

## W3.1

# The past, present and future of the National Land and Property Gazetteer

Simon Barlow, NLPG Hub Manager, Intelligent Addressing Ltd

## Introduction

This paper outlines the build process for the National Land and Property Gazetteer and reports on some of the issues encountered along the way. It will outline the main phases of construction and identifies some of the inconsistencies that needed to be worked into the process. Finally it provides some detail on the update process which will prove critical to the NLPG's importance in the wider arena.

## Context

For services like the National Land Information Service (NLIS) to work there was a requirement to have a single national index of all places. An electronic national land and property enquiry service needed to have details of all residential, commercial and industrial property as well as land parcels. For example many searches that a local land charges officer undertakes relate to properties that have not even been built yet so details of the site that the properties would reside on needed to be contained within this index. To this end the NLPG could not simply be a list compiled from only those local authorities that already had their property holdings entered onto a local register. Or indeed a list that was added to every time another local authority digitised all their land and property data. The requirement for the NLPG was, as its name implies, for an index of all land and all property.

This requirement was taken on board and a deadline was set for all local authority data to be submitted to a central collation hub. Also at this hub national address datasets from the Valuation Office were collated for Council Tax and Non Domestic Rates. In order to bring all this data together a common standard was adopted, namely the British Standard for addresses – BS7666 (2000). Adopting BS7666 meant that all property and land data needed to be matched to the National Street Gazetteer. A copy of all the local street gazetteers compiled by each of the Highways Authorities in England and Wales in the National Street Gazetteer was downloaded to the hub. It was at this point that we hit the first hurdle. There were two street gazetteers missing – one in Wales and one in England. A series of phone calls, emails, letters and meetings ensued followed by more phone calls, emails and letters. Eventually the relevant street gazetteers were obtained and have now been incorporated into the NLPG. The problem is a little more complicated north of the border as there are currently 14 street gazetteers missing from Scottish authorities – a problem we will hit later this year when the Scottish Authorities link to the NLPG!

## Phase One: Local Authority Data

The first phase was to identify all local authority sites that were already maintaining a local land and property gazetteer to the BS7666 format. In undertaking this exercise we were quite surprised at the level to which some local authorities had already progressed in embracing BS7666 and the concept of a single index of places within their administrative area. A number of local authorities have been maintaining a land and property gazetteer for a considerable number of years and over that time had been integrating addresses with GIS platforms to add more value to the data they were maintaining. For some the introduction of a British Standard for addressing in 1994 meant that they could then secure funding and lead the way in working towards adopting this in their own systems.

Others we found were simply maintaining a list of addresses used by the council as a necessity and had not heard of BS7666. In these cases we encouraged them to convert their address data to BS7666 in order for them to link into the NLPG. For some this was a relatively straightforward task as they were already holding this data in MS Access (or in a couple of cases MS Excel!). If a dataset was being held to a BS7666 format then creating the necessary data transfer file from this software was a simple comma separated export of the data.

Perhaps the most significant issue we came across when verifying all the local authority datasets was the interpretation of BS7666. The standard lends itself to be quite open in its interpretation and anyone new to the subject faces a multitude of acronyms that seem to complicate the purity of a standard address. For example BS7666 has brought in the concepts of Basic Land and Property Units, Unique Property Reference Numbers, Land and Property Identifiers, Unique Street Reference Numbers, Primary and Secondary Addressable Object Names and Provenance to name but a few. This is a long way from the historical picture when addresses have been held either as single address strings (i.e. all street number, house name, street name, town name and post code elements concatenated), or as simple mailing addresses with no formal structure as to what goes on each of the address lines.

The biggest issues have been with populating the Primary and Secondary Addressable Object Names (PAON and SAON). When looking at an address in this format the address elements appear to read in a completely different approach to a standard mailing address of say 6 address lines or a single address string. Added to this there is a character formatting standard to both these fields which in a large number of local authority data submissions was ignored.

Following on from this the street name, locality name and town name does not appear in a BS7666 formatted address. All this is referenced to a number contained within the National Street Gazetteer. This has been the second issue we have encountered both with the national datasets that have been integrated into the NLPG and with local authority datasets. More often than not the combination of localities and towns for streets in a local address list differs from the localities and towns contained for the same street in the national street gazetteer. In these instances it is difficult to match property data to the NSG in order to conform to BS7666.

For some people this may all add up to a situation where they may consider the use of BS7666 inappropriate faced with the above complexities. However after learning the basics of BS7666 the structure is very rigorous and can adapt itself to all situations of addresses that have been thrown at it so far. Importantly anyone involved in maintaining a corporate address list should have detailed knowledge of BS7666 if we are to move towards joined-up government and knowledge management.

