Using FME to quality assure the HMLR Local Land Charges register

Rod Else

HM Land Registry Snr Data Quality Manager

Jamie Di-Cataldo

Data Quality Analyst

Our ambition is to become the world's leading land registry for speed, simplicity and an open approach to data

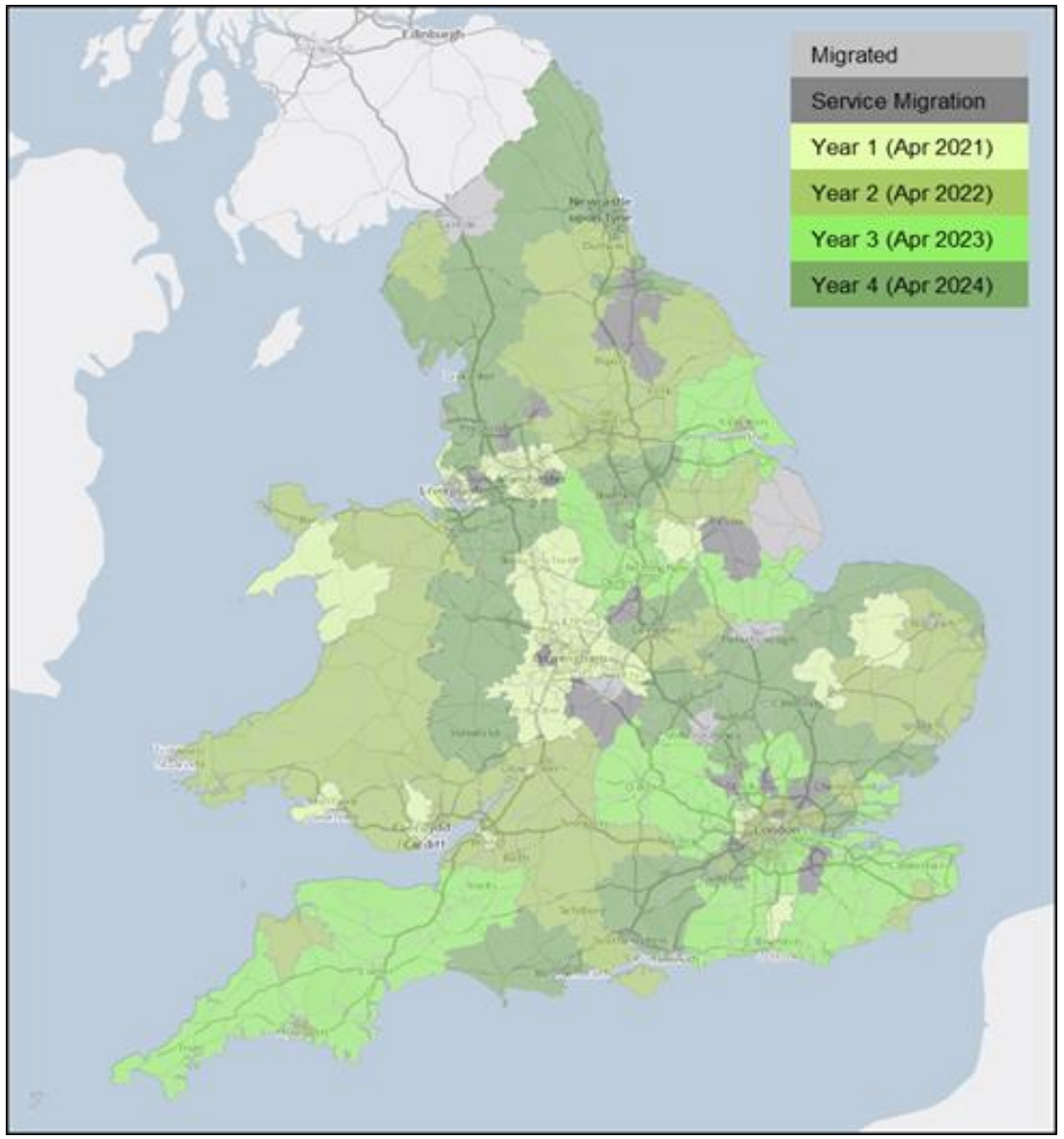
Local Land Charges : a diverse portfolio

Approx. 80% Planning

Also includes:

- Listed Buildings
- Ancient Monuments
- Tree Preservation Orders
- Smoke control Orders

