

# Address Matching Overview

This page is intended to be a global overview of the address matching theme - the "Executive Summary" of the theme and all of its aspects.

- [Glossary](#)
- [SMEs](#)
- [Activities for this Discovery](#)
- [References / Materials / Meeting Notes](#)
- [Data Sources](#)
- [ONS Address Index](#)

## Glossary

**Register UPRN** - We will use the language "Register UPRN" to refer specifically to UPRNs on the EPBR. This is because a new Register UPRN can be created during the assessor process that is unique to the EPBR, and will not necessarily relate to what any other database (e.g. Royal Mail PAF) might call a UPRN.

Also see comment from [Phil Beschizza](#) on this page [Local Government user research](#), reproduced here: The register UPRN is an internal administrative number generated by the Energy Performance of Buildings Registers. There is an obligation on the Register Operator to generate and verify UPRNs for new addresses within a short timeframe or to produce sub-UPRNs for building units in larger commercial premises. The Register UPRN is not the same as more commonly used UPRNs, such as the National Land and Property Gazetteer or O/S UPRN.

## SMEs

[Jason.Ward@communities.gsi.gov.uk](#) (MHCLG)

[Phil Beschizza](#) (MHCLG)

[Steve Peters MHCLG](#) (MHCLG - more of the data perspective)

Paul Downy (MHCLG) (from perspective of bridging to ONS?)

## Activities for this Discovery

- Complete this executive overview (this page).
- Outline Address Matching options and data sources. e.g. compare existing use of Royal Mail PAF inside Landmark with other data/services e.g. ONS Address & Business Index, Ordnance Survey etc.
- Understand Address Matching activity internal to MHCLG, in the data team.

## References / Materials / Meeting Notes

All of the collateral we know about on the issue, including from previous Discovery etc. This section might also cover for "work already done". It represents all of the collateral we are aware of.

- On the page [Landmark documents](#) we have added a column to identify documents related to address matching, some of which are referred to here:
  - Two Landmark documents "Domestic / non-Domestic EPC Register Services to Accreditation Schemes", see [Landmark documents](#).
  - Two Landmark documents for Address Matching Proposal and Address Matching Training Manual, see [Landmark documents](#).

## Data Sources

This summarises data sources that various entities use for the address matching problem.

---

Data Source Description	Remark	Landmark Addressing Solution	MHCLG Analytics Group	ONS Address Index	ONS Business Index	BEIS
<b>Royal Mail PAF (Postcode Address File)</b> <a href="#">here</a>		Used as master dataset	Implicitly used in OS AddressBase Premium.	Yes		
<b>OS AddressBase products</b> <a href="#">here</a> AddressBase Products include:  1. AddressBase 2. AddressBase Plus 3. AddressBase Premium 4. AddressBase Islands			Uses AddressBase Premium  (uses PAF and Local Authority data)			
<b>OS Address Layer and Address Layer 2</b>	Formal notice of withdrawal issued on 31 October 2014.					
Business address information collected by <b>Experian</b>		Supplement				
The <b>National Land and Property Gazetteer (NLPG)</b> and the <b>Local Land and Property Gazetteer (LLPG)</b> , which are maintained by Local Authorities.			Implicitly used in OS AddressBase Premium. (t.b.c.)	Yes		
The <b>Valuation Office Address (VOA)</b> database, used for council tax purposes.						
National Business Database		Supplement				
QuickAddress (QAS) / Data Plus		Supplement				

## ONS Address Index

- Context: Valtech was involved with supporting the ONS Address Index project, including the data science behind it.
- Underlying data science approach: apply [Conditional Random Fields](#) to correctly identify subsequences of an address, then send those subsequences to [Elasticsearch](#) with bespoke subsequence boosting.
- Data sources: see above table
- Source code:
  - ONS Address Index on GitHub, see [here](#)
  - ONS Address Index, some of the Data Science [here](#).
  - ONS Address Index API [here](#) on GitHub. The 'address-index-api' is a front end REST API for the index in Elasticsearch. It includes a simple UI for searching.
- Key contact at the ONS: Alistair Calder ([Alistair.calder@ons.gov.uk](mailto:Alistair.calder@ons.gov.uk)) and also speak to Richard Smith (contact via Alistair).
  - Could ask for a demo over Zoom/Skype.
  - Run demo yourselves: Would need a copy of AddressBase (the free downloadable one might do for a test). Budget 1-2 days.