique street number references), street classifications and state coding schemes;
- **Identification of attributes of each entity:** for streets and elementary street units (if included), the attributes need to be defined, including whether attributes identified in BS 7666-1 as optional are to be included, and if so the rules for their inclusion, any additional attributes to be recorded, field lengths and domains (allowable values);
- **Links to other datasets:** any cross-references to other datasets, and how they are to be established;
- **Rules for inclusion of instances of streets and elementary street units:** and in particular how they are identified and referenced;
- **Identification of the life-cycle of a street:** the stages in the life of a street, and how these affect gazetteer entries;
- **Identification of the source of the data:** a process needs to be established to collect the data. It should include the creation of new instances including how they are referenced, for example the process of street naming.

4.1.3 Acceptable quality levels

BS 7666 does not specify acceptable quality levels (AQLs), threshold values applied to the results of testing data quality to determine whether the data meets criteria determined from the data specification. These need to be established as part of the

gazetteer implementation. In setting AQLs, the user and usage of the gazetteer should always be borne in mind. AQLs are discussed in 4.2 of Section 3 of these Guidelines.

4.2 Definition and identification of streets

A fundamental issue in building a gazetteer is identification of a street and elementary street unit, and where these start and end.

A street is a “way or thoroughfare providing a right of passage on foot, by cycle or by motor vehicle”. A street instance is identified initially by a street name, a highway authority street number, or identification of a piece of unnamed street on a map. The following rules should be used:

- **Roundabouts:** these should be included, and may also be part of streets that pass around them, but mini-roundabouts should not normally be included, unless named (e.g. the Magic Roundabout, Swindon);
- **Dual carriageways:** these should normally be a single street, unless the two carriageways have different names or route numbers²;
- **Link and slip roads:** these should be included as part of the street to which they relate, unless separately named. Where the slip road relates to two streets, it should be considered as part of the one onto which it leads³;
- **Cycle routes beside carriageways:** these should only be recorded where they are separated from the carriageway
- **Non-streets:** things that do not fit the definition of a street should not be included in a street gazetteer, e.g. waterways, railways, buildings, groups of buildings.

Streets may need to be subdivided in the gazetteer. In general, a street instance should extend to:

- a) The end of a street where it is a close or a cul-de-sac;
- b) The point at which there is a change of designated street name;
- c) The limit of a street description;
- d) The point at which a street number ceases to apply;
- e) A change of locality, town or administrative area (note that a street need not end at these boundaries – see 4.4);
- f) The boundary of the territory of use.

4.3 Identification of elementary street units

Elementary street units (ESUs) are subsets of streets forming a continuous length. They are recorded primarily to relate the street reference to a physical extent, and to enable the recording of overlaps between different street references. They also enable the construction of street network topology.

² The individual carriageways will be held as separate elementary street units (see 4.3).

³ The individual link or slip roads will be held as separate elementary street units (see 4.3)

The following rules should be used:

- a) where there is a junction, the street should be broken into separate ESUs;
- b) where there is a change in the characteristics of the street (e.g. the introduction of access restrictions), the street should be broken into separate ESUs;
- c) where ESUs are recorded for a roundabout, there will be one ESU between each pair of adjacent streets coming into the roundabout;
- d) the two carriageways of a dual carriageway should be recorded as separate ESUs;
- e) link and slip roads should be recorded as ESUs.

It is important that elementary street units end at features that are identifiable in the real world. In some remote rural areas where it is required to break a long street, this might be a stream, bridge, bend, or boundary marker.

4.4 Street names and descriptions

4.4.1 Street names

The designated street name is usually to be found on the name plate on the street. However, these may not always be correct, and may differ between the ends of the street. Unofficial street names are ones that have not been adopted by the appropriate Highways Authority but may be in common usage, e.g. “The Great North Road”. Street names, whether designated or unofficial, should be recorded in full. Abbreviations and punctuation should not be used unless they appear in the designated name, e.g. “Dr Newton’s Way”. Only single spaces should be used.

4.4.2 Street descriptions

Street descriptions should only contain what is necessary to identify the street, such as the type of street, a description of the end locations or identifiable landmarks. Where the street lies within a large public or private establishment or estate, it should be identified to that site.

Examples of street descriptions

- Underhill Lane, track to South Farm;
- Smith Street, alley behind houses on the north side;
- Southmoor Trading Estate service road;
- M1 from junction 3 to junction 4;
- Grand Canal towpath on north side between Old Bridge and New Bridge.

4.4.3 Street numbers

A street number shall comprise the road or route number allocated to a street by a relevant authority such as the Department for Transport, the Scottish Executive, the Welsh Assembly Government, Northern Ireland Roads Service or the local highway authority. A county/highway authority reference should be prefixed by the additional character “Z”.

4.5 Descriptive identifiers

4.5.1 Components

The descriptive identifier holds the key to street references. It is there to give the street an address and to distinguish between streets with the same name.