Starting Position LLC Register Service



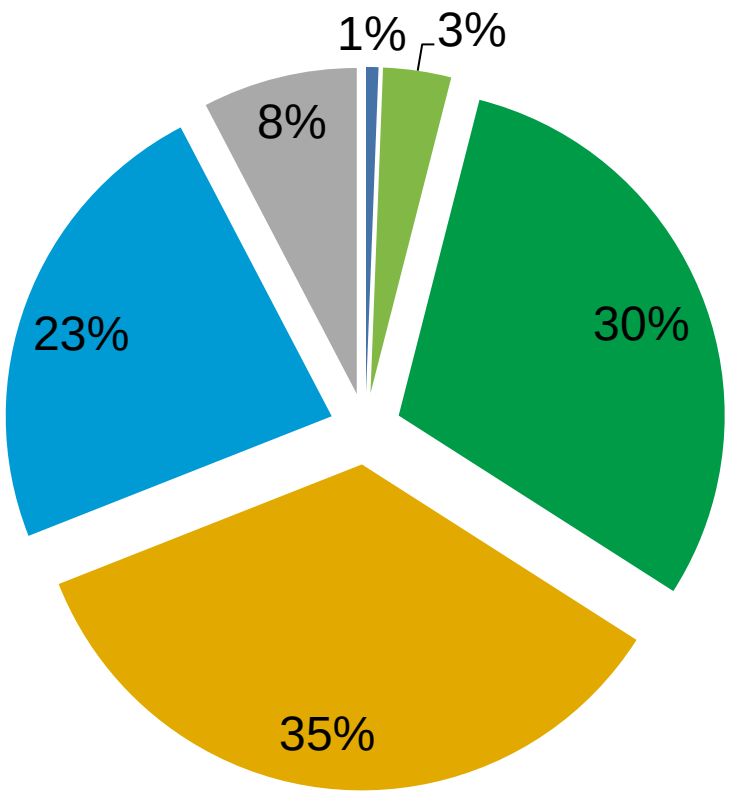
Local Land Charges data held and maintained independently by over 300 Local Authorities across England and Wales

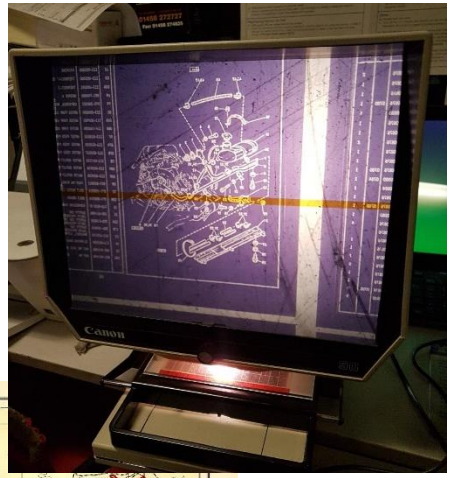
Dataset estimated to be over 30 million registered charges

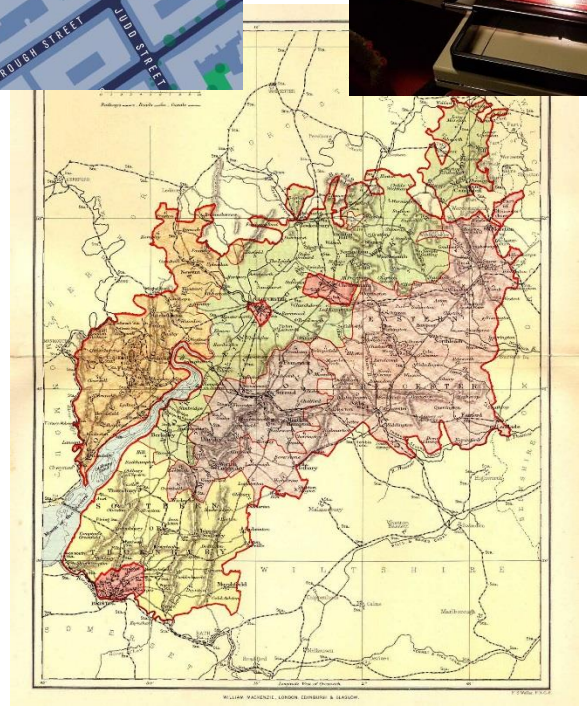
Primary legislation passes statutory function to HMLR; initial register build and development of migration approach complete