In compiling the NLPG experience has been gained in how various local authorities have tackled these issues and anyone looking to implement a Local Land and Property Gazetteer (LLPG) really must have a grasp on the following before it can be put to use corporately:

- An understanding of the concepts and workings of BS7666. If you are part of a team working on the LLPG then everyone involved needs to have knowledge of the relationships between BLPUs and LPIs etc.
- In order to synchronise your LLPG with the NLPG data needs to be sent and received in the NLPG Data Transfer Format specification. Many software suppliers have now incorporated this import/export functionality into the products on offer to local authorities. This to a large extent has removed the requirement for local custodians to develop their own tools. However if you are creating your own database structure to hold your LLPG then this is something that you will need to become familiar with.
- As a result of merging a number of disparate address datasets you will undoubtedly come across duplications of the same addresses. These will need to be rationalised in your LLPG. A simple check is to put the data into MS Access and sort the address information by street name/USRN, building name and street number, which will quickly identify these potential duplications.
- As a result of adopting BS7666 this means that a property or land parcel cannot be included on your LLPG unless it is referenced to a street as defined in the National Street Gazetteer. One of the most

important tasks to undertake when creating your LLPG is to resolve all the street mismatches and anomalies. This is something that a local authority will have to do by liaison with their relevant Highways departments. Experience has shown that there has not been one single local or unitary authority that has not had to add new streets to their NSG or seek to modify existing locality, town and county combinations.

- If you are using Ordnance Survey co-ordinate products to help produce the spatial element to your gazetteer then make sure that the tiles supplied to you are “cookie-cut” to your administrative boundary. This will mean that you will not have to resolve a considerable number of addresses/co-ordinates that are maintained by a neighbouring authority.

Local Authorities have been using a variety of software products, consultants and data matching companies to create their LLPGs. In fact a large number of the local authorities that are currently linked into the NLPG have undertaken the work themselves by utilising internal resources. There are no definitive ways of completing the work but it is clearly understood that the creation of a LLPG is not a small task for most authorities. In fact the supply of core address datasets to a software supplier or data matching company is the straightforward part and also the relatively inexpensive aspect. The costly and resource intensive parts actually arrive in physically maintaining the dataset and resolving the anomalies from the initial creation. However most are relieved to know that once the initial creation and anomaly resolution has been achieved this will never have to be undertaken again as all the cross-referencing and alternative/alias address information will have been captured.

### Phase Two: Draft data collation

For those local authorities that did not have a LLPG a “draft LLPG” was created from national addresses datasets. Data was compiled from the Valuation Office which provided details of all rateable residential and non-domestic properties as well as other entities such as public conveniences, bus shelters and advertising rights. This data was then converted to BS7666 and matched against the NSG. The initial match gave approximately an 85 per cent hit rate which has been improved over time to nearly 98 per cent.

As indicated above there are not an insignificant number of streets that do not appear on the NSG many areas. In order to get around this problem the Hub allocated temporary street references to these streets so that the valuation office data could be incorporated into the NLPG. As a local authority takes ownership of their draft data then these temporary streets will eventually disappear as the correct or new street is added to the relevant NSG.

Once all the data had been converted and matched UPRNs were then allocated to each of the addresses. As a result of receiving in as many LLPGs as possible we were able to adopt the UPRNs these local authorities were using as the national UPRN. Following this we then assigned unique numbers to the draft data – numbers that did not conflict with those in use by local authorities. These UPRNs have now been fixed and are in use by end users of the NLPG.

Once all the LLPGs had been collated into a central hub this data became the reality for these areas. The “draft” data was then merged on top of the local authority submissions in order to produce the definitive NLPG for England and Wales. At all times local authority data takes priority over any draft data. Consequently any addresses found within the draft data that are over and above those supplied by a local authority are flagged as “candidate” addresses. These addresses do not make it out to end-users until they have been confirmed by the relevant LLPG custodian as being valid.

Candidate data may be of two types:

- Alternative addresses – where a match has been made to an address supplied by a local authority. This could take the form of an alternative building name or street number or in some cases a differing postcode.
- New addresses – where an address from a national source is put forward to the NLPG and no match on street name, primary addressable object name or secondary addressable object name can be made.

These records are given a new nationally allocated UPRN and flagged for the appropriate LLPG custodian to verify.

