# **8th March 2022 - Data linking and connectivity across the public sector webinar**

## The value of Unique Property Reference Numbers (UPRNs) in data linking and connectivity

## **Questions and answers from ‘Teams’ chat function:**

|  |
| --- |
| Q: Are UPRNs unique for properties even within the same building (e.g. flats in the same stair?) |
| A: Yes, every property has its own UPRN. Including individual flats within a block |
|  |
| Q: Do you mean household or same address |
| A: Property within a building |
| Additional Q: Yes it would be useful to know how this works for blocks of flats, where many households share the same address |
| A: Each flat/household within a property has its own UPRN |
|  |
| Q: Would e.g. a flat 8 and flat 9 within the same building (same common entrance) have different UPRNs? |
| A: Yes they would have two separate UPRNs |
|  |
| Q: Would all flats have an identical location? |
| A: All Flats eg. Flat 1, Flat 2 etc will have their own UPRN, and be "child records" to the "block" which will typically be a "Parent" UPRN such as "Forest lodge" |
|  |
| Q: But they may have the same 'location' as in geographic coordinates - which could affect how the data can be used? |
| A: As each property has its own UPRN, it won’t affect how the data is used |
|  |
| Q: What happens to the UPRN of a property if it gets split into two i.e a house being converted to two flats? |
| A: The ‘parent’ UPRN is retained and additional UPRNs attributed to each of the flats. See [Lifecycle of a UPRN](https://www.geoplace.co.uk/downloads/The-UPRN-lifecycle-V3-2015.pdf) |
|  |
| Q: Is the UPRN classed as 'personal data'? |
| A: No |
|  |
| Q: When you mention household are you referring to the Census household definition? According to 2021 Census "A household is defined as one person living alone or a group of people (not necessarily related) living at the same address, who share cooking facilities and a living or dining room (or area). |
| Q: Also, wondering about houses of multiple occupation. |
| A: BLPUs are digital representations of physical structures and are not intended to reflect the individuals or groups of people at a location.  The structure of an HMO property would be like the structure of a block of flat discussed above, with a parent and child records. However, HMOs are given a different classification so they can be identified in the data. Blocks of flats are classified as a ‘PP’ parent with ‘RD06’ children and HMOs have an ‘RH01’ parent with ‘RH02’ for each bedsit. If the internal structure of an HMO is unknown, then it will be classified ‘RH03’. |
|  |
| Q: Is a UPRN linked to USRN? e.g. if a property has a front street, back street & is on the end of a street so side onto another street? |
| A: It will be linked to the USRN on the street that the property is addressed from |
| A: Yes, each UPRN is linked to a USRN and this is listed as an attribute in the data |
|  |
| Q: If raw data is collected with just address text (as I would assume most is) how can we be certain that the same UPRN is being correctly applied to the same addresses across multiple datasets? Not all data is as 'perfect' as the examples you have provided? |
| A: the key is not to use free text entry in the first place. Use AddressBase as the lookup tool to select/enter addresses and the UPRN is embedded in the dataset from the start. If you have legacy free text data, it's a one-off matching exercise to add UPRNs |
| A: ‘Alternative’ addresses can also be held within the record which can help to identify one property known as different things |
|  |
| Comment: Most of the data I work with is not provided with UPRN data and we haven't got control over the data collection. So this is a problem we face quite often. |
| Comment: The strength of the UPRNs stability may also be a weakness, many people we are interested in live chaotic lives, moving between homes frequently. |
| A: I don't think that stability is a downside with UPRNs. Identifying citizens and linking them to where they live is a whole different challenge! The citizens move around; the properties stay the same (mostly) - so you want the UPRNs to be stable. It's the links between citizens and UPRNs that change frequently and need to be kept up-to-date. No small feat doing that, of course...! But far better to use UPRNs than free text addresses on citizen records wherever possible. |
| It's possible to store multiple UPRNs against an individual's record if necessary as they would with all the different addresses (although it will be dependent on the database configuration at their end). |
|  |