The descriptive identifier has three components:

- **Locality name:** the formal or informal name of the local area where people live or where certain activities take place, e.g. East Wellow, Maybush, Chiswick;
- **Town name:** the name of a city or town, or a suburb of an administrative area that does not form part of another town, e.g. Luton, Gosforth, Ealing;
- **Administrative area name:** the formal name of the highest level local administrative area (i.e. a County, Unitary Authority, Metropolitan District, Council Area (Scotland) or District Council (Northern Ireland)), or Greater London, or an island or group of islands (e.g. ‘Shetland Islands’). The name of the area should be used rather than the name of the administrative authority.

The issue is how to identify these for a particular street, given that locality and town are not necessarily well-defined in terms of their existence and extent. Whilst it is only necessary to provide sufficient of locality name, town name and administrative area name to make the street name unique, there is merit in providing more than the minimum where this helps to locate the street. For example, even where there is only one occurrence of a particular street name in the administrative area, it will be helpful to have either a locality name or a town name to locate that street in the real world. Because a town is usually more well-defined, a town name should be used where possible.

Streets normally do not need to be broken at locality boundaries, but it may sometimes be helpful to do so for addressing purposes, where properties are numbered separately in each section, and it is necessary to identify the locality in order to identify the properties uniquely.

There are essentially two different approaches to creation of descriptive identifiers, the top-down approach and the bottom-up approach.

4.5.2 Identification of components

Whilst it is possible to identify the locality, town and administrative area for each street on an ad hoc basis, this approach is not recommended, apart from where deemed necessary to deal with problem areas. Instead, it is recommended that a structured approach to identifying the areas is adopted. The territory of use of the gazetteer is broken down into a set of levels as follows:

1. the administrative areas are identified:
These will be contiguous, as they all have well-defined boundaries;
2. towns are identified:
The definition of these will vary from place to place, as what is a town in a rural area will be too small to be a town in an urban area. Towns are disjoint,

and will not together cover the whole of the territory of use. Typically there will be 10-50 towns in an administrative area.

3. localities are identified:

These will be local area names identified from address records, maps or other sources. They will include villages, suburbs of towns, housing estates, industrial estates (e.g. Park Royal), large trading estates (e.g. Bluewater Shopping Centre), airports (e.g. Heathrow) and university campuses (e.g. Nottingham University). Localities may be sub-divisions of towns, associated with towns, or stand alone within the context of an administrative area.

It is recommended that lists of allowable values be established, and that only these are used in creation of addresses. It is likely that in rural areas, some additions will have to be made in order to identify unnamed streets that provide access to properties.

4.6 Linking to other datasets

The Standard explicitly allows links between street records in the gazetteer and other datasets. How this is implemented will depend on the form of the other dataset and the nature of the corresponding related object. The linkage will be geographical in nature, but the other object does not have to be exactly the same nor coincident spatially. The relationships may involve time as well as space.

In all cases, the link will take the form of the identifier or identifiers of the related object or objects in the other dataset being recorded in the street gazetteer. The nature of the relationship and the dataset to which the data is cross-referenced should be identified in the metadata (see **3.4.3**). This relationship need not be one-to-one (i.e. an object in one dataset may be related to more than one object in the other dataset, and the two objects may not correspond exactly).

Possible types of cross-references to streets are:

- direct cross-references to other street records to link other data, such as topographic map data, with the corresponding street identified by the appropriate identifier or identifiers. An example of this would be Ordnance Survey TOIDs.
- Identifiers of street records in other systems in the same organisations, for example historic gazetteers or highway records;
- Links to other related objects, for example Public Rights of Way.

Streets and street references are used in many applications. BS 7666-2 contains explicit links to a street as recorded in a street gazetteer and referenced by Unique Street Reference Number or full descriptive identifier, for purposes of addressing Basic Land and Property Units.

4.7 Data maintenance

A street gazetteer is not a static dataset, but a continually changing description of a set of real-world objects. Consequently, it is essential that a maintenance regime is established. There are three main stages in the life-cycle of a street record in the gazetteer, creation, change and closure. Different procedures are required for each.

- **Creation:** a business process needs to be devised to identify new instances of streets, and to collect the necessary data about them. This will involve some level of interaction with the life-cycle of the street, including with the Street Naming and Numbering process. It will need to be decided at what point in the life-cycle, the street is entered into the street gazetteer.
- **Change:** change to a gazetteer record can occur for many reasons. They essentially fall into two categories, those representing real-world change, for example as a result of changes to the alignment of the street, and those due to correction of data or insertion of missing data. The changes should include the date information. The USRN should not be changed. The metadata will need to be updated at regular intervals.
- **Closure:** a street ceases to exist when it is physically closed, or it ceases to exist as a named entity (e.g. where it is merged into another named street). The gazetteer record for this instance is amended to change its state, and to input a value for the end date. Historic records should not be deleted, as they may still be of interest, but may be archived.

4.8 Data quality

4.8.1 Requirements

BS 7666-1 specifies a requirement for a data quality report. This is an assessment of the quality of the data in terms of the following:

- Lineage
- Currency
- Positional accuracy
- Attribute accuracy
- Completeness
- Logical consistency

The data quality report should contain details of any tests carried out, including the test methods, the date of the test, the name of the tester and details of any source material or other information used. Where tests are carried out on samples of the data, these should be chosen at random, and the method of generation of the sample recorded. The general aspects of reporting data quality are discussed in Section 3 of these Guidelines. This section provides some additional specific guidance for street gazetteers.

