

GeoPlace Leeds May '17

Using your addresses for Census 2021

(and so much more)

Alistair Calder

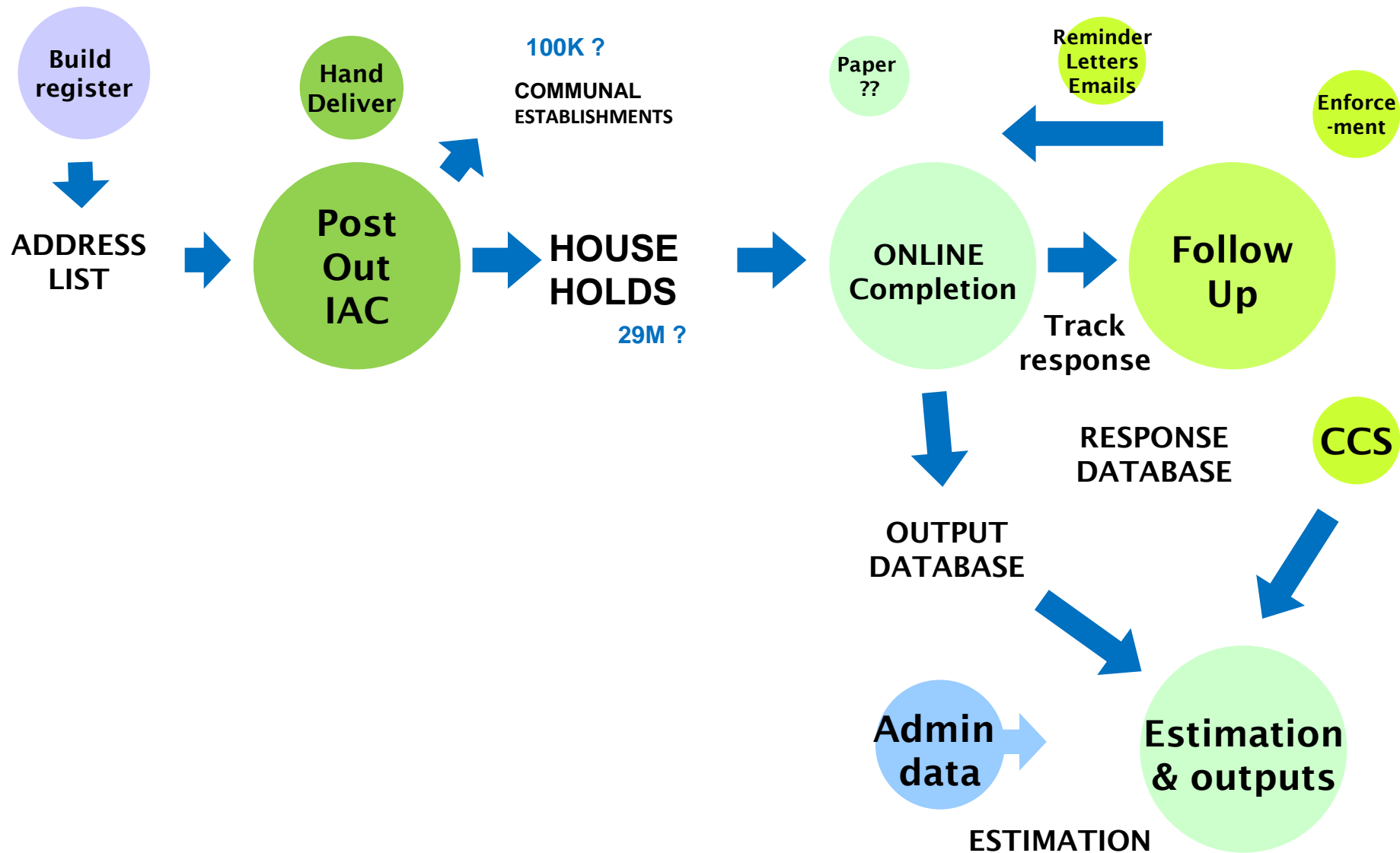
Head of Addressing & Census Methodology

Addresses

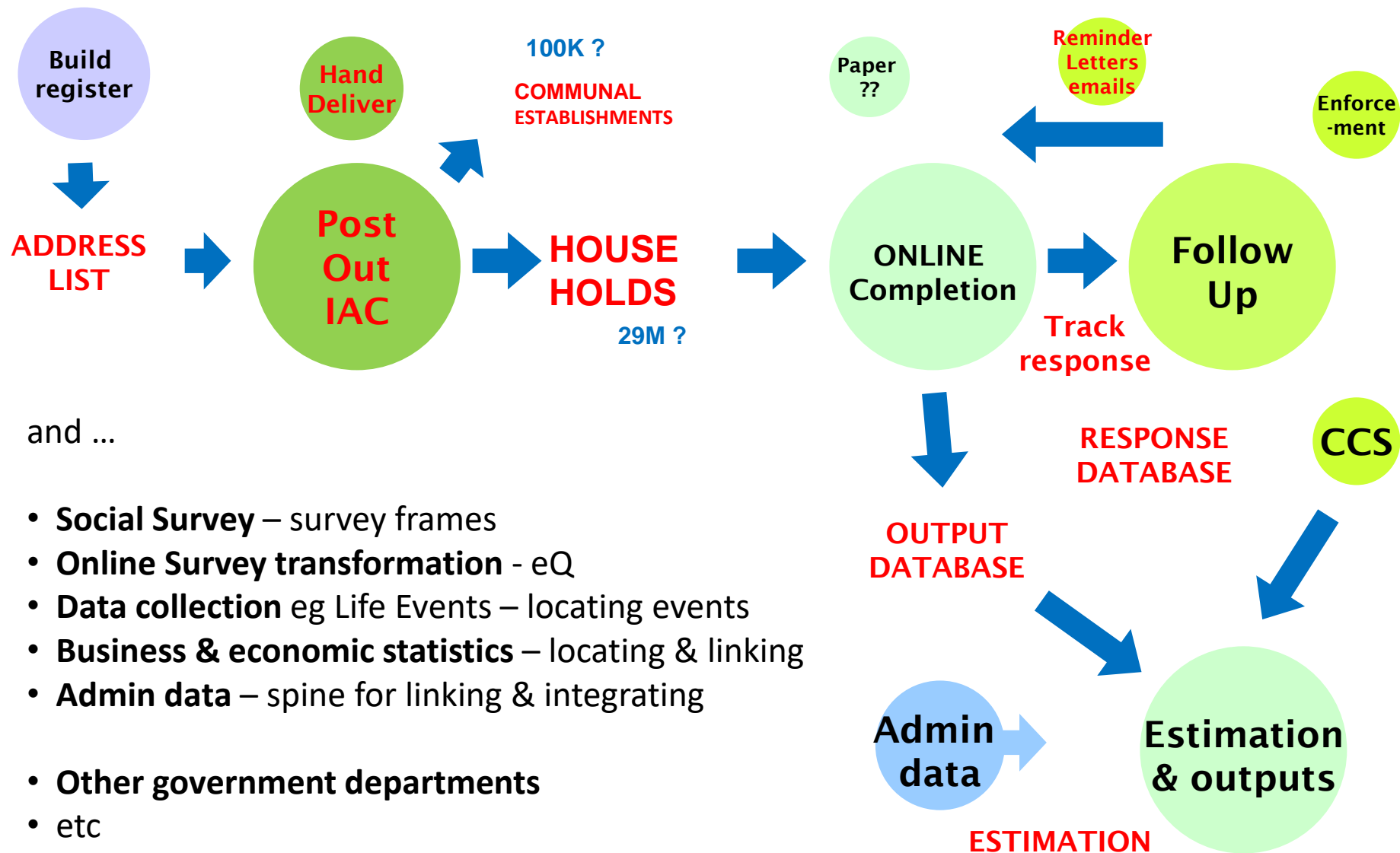
- What we need – and why it has now become easy
- Issues – and why it is still really hard
- Government addressing – it's AddressBase
- What it means for you – how can we work together ?

(& an annexe)

CENSUS OPERATION 2021



CENSUS OPERATION 2021