HMLR has started a 4 year plan to extract, transform and load the data into a single, online, geospatial register

A fragmented and disparate dataset







Maintain quality whilst achieving delivery

HM Land Registry

Year 1
44 Local Authorities

Year 2
89 Local Authorities

Year 3
93 Local Authorities

Year 4
98 Local Authorities

Data

Reputation

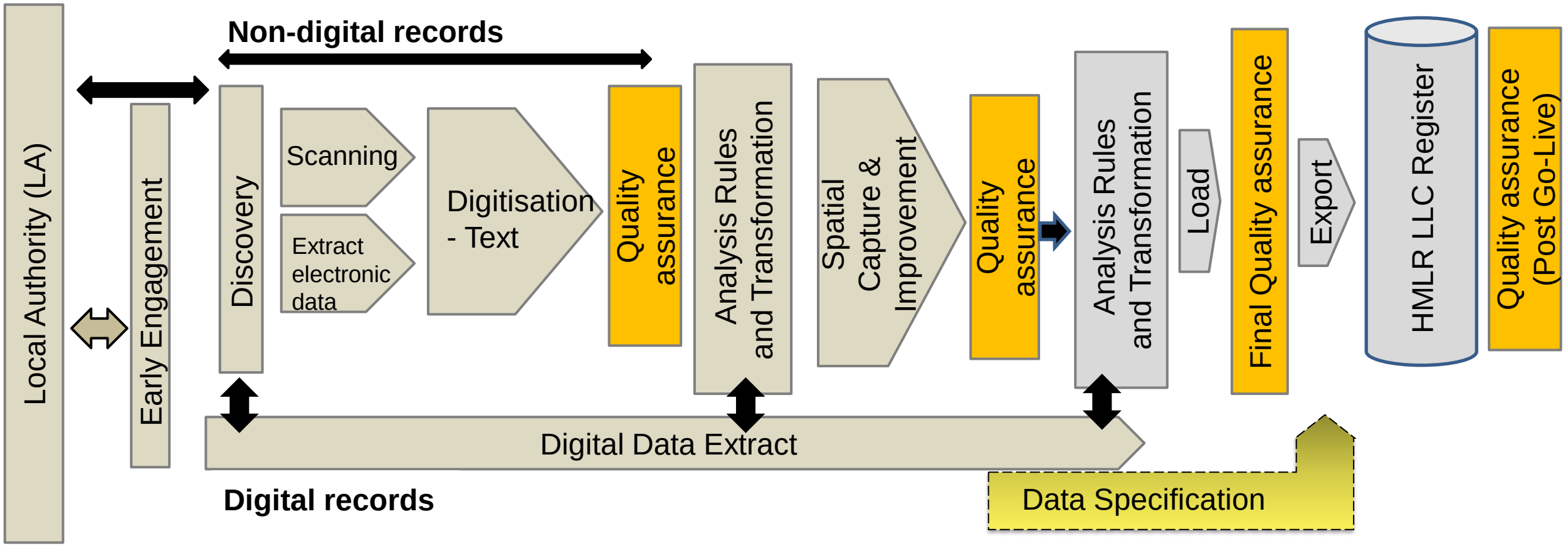
Legal

Customer

Cost/VFM

High level data workflow process

HM Land Registry



Key Principles:

- **Test Early**
- **Feedback quickly**
- **Automate wherever possible**

HM Land Registry

Using FME in our QA process

Why do we use FME?