| Q: Does the mandated use of UPRNs mean we should always ask for it if thinking of a new address data share? And if so, are there reasons people could refuse to supply it/a risk people still don't know what it is? |
| A: It should be used to identify geographic locations. UPRNs are derived from AddressBase and Open UPRNs data product. They are not personal data. Consumers can check their UPRN using [FindMyAddress](https://www.findmyaddress.co.uk/search) |
|  |
| Q: Were UPRNs used in the latest census? |
| A: Yes - local authority Custodians and GeoPlace worked very closely with ONS - see <https://www.geoplace.co.uk/local-authority-resources/guidance-for-custodians/how-to/census> and <https://www.geoplace.co.uk/press/2021/developing-an-excellent-address-list-for-census-2021> |
|  |
| Q: Would a UPRN be assigned to the whole block of flats as well? I look at entire blocks rather than individual flats in my work, so separate UPRNs for flats would make matching difficult for me |
| A:Yes we would set up what we would call a parent Record that would have a UPRN in its own right, this is the building as a whole, coming off building (Parent) would be child records so separate flats and each would have their own UPRN's. I hope this helps. |
|  |
| Q: Once a local authority makes you aware of an address, what is the turnaround time before this data becomes available and which product will it appear in first? |
| A: It is reflected in the GeoPlace databases straight away. It is reflected in AddressBase Core weekly and AddressBase Premium every 6 weeks. Availability is changing to daily towards the end of this year [See this infographic](https://static.geoplace.co.uk/downloads/GeoPlace-GeoHub-Infographic-Landscape-Updated_1.pdf) |
|  |
| Q: What are the main challenges for local custodians in assigning UPRNs to different types of buildings? |
| A: Custodians work to the address standard BS7666 as well as nationally agreed Data Entry Conventions |
|  |
| Q: How does business logic assist in fixing data issues? Is this automated? |
| A: In some cases we will fix known issues within the ETL, but usually we will produce reports of known issues for the data owner or fix at source |
|  |
| Q: All Flats eg. Flat 1, Flat 2 etc will have their own UPRN, and be "child records" to the "block" which will typically be a "Parent" UPRN such as "Forest lodge" |
| A: Yes this is correct, depending on the layout of the block, the flats and the parent could be stacked up on top of each other, however local authorities do place these at the "Visual centre" of the property, meaning they will obtain the plans for new flat developments and locate the individual flats to their correct location based on the plans and building footprint |
|  |
| Q: Can you elaborate on this if possible, are you saying that you have come across a flat 1 and flat 2 being the same geographical location? |
| A: Ideally the source of address data should be from a data product such as "AddressBase" - this ensures the addresses in eg a system or website (address lookup) has the UPRN at source in the system and you can use it without any uncertainty. IF your system is not capable of ingesting address data (AddressBase) then the manually entered addresses will need to be matched to the UPRN first. There are many ways to do this (software / services) but the key here is to move towards a sustainable process, i.e. getting address data with its UPRN used at point of entry to avoid continuous matching. We can help with address matching if need be. |
|  |
| Q: What was the process of matching the text addresses to UPRNs? Is this process open? |
| A: (From a local authority Custodian) We use a tool by Aligned Assets called 'iMatch' which assists with fuzzy matching, so it's a semi-automated process. To give you an idea about time and effort, we would probably allocate around 1/2 day to a day for matching a dataset of c.100k records |
| A: (From a local authority Custodian) We would use address matching software called ClearCore by Infoshare, this would take around half a day or a day to match a large number of records. |
| GeoPlace can also help with address matching |
|  |
| Q. How do you deal with complex parent/child UPRN hierarchies? e.g. a flat in a tower, which is part of a larger structure. |
| A: We accept multi level relationships - colloquially known as grandparents |
|  |
| Q. How and when is the link between UPRNs and their associated MasterMap Topographic Identifier (TOID) made and maintained across the UPRN lifecyle? |
| A: There are no links between UPRNs and TOIDs recorded in the data. UPRNs have coordinates so it is possible to find an associated TOID via spatial analysis |
|  |
