Ram

You asked me for a brief review of the status of BS7666 and it's Compatibility with xAL.

I have tried to describe the BS7666 standard as succinctly as possible and have tried to identify the issues as to compatibility with xAL.

I have had discussions with Dr Rob Walker Chairman of the BS committee responsible for BS7666 (IST/36) and he is amenable to the consideration of xNAL as part of that standard, for interoperability, though this would be dependent on somebody being found to fund the activity.

It is likely that any standard for names personal and corporate would come under his committees and he has agreed that xNL could be a good starting point, again if funding and politics don't get in the way.

I currently am a member of the technical Panel 3 of IST/36 that deals with BS7666.

I have highlighted problem areas in passing data from BS7666 in red.

Structure of BS7666
There are four parts to BS7666

Part 1 - specification for Street Gazetteer
Part 2 - specification for a Land and Property Gazetteer
Part 3 - specification for Addresses
Part 4 - specification for recording Public Rights of Way

In this context part 4 is not relevant since any Public right of way that had property would by that very reason alone be included in any street gazetteer.

Street Gazetteer

The data for the street Gazetteer is as follows:

a)   Street
b)   Locality
c)   Town
d)   Administrative area

notably there is one less thoroughfare possibility and one less locality possibility than UK postal addresses in the Post office file PAF (Postcode Address File).

The lack of a dependent thoroughfare type will lead to significant difficulties in handling some addresses since dependent thoroughfares must become primary thoroughfares, this is often inappropriate in say a postal file, but may be acceptable for Local Government asset management.

No difficulties were reported in this area when the National Street Gazetteer was created. This is probably due to the fact that the NSG was created by the Highways function and the majority of dependent thoroughfares in the postal file, are parades, arches and terraces entities that are not of interest to highways units, who would view them as a superfluous alias for a section of one side of the road. These types of entities are not appropriate to be recorded as buildings.

Land and Property Gazetteer

The gazetteer entity is defined as :

Postcode	Postcode as assigned	Alphanumeric	8	"Optional"

	by the Royal Mail for the address			(Where one exists)
Posttown	Posttown as assigned by the Royal Mail for the address	Alphanumeric	30	"Optional" (Where one exists) c
Secondary addressable object name	Number, name or description of the SAO	Alphanumeric	100	"Optional" (Where one exists)
Primary addressable object name	Number, name or description of the PAO	Alphanumeric	100	"Mandatory" (Always exists)
Unique street reference number	Identifier of the street that provides access to the address taken from the related street gazetteer	Integer	8	"Mandatory" (Always exists)
Level	The vertical position within a structure or relative to another address	Alphanumeric	30	"Optional" (Where one exists)
Official address marker	Indicator of whether the address recorded in the LPI corresponds to the entry in the official name and numbering register **	Alphanumeric	1	"Optional" (Where one exists)

** in practise in many authorities no such register exists and where it does exist it covers only the last few years and will not normally be updated for ad-hoc changes such as when property name changes with ownership changes.

The PAON and SAON fields are further split as the table below

Character position	Usage
1-4	Number
5	Suffix
6-9	End of range number
10	Suffix to end of range number
11-100	Text

There are some obvious problems with this in particular the assumption that suffixes are single character, and that a number is numeric only. In the UK we have cases of multi-character suffixes and prefixed numbers.

This splitting of fields means that we have in reality six fields being compounded as one:

Building Number
Building Number Suffix
Building Number Prefix
Building Number Range
Building Name
Organisation Name
Building description

Where number and range may include suffixes. However it is not difficult to disassemble numbers and number ranges.

A more difficult problem is when the text element of SAON/PAON is a description or an organisational name not a building name. In this case only by using some interpretation rules will it be possible to split out these cases since there is no indicator for content type.

This is not as difficult as it first seems since the gazetteer should be built according to the following rules

> *"The primary addressable object name shall be defined according to the following descending order of priorities:*
>
>> *1) a designated premises number for the Address;*
>>
>> *2) a premises name for the Address;*
>>
>> *3) an organization name for the corporate occupier of the Address;*
>>
>> *4) a description of the Address.*
>
> *NOTE if both the designated premises name and premises number exist, then both may be recorded. "*

Organisation name should be used only when it is in effect the building name (e.g. "Harrods")

However there are certain to be a proportion of addresses that are ambiguous and cannot be resolved in software.

Address Files

An address is based upon a logical data model comprising the following entities:

    addressable object, with sub-types:
    primary addressable object;
    secondary addressable object;
    address;
    street;
    locality;
    town;
    administrative area;
    Postcode.

Created Gazetteers

To date there have only been two Gazetteers created to these standards the NSG (National Street Gazetteer) and NLPG (National Land and Property Gazetteer), these are Local Authority created files that have not been through any open and independent quality assurance procedures and where there is currently a very high level of political posturing.

The essential issue is the acceptability of these files to underpin all government activity in the UK, a significant number of us in the GI community in the UK have made representation to the eEnvoys office that these assets are unsuitable as they stand for the proposed purposes.

Often NLPG is taken as being synonymous with BS7666, this is not the case NLPG is one specific implementation of Part 2 and doesn't preclude other Gazetteers.

Overall

There are some limitations in BS7666 that mean that converting data from that format to xAL would be difficult and could require significant programming, however moving from xAL to BS7666 is easy so long as the address elements selected from xAL are limited to those that are in the BS7666 structure explicitly or implicitly.