HM Land Registry

To streamline the millions of records of land charges we have to Quality Assure

Automate as much testing as possible to reduce manual work

Our FME checks

HM Land Registry

Sampling

Geometry Validation

Polygon Shape

Textual Descriptions

Postcode Validity

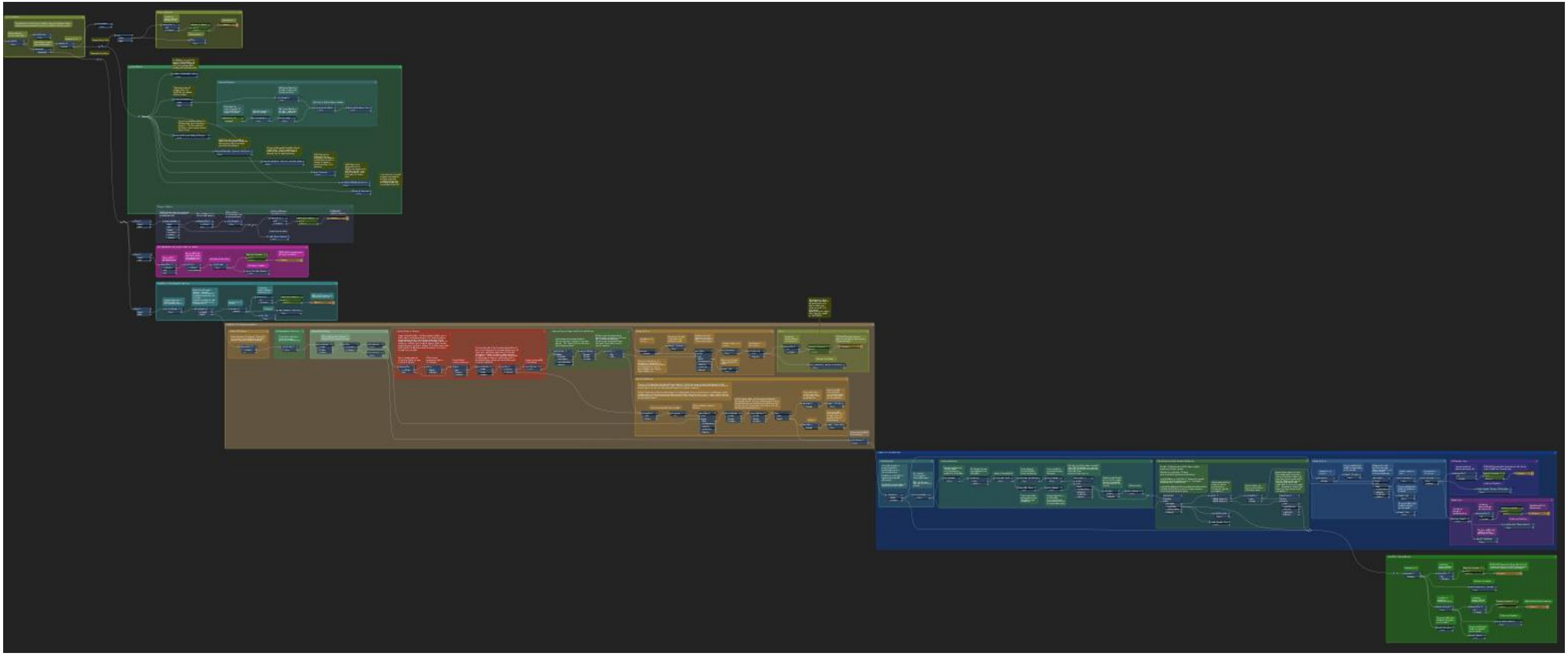
Feature Type

Charge Location

Charge Type

FME QA Process Overview

Each bookmark forms part of a different test until we reach a manual check to ensure 100% test coverage of charges



Case Study 1: Postcode Geometry Validation

Postcode Testing – Postcodes Explained

PE	Peterborough
PH	Perth
PL	Plymouth
PO	Portsmouth

- Similar to ZIP codes in USA or PLZ in Germany
- Alphanumeric codes devised by the Royal Mail (1959) eg. Big Ben in London would be “SW1A 0AA”
- Used to designate an area or a single major delivery point

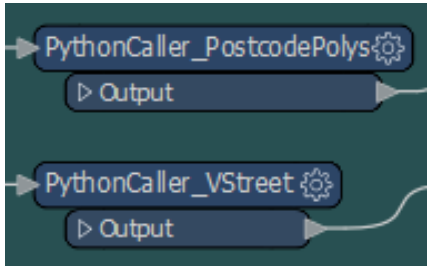
Postcode Testing – How we do this in FME

HM Land Registry

- Extract postcodes from geographic description on a charge

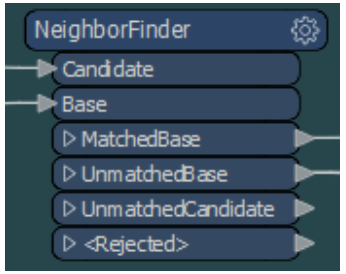
<u>Postcode Geometry</u>

- Get postcode geometry from OS source
- Python callers are used to look up first part of postcode against a vertical_street table
- This is to track any buildings which might have more than one postcode