| What are the main challenges locally with maintaining an up-to-date, comprehensive set of UPRNs? And is there a community of practice for practitioners to share experience/best practice on addressing challenges? |
| A: GeoPlace runs a closed community of practice for local authority Custodians |
|  |
| Q: My understanding is that the coordinates for a single building/structure can be used for all the flats inside. |
| A: That depends on the physical layout of the building. If the flats are clearly separated, most custodians will create the locations accurately even for flats. Obviously this isn't possible for high rise buildings etc, but generally it's possible to do accurately it will be done |
| It is common for stacks of flats (eg. when each floor has the same layout) to share the same coordinates |
| Further comment: I wouldn't use the same coordinates for the flats and the single building, I would create an accurate location for each individual flat as the site and layout plans I receive from a developer will show me exactly where a single unit is located. I know this may not be the case for some authorities. |
| Reply: I think most people facing the task today would do the same because the understanding about the value of accurate data and the potential for analysis and insight is there. But the majority of coordinates (20+ million) were collected over many many years. |
| This is why local authorities are the right people to manage this data - they don't only update it at the point of capture, they're reviewing and improving the data all the time. |
| Yes you are correct and this is why the local authority would manage the address data, I make sure that the data is reviewed and improving all the time, this is working closely with many internal integrated systems and constantly communicating with external contacts (GeoPlace, Royal Mail, VOA). |
| There is the assumption there that we have any control over the initial process. Working with administrative data means that we often don't have any control over the end-to-end process of data collection. We can make suggestions but can't control if this is taken up. I agree it would be great if everywhere moved towards so this, but at the moment this still not happening in the majority of cases (at least for the data I work with). |
| Q: We would have the same issue. I'm not sure some of our data providers would have access to AddressBase e.g. is there an easy way for a building owner to look up their building's UPRN? |
| A: Yes, use https://www.findmyaddress.co.uk/search |
| A: Really interesting to hear, is this more of a "technical" issue with say your system supplier not knowing about the data, or is it more of an organisational issue in that people don't see a need to use the UPRN or are aware of it? The Central Government Mandate is there to hopefully help provide this impetus, but if it doesn't gather traction then please let us (GeoPlace) know - we would be happy to reach out to the wider departments in question |
| In many of our instances, a lot of address info will be 'user provided' - so, for example, writing your address on a form when signing up for something. We can't expect the whole population to a) know their UPRN or b) fill out their address perfectly every time. I guess our question is how we ensure consistent accuracy in cases like this? |
| A: This is where data matching exercises can be carried out to ensure the data collected will be assigned the correct UPRN, GeoPlace is more than happy to help if you have further questions regarding this data. [support@geoplace.co.uk](mailto:support@geoplace.co.uk) |
| You can check UPRNs with this tool - <https://www.findmyaddress.co.uk/search> |
|  |
| Q: If I got some questions about the geo-reference data with UPRNs, may I ask what is right email address to connect with GeoPlace? |
| A: Happy to talk - contact us at [support@geoplace.co.uk](mailto:support@geoplace.co.uk) |
| Here are some FAQs as well - <https://www.geoplace.co.uk/addresses-streets/data-in-use/faqs> |
|  |
| Q: In our linkage work we've come across issues with linking flats; we very often ended up having the same UPRN for a number of flats and cannot accurately link them. We've been trying to find the source of this problem but without too much success |
| A: the source of the problem is with the data capture: if the collector of this data used AddressBase as a lookup, all of the flats would be available with different UPRNs at the point of capture. |
| A: This may be the case that the "Parent" UPRN for the entire block of flats is being selected by users, such as "Sunshine Lodge" rather than "Flat 1, Sunshine Lodge" - happy to talk offline to help with this (luke.studden@geoplace.co.uk) |
