ODUG Benefits Case: Housing Land Availability Assessment – Data Constraints

Introduction

It is widely acknowledged that an inadequate supply of housing land is a constraint on increasing the number of new dwellings for a growing population – a priority for the public and, therefore, all political parties in the United Kingdom. Lack of supply is a strong contributory factor to the current housing affordability crisis.

The exam question Government and Local Authorities must address to construct a coherent plan for the supply of new homes is: Where should new housing be built? Getting the answer to this question right is hampered by a lack of coherent information on building land availability.

This is confirmed by the number of house builders who advertise on billboards, television and through other media for "housing land wanted". Other organisations such as the BBC have run campaigns to crowd source information on "wasted land" which could be used for house building or other beneficial purposes.

This raises the questions: Do we, as a nation, really not know what land is available for building more houses? or, Do we know but have a defective information infrastructure that prevents the necessary information from being brought together as an open resource? I argue that it is the second of these. All the information necessary to prioritise potential house building sites nationally exists in different places, however our current information policy prevents it from being brought together and used for this purpose.

The information sources

It may be argued that Great Britain is the best mapped country in the world at a large scale and with a high degree of currency. Ordnance Survey (OS)'s MasterMap is a unique database structured as a set of closed areas (polygons) each of which is classified. Every building, every piece of road, every stretch of water is recorded in MasterMap. Therefore every potential building plot in the country exists on MasterMap, usually as a set of adjacent polygons. Because MasterMap is a database, not a "map", it is possible to select all polygons with a particular classification. So, for example, you can retrieve from MasterMap every polygon that represents a building.

In some cases it is possible to identify what the building is, but this is not information that is stored in the MasterMap database, however every building, or part of a building (where one building is represented with multiple polygons) has a unique identifier, the TOID (TOpographicIDentifier). This allows anyone, including OS, to record information against any polygon in the OS MasterMap database, not just buildings. For many buildings OS also hold an address and a UPRN (Unique Property Reference Number).

As every polygon that makes up the land surface of Great Britain is recorded on MasterMap it becomes almost trivially possible to identify land not suitable for building. For example every polygon which is classified as water, railway, or highway can be excluded from any first search for housing land, because, with the exception of large scale regeneration, the cost of building on these features would be prohibitive.

If one were to search the OS MasterMap Database for these three feature types and they were plotted as a map what would emerge would be a series of networks dividing the land surface up into closed areas. The characteristic of these closed areas is that they are not highway, railway or water,

although some of these closed areas may have 'holes' where a lake or pond unconnected to the surface water network exists.

These closed areas are known as "blocks", an American term relating to the common rectangular street networks making up many American cities. In that case a block is an area of land surrounded by streets. In our case it is any area of land entirely surrounded by streets, water, or, railways as well as any other impassable feature such as a cliff.

OS MasterMap also allows us to know what is happening in the blocks. So by plotting buildings within the blocks it becomes easy to calculate how built up they are. This is done by, again, a trivially simple piece of analysis, calculating the total area of the buildings in a block and dividing it by the area of the block. This produces a statistical measure that can be expressed as a percentage of how built up a block is. While that information does not exist explicitly in MasterMap, it can be very easily inferred.

Much, much more can be inferred about the characteristics of blocks. Building size and configuration make it relatively easy to guess what a building is. Detached, semi-detached and terraced houses can be identified. Garages and outbuildings can be inferred fairly reliably. Commercial buildings such as supermarkets, warehouses or factories can be identified. By combining this information, blocks can start to be classified. Only certain types of block will be suitable for new housing. These can be filtered out and provide a short list of potential sites.

Whilst this analysis can be carried out using OS MasterMap, other than in research studies or proofs of concept, it has never, to the best of my knowledge, been done comprehensively to identify potential housing land. This is because, even though the OS MasterMap database is extremely detailed, there is information missing from it and the additional data it is held by other organisations, or part of a different data model.

When considering new building the configuration of the land is a major determinant of whether it is possible to build at all and as to the actual building costs. Ordnance Survey records height both as individual spot heights and as interpolated contour lines. So if a question needs to be asked about whether a block is flat enough to build on a slightly more sophisticated (but still routine) piece of analysis needs to be carried out to calculate how much of a block is flat, how much is sloping and how steep the slopes are.

So, again, using Ordnance Survey data it would be possible to classify potential house-building blocks on the basis of how much land they contain which is flat enough to build on cost effectively. This would immediately 'filter out' hilly or mountainous areas unsuitable for building.

Other data

The Ordnance Survey does not hold all the data that could be useful for a national housing land availability model, many other organisations hold important information. For example:

- The Department for Communities and Local Government (DCLG) holds information on restricted and protected land including green belt, SSSI (sites of special scientific interest) and AONB (areas of outstanding natural beauty). These could, again, be plotted on MasterMap so potential housing blocks could again be filtered to exclude those.

- The Environment Agency holds information on water courses and flood risk which can again be added. This data will be made open soon.