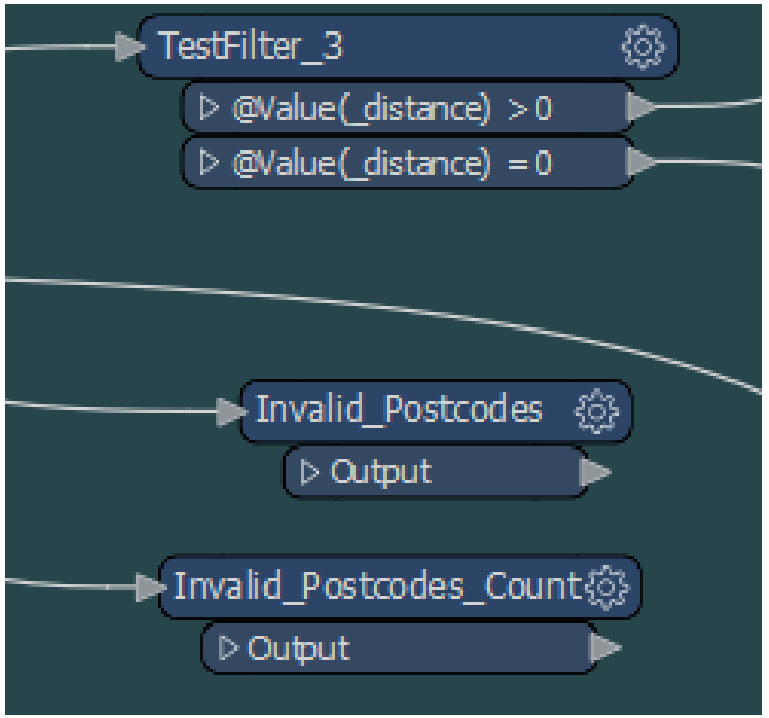
- Use a “NeighborFinder” transformer to find postcode geometry with same postcode

- We then use a test filter to identify any LLC charges which do not match their postcode geometry