| Further comment: Difficult to help much without exact examples, but I'd start by looking at fuzzy matches - sometimes these remove necessary characters so multiple address look to the same to the matching algorithm |
|  |
| Q: You have gained more coverage of UPRN on your data but has any quality assessment been done to make sure this additional coverage is all 'good quality'? |
| A: There is a very rigorous data assurance process - as shown by this infographic - <https://www.geoplace.co.uk/about-us/who-we-are/our-work/geoplace-central-source-for-uk-addresses-and-streets> |
| GeoPlace works very closely with LAs to support and quality assure data under the joint governance process called the Data Cooperation Agreement - <https://www.geoplace.co.uk/local-authority-resources/guidance-for-custodians/how-to/data-cooperation-agreement> |
|  |
| Q: The central government Mandate isn't something I've come across before, do you have a link to any further info on it? |
| A: <https://www.gov.uk/government/publications/open-standards-for-government/identifying-property-and-street-information> |
|  |
| Q: I then wondered how could a member of the public find out themselves the UPRN of their owned / rented property, if they wanted to? |
| A: they can use this tool - <https://www.findmyaddress.co.uk/search> |
|  |
| Q: I understood AddressBase was the official address and UPRN, not PAF! Am I mistaken? |
| A AddressBase Premium and OS Places for example contain both the local authority address and the Royal Mail "PAF" address - there is an option to select either |
| A: The local authority maintains the definitive record of street names and property addresses, while a postcode comes from Royal Mail. |
|  |
| Q: Really helpful examples of the pitfalls of address matching. Do you know of any good resources to help with free text address matching? We're doing similar work in SQL currently. |
| A: If you are part of government Ordnance Survey is here to help you use the OS Places API and AddressBase with all the different flavours |
| We should move away from address matching in silos. I think this should involve openly sharing address matching algorithms, ideally extending to source code.  Building new standards (e.g. for assigning match confidence scores) could be important too. Probably a bridge too far, but I'd also openly share lists of input address strings, the matched UPRN and confidence scores - so the community can collaborate on improving data quality and algorithms based on real world examples. |
| I am not 100% on this. We went through some learning to do what we needed to do on relatively small lists. I think there is a team at HMRC who are doing work on this area. I totally agree that doing this in silos is not optimal but appreciate there are also GDPR and clearance challenges in some areas. |
| I agree, it would be really helpful to have an open source algorithm to have a single source of truth where we need to assign a uprn for free text address. |
| About the point of address matching in silos, we have at the NHSBSA recently developed an address matching R package. All of the address cleaning and matching functions and code openly available on our github page. Configured to work with data stored within a database so large files can be matched: <https://github.com/nhsbsa-data-analytics/addressMatchR>. We used this to match patient records against AddressBase, feel free to get in touch if you have any questions/queries/how to apply within your own specific use case |
| Agreed. Because at the moment two organisations could (if using different, imperfect algorithms) provide the same UPRN for different addresses or different UPRN for the same address |
| This would be picked up by GeoPlace, investigated by GeoPlace or passed to the relevant local authority, so such errors would be picked up |
|  |
| Q: How far back in time (i.e. outdated, demolished, changed, etc) can we match addresses to UPRNs using OS Places API? |
| A: When data is available from varies by local authority because they did not all start producing gazetteers at the same time. Generally, properties which have been constructed since about 2000 will be present. Properties which were constructed before then but were still in use have also been added and many have retrospective start dates – some a few centuries ago. |
|  |
| Comment: From a Street Gazetteer perspective the bridges, level crossings and structures etc. are included within USRNs. In Powys CC we have captured these to ESU level |

Yes, every property has its own UPRN. Including individual flats within a block

Yes, every property has its own UPRN. Including individual flats within a bloc