Candidate addresses will only be forwarded onto a local authority when no match can be made against the NLPG. There has been a great deal of rigour put in place to minimise the number of candidate records being sent to a LLPG custodian. For example an organisation holding national address data can only submit candidate data to the NLPG in the Data Transfer Format. This requires them to convert their address data to BS7666 and match it against the NSG. The matching is then done at the NLPG hub where the relevant UPRN is sent back to the organisation. This means that any records that do not match will arrive at the local custodian with a grid co-ordinate, pre-matched to the NSG and converted to BS7666 so that the verification process is made easier.

### Managing candidate entries

Intelligent Addressing are often asked how many candidate records are likely to be received by a LLPG custodian and how should they be verified. A number of local authorities are now setting up the necessary internal channels in order to route these issues around the council in order to validate incoming addresses. For example one authority LLPG manager has tried to identify a contact in each of the linked departments for planning, council tax, electoral roll and environmental health that they can contact should they need to verify an anomaly in the LLPG. These contacts are briefed on how the LLPG works within their authority and the basics of BS7666. In much the same way the LLPG could be used as a triggering device within the council to notify other departments that new addresses have been added to the gazetteer and that the appropriate demands/letters etc should be sent to these new occupiers.

In terms of the number of candidate records coming through to a local authority this is really dependent on the number of linked applications within the council. The more local authority datasets linked the more likely it is that any alias address information for a property has already been captured and linked into the LLPG. Consequently new addresses coming through will be instantly matched. As discussed earlier the “draft LLPGs” were created from Valuation Office data for Council Tax and Non Domestic Rates. Therefore a LLPG created from the Council Tax and Non-Domestic Rating lists and the Electoral Register will contain nearly 100 per cent of all the addresses within any administrative area. The inclusion of other datasets such as planning, environmental health, social services and refuse collection lists will undoubtedly add greater definition to the LLPG and reduce the number of candidate addresses coming through from end users.

### BLPU extents

With the application of the initial NLPG Data Transfer Format specification the NLPG was originally designed to hold point data. Some local authorities have now been busy capturing polygon extent data and consequently the NLPG Technical Working Group introduced a new Data Transfer Format specification (version 6.3) to enable these local authorities to submit BLPU extent information. This means that a local authority can link to the NLPG at any level. Currently as there are only a limited number of local authorities that have captured BLPU polygon extents to BS7666 format the NLPG will not issue these out to end users.

### The Future of the NLPG: Updating the NLPG

In order for services such as NLIS and other national address based services centred around the NLPG to become more and more efficient a frequent feed of change information is required. The success of the NLPG is therefore contained within the update mechanisms and local authorities linking in their own LLPGs to the national picture. Part of the organisational framework needed to be put in place within a local authority creating a local land and property gazetteer is the link to the NLPG. This will ensure that the LLPG is both kept up-to-date at the local level and also in synch with the national dataset. The final stage of this process enables changes made at the local level to be fed back to end users of the NLPG such as NLIS. In essence the NLPG Hub acts as a collation hub for 475 separate local gazetteers so that one update can be sent to end users of the NLPG rather than 475 different update files.

Once a LLPG has been incorporated into the NLPG this becomes the definitive source of information for an authority’s administrative area within the NLPG and an authority takes on the responsibility of maintaining the accuracy of this element of the national dataset.

The NLPG has been designed so that large volumes of data do not have to be transferred between parties linking in to the NLPG. To this end change only updates from a Local Custodian need only be submitted to the NLPG and not a complete resupply of the entire LLPG. Any new properties being built in the area should be flagged as inserts. Changes to existing records are notified by submitting status/change pairs. Any record contained in the LLPG and NLPG in error is marked for deletion and must have a close date. The UPRNs associated with closed and deleted records are historicised but importantly the UPRN is never used again.

Local authorities are now beginning the update process with a large proportion opting for a monthly update regime. Each individual authority will initially agree a timetable of updates that supports their ability and constraints in terms of software, resources and physical changes made to their gazetteer. Over time the updates will become more frequent and for some, dynamic.

## Access

The NLPG will be made available to all local authorities signed up to the process via a secure web environment. This will allow authorities to verify USRNs, UPRNs and addresses from neighbouring authorities and also to check details of properties that they may have an interest in elsewhere in the country. This secure environment will also allow LLPG custodians to automate the update process of synchronising their LLPG with the NLPG through use of file transfer protocols.

The NLPG is already a continually changing national index of places. There will be over 100 LLPGs linked into the NLPG by the end of the year and well over half of all local authorities are already in the process of having or creating their own LLPG. The process of updating and monitoring the content and quality of the inherent data has already begun but there are still many challenges ahead in ensuring that the NLPG achieves its goal of becoming the referencing base for all address information in the country.