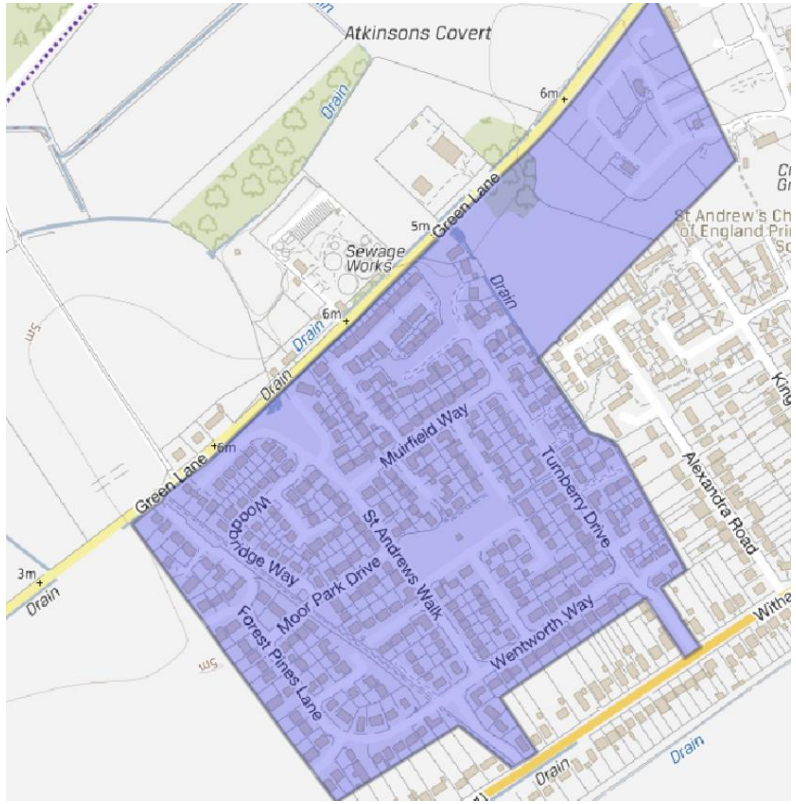
Postcode Testing – Results

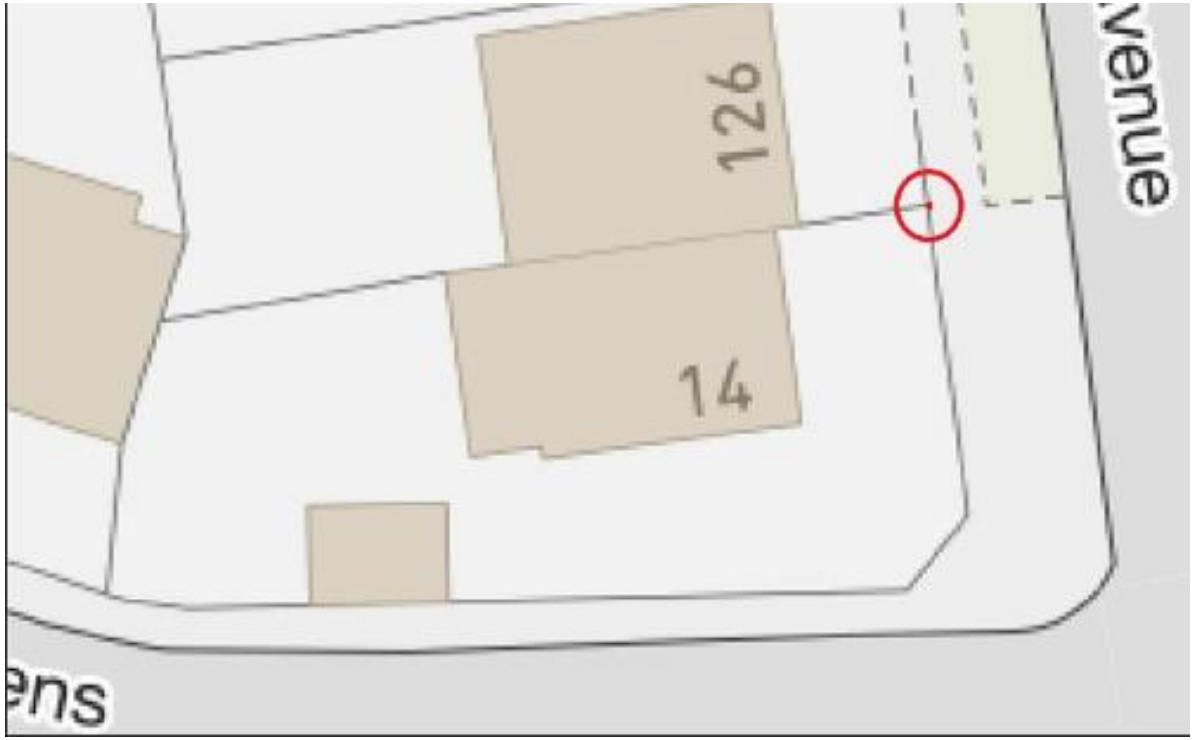
- Incorrect postcodes are produced in a list of “invalid_postcodes”
- Matching postcodes are then output in the form of an Excel output as well as a geopackage
- Anything outside postcode tolerance is filtered for manual sample testing, using LLC sampling rules



Case Study 2: Large & Small Polygons

Large & Small Polygon Problem





Large polygon

- A charge on a single house can cover this whole area

Small polygon

- A charge for a single-storey extension on no. 14 is represented as a small dot on the boundary between no. 14 & no. 126

Risks with Small & Large Polygons

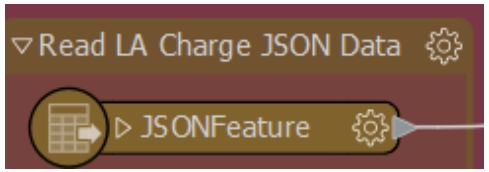
- Confusing for customers
- False Positives on charges for Large Polygons
- Data Quality
- HMLR reputation damage

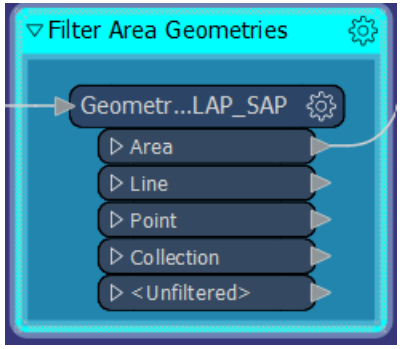
In order to provide the best service, we need to standardise the data!