4.8.2 Lineage

The lineage statement should describe how the gazetteer was created and how it is updated. These should include details of the processes and the sources of information.

4.8.3 Currency

The currency statement should describe the date at which the data was considered to be current. Tests should be carried out to check if this is actually so, as it may be that in certain areas, updates have not been carried out. Where this is the case, it should be recorded in the quality report.

4.8.4 Positional accuracy

Detailed positional accuracy checks can only be carried out on samples of the data, comparing the recorded position of extremity points with the true position as defined on a map, or by GPS survey. However, simple checks of coordinates can be carried out to determine whether the points are within the geographic extent of the gazetteer, and the locale of the various spatial units (administrative area, town, locality) identified in the descriptive identifier.

4.8.5 Attribute accuracy

It is not generally possible to test the accuracy of all attributes. However, some key attributes should be checked on a sample basis, especially the street names and other key identifiers. It is also possible to test attributes for valid values. In particular, the following should be tested:

- Any identifier, such as a cross-reference, to confirm that it exists;
- Designated Street Names and Highway Authority road numbers with other official records;
- Dates as being within pre-defined ranges;
- Codes having valid values from code lists.

4.8.6 Completeness

Completeness errors can be of omission (missing items) or commission (additional items). The following should be checked and reported on:

- Whether all streets have been included (testing for this is probably best done on a sample area);
- Whether there are any duplicate streets (including whether there are duplicate records with minor differences in attribute values, for example extremity point coordinates);
- If elementary street units are included, that all streets have elementary street units;
- That all mandatory attributes of streets and elementary street units (if recorded) have values.

4.8.7 Logical consistency

The particular aspects of logical consistency that are important for a street gazetteer are as follows:

- streets recorded in the gazetteer are actually streets as defined by the gazetteer scope;
- street references (unique street reference numbers and descriptive identifiers) are unique within the territory of use;
- associations between an elementary street unit (ESU) and a street: for every ESU there must be one or more streets associated with it (and identified by a valid USRN);
- ESUs with the same street reference, when combined make up the complete street;

- attributes are in the prescribed format;
- attribute values are consistent with the domain set out in the data specification;
- descriptive identifiers are recorded consistently in terms of whether optional fields have values;
- dates are consistent with respect to each other (e.g. end date, if it exists, is not before start date).

4.8.8 Quality management and Quality improvement

It is essential that a formal quality management system is introduced for the gazetteer creation and maintenance process. A simple flow model of gazetteer creation and maintenance is given in Section 3 of these Guidelines. Built in to this scheme should be the maintenance and improvement of the quality of the data. Users will be looking for continuing improvements in data and this means that there will be a requirement for further quality improvement. To achieve this there will need to be a mechanism for the following:

- Feeding back and acting on errors found in the data;
- Feeding back on improvement to processes by operators of those processes;
- Learning and applying lessons from the use of current processes;
- Managing change whether to the data specification or the acceptable quality levels (AQLs).

5. Common misunderstandings and errors

Because of the nature of street gazetteers, certain types of errors tend to be prevalent. Many of these errors can be overcome by a more rigorous approach to managing quality. These errors include the following:

- Objects being entered into a street gazetteer that are not streets, such as canals, railway lines, high-rise buildings and cycle lanes within carriageways;
- Entering the same street in the gazetteer several times, leading to duplicate records, often with slightly different data – for example the same street having several “unique” street reference numbers (USRNs);
- Excessive sub-division of streets (as separate streets) adding little or no value for the user and complicating the maintenance of the gazetteer;
- Elementary street units that when combined do not make up the whole of the street;
- Missing mandatory attributes, such as extremity points;
- Inconsistent descriptive identifiers, for example where two neighbouring streets in the same locality have different entries for locality, town and administrative area, even though there are no area boundaries interceding.

6. Conformance issues

Annex A of BS 7666-1 defines a set of conformance levels for street gazetteers. Three conformance levels are defined:

- Level 1 - a gazetteer containing only streets;
- Level 2 - a gazetteer containing streets and elementary street units with no intermediate points;
- Level 3 – a gazetteer containing a full description of the elementary street units.

Each of these is progressively more detailed than the level below.

To claim conformance to the Standard, it is necessary to demonstrate that a gazetteer fulfils all the requirements outlined in the conformance statement. Such claims should be independently verifiable.

Essentially, a level 1 gazetteer must have:

- a record of all streets within the scope of the gazetteer;
- all the mandatory attributes for each street record;
- gazetteer metadata;
- a quality report.

For a level 2 gazetteer, in addition to meeting all the requirements of level 1, the gazetteer must also have:

- a record of elementary street units for each street;
- all the mandatory attributes for each elementary street unit.

For a level 3 gazetteer, in addition to meeting all the requirements of level 2, the gazetteer must also have:

- street centre lines for each elementary street unit.