- The Department for Transport hold details of transport land including roads, railways, airfields, ports.

Each of these data sources could be added to a national housing land availability model, to continue filtering out potential house-building blocks.

Once all government data sources are included the proportion of the national land surface available for house building will become quite low. This allows the search process to be refined to include only on those sites which remain feasible for building. At this stage the simple binary classification suitable for housing / not suitable for housing becomes weak as sites need to be prioritised in terms of build cost, desirability of building, and public opinion. Much green land currently used for farming or recreation may not be in any protected category such as green belt, but may still be valued as an un-built resource and so its use will be restricted. For example: the re-use of large gardens for house building is controversial and policies for garden plots vary between local authorities. Nevertheless, once all land which is currently not available for building is filtered out prioritising the remaining sites is a feasible operation, to enable a fully informed debate about where we should, and should not, consider building new homes.

SHLAA (Strategic Housing Land Availability Assessment)

Local Authorities (LAs) carry out land availability exercises; however their starting point is not as comprehensive as that proposed above. LAs regularly ask all land owners who would like to develop their land for housing to declare all the sites that they would like to have considered. The LAs then consider the constraints on building on those sites and prioritise them in order of deliverability under current policies.

SHLAAs are made available as text documents with illustrative maps. To the best of my knowledge there is no single aggregated SHLAA for any of the countries in the UK. So the SHLAA is mainly of interest to those declaring land and taking part in the local planning process. SHLAAs exclude windfall sites which are undeclared but become available for any reason, also land not currently on the market but where owners, individually or collectively, may be persuaded to make it available if the price was right.

Some enterprising builders assemble sites in blocks of houses with large gardens by acquiring one house for access and making offers that are hard to refuse for slices of neighbouring gardens to create a viable housing development site.

Putting aside whether such 'garden grabbing' is desirable, or accords with local policy, it would, nevertheless, be helpful to automate the method of identifying garden grabbing opportunities, so that such opportunities could be exploited or resisted. The detail in OS MasterMap makes that possible.

Land ownership

Once land suitable for housing has been identified the process of finding owners, making offers and identifying future house building sites is dependent on data from the Land Registries; Her Majesty's Land Registry (HMLR) and the Registers of Scotland in the UK. It would be entirely feasible to make the UK's land registers open, as the cadastres are in many other counties.

In fact HMLR would like to make it's INSPIRE Polygon Index Map available as open data. This is the map of all land parcels each with a unique identifier from which individual land records can be accessed. However, the Ordnance Survey currently restricts the use of the HMLR Polygon data under its 'derived data' Licensing restrictions. Entrepreneurs, who would like use this HMLR data to make the bulk identification of potential housing land faster cheaper and easier, and for other useful purposes, are unable to do so.

Ordnance Survey are blocking such bulk data releases on the grounds that HMLR's land parcels were "derived" from Ordnance Survey Mapping, and may be used as a substitute for MasterMap or other Ordnance Survey products. This is not the place to discuss the intricacies of the relative positions, but is a good example of how the Ordnance Survey Derived Data License is precluding the use of open public sector data and third party data, both open and closed, in ways which have a great benefit to our society.

Consider, for example, the Environment Agency Flood Data which is to be made open very soon. How can it be put to best use when the best quality mapping data is not also openly available? The flood data is likely to be linked with open mapping solutions, in preference to MasterMap, because these solutions are unencumbered by restrictive licensing arrangements.

Conclusion

If a political decision was made to make the identification of developable housing land a priority a mechanism would be arrived at to make all the discovery and analysis discussed above quick, easy and inexpensive. That could be achieved by making all the necessary data available as open data under an unrestricted Open Government Licence (OGL) to allow agencies, individuals or entrepreneurs to carry out a national Housing Land Availability Assessment. Many other, associated and useful analyses could also be carried out.

Instead the ability of Ordnance Survey to fund some of its operations by restricting access to its data to all but the highest value users and to force all agencies whose data is tied to, or based on, Ordnance Survey data to restrict data access in the same way, appears to be taking priority over solving the national housing land availability crisis and other nationally important questions faced by our society.

The government and all political parties, prior to the next election, will have to decide whether to stick to closed data policies based on high transactional costs and low availability of data in bulk. Or whether, alternative sustainable funding mechanisms can be identified to maximise the benefits which can be delivered from the National Information Infrastructure we already have, but currently restrict the use of.

Note(s)

The Royal Town Planning Institute (RTPI) has piloted the concept of "A Map for England" which would have brought together all planning constraint maps onto a single base which would have allowed the planning constraints and policies applying to every polygon in OS MasterMap to be identified by simply pointing at the appropriate point on the map. This was shown to be technically feasible but impossible to achieve with the current commercial and licensing models adopted by Ordnance Survey and HMLR.

Ordnance Survey spends between £2m and £3m per annum (out of a total operation budget of about £105m) on legal costs to prevent unauthorised access to its data.

Bob Barr

Member, Open Data User Group
July 2014