Large & Small Polygon Process Map

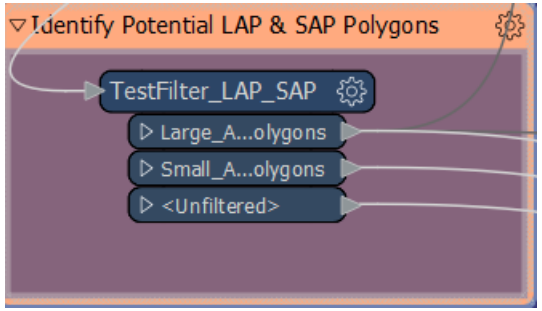
HM Land Registry

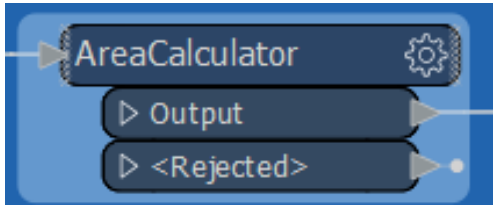
Input



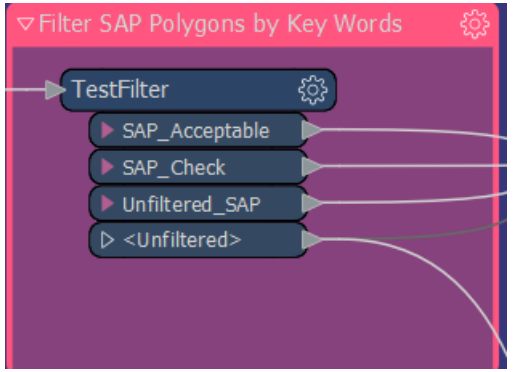


Identify LAP

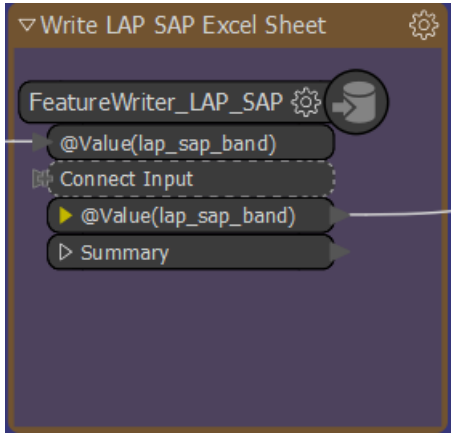


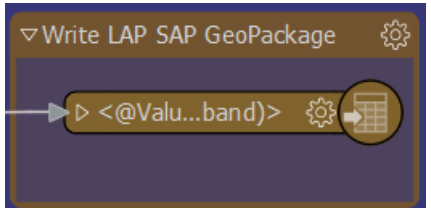


Identify SAP



Output





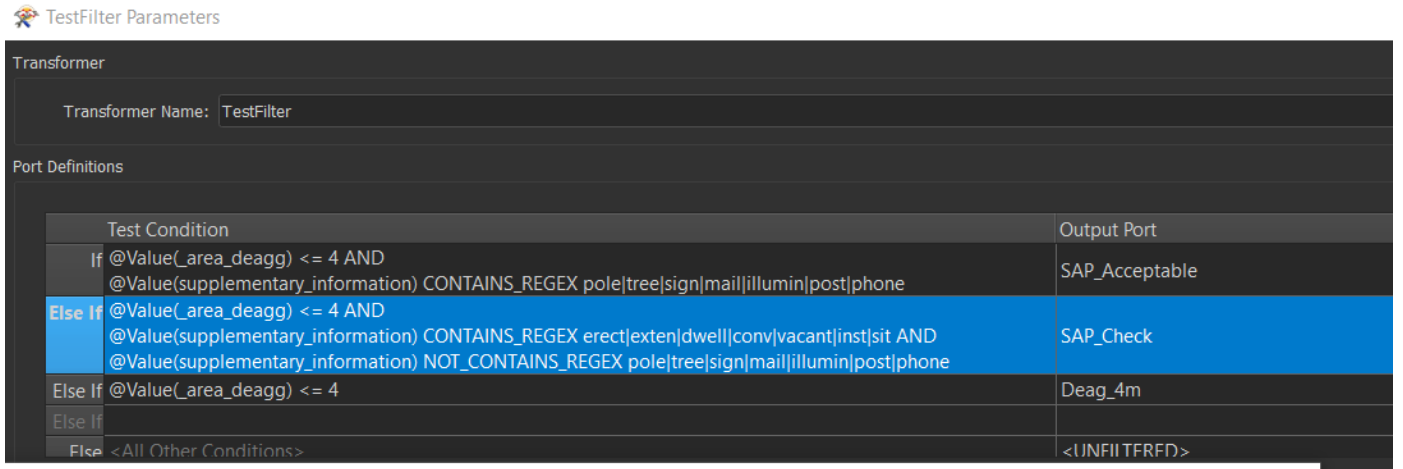
Using keyword filters to identify LAP & SAP

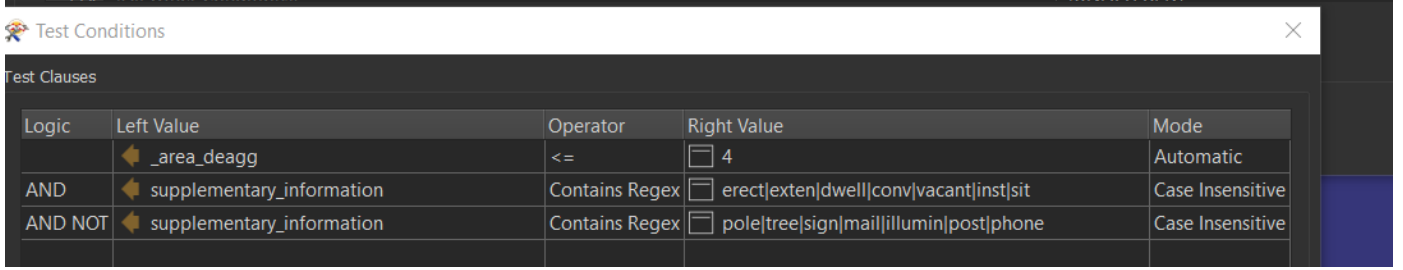
HM Land Registry

using regular expressions:

Include	Exclude
Erect	Year
Extension	Pole
Change	Tree
Dwelling	Sign
Convert	Royal Mail
Vacant	Illuminated
Install	

Small filter

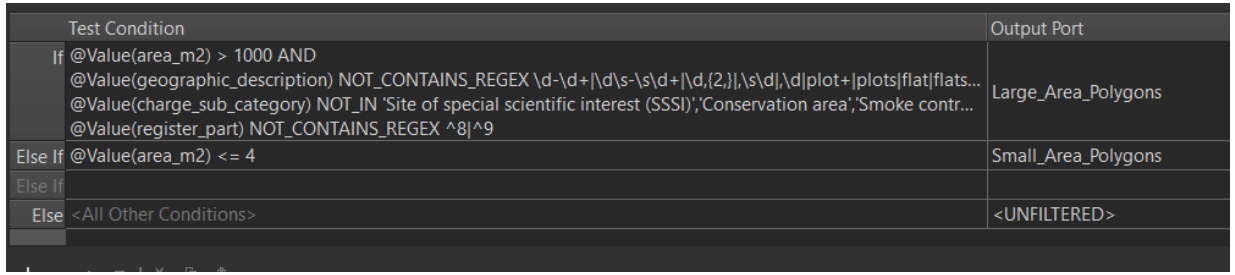




Large filter

To be excluded in large filter

Activity	Listed Types	Notes
Mining	Old register part 9	
Special control advertising		No way to identify
Outline planning		No way to identify
Civil Aviation	Old register part 8	
Planning	Conservation Areas	
	No permit dev. Article 4	
	Planning Agreement (section 106)	
Other	New Towns	
	Pipeline	Managed by xxxx
	Protected Sites	SSSI etc
	Smoke control Orders	



FME Output: How do we use this data?

Quality Theme	Quality Criteria
Completeness	Omission
	Commission
Thematic Accuracy	Classification
Logical Consistency	Conceptual (adherence to Digitisation Practice Guide (DPG)
	Domain (adherence to wider rules)
	Topological Consistency
Positional Accuracy	Relative position

Used to match against our company test suite

In Particular:

Quality Theme	Quality Criteria	Example Quality Measure	AQL
Positional Accuracy	Relative position	Is there an acceptable relationship between the land description and the geometry	99%

Acceptable Quality Level (AQL): Is the geometry in the correct location for 99% of the charges?

Questions?

Jamie.di-cataldo@landregistry.gov.uk

HM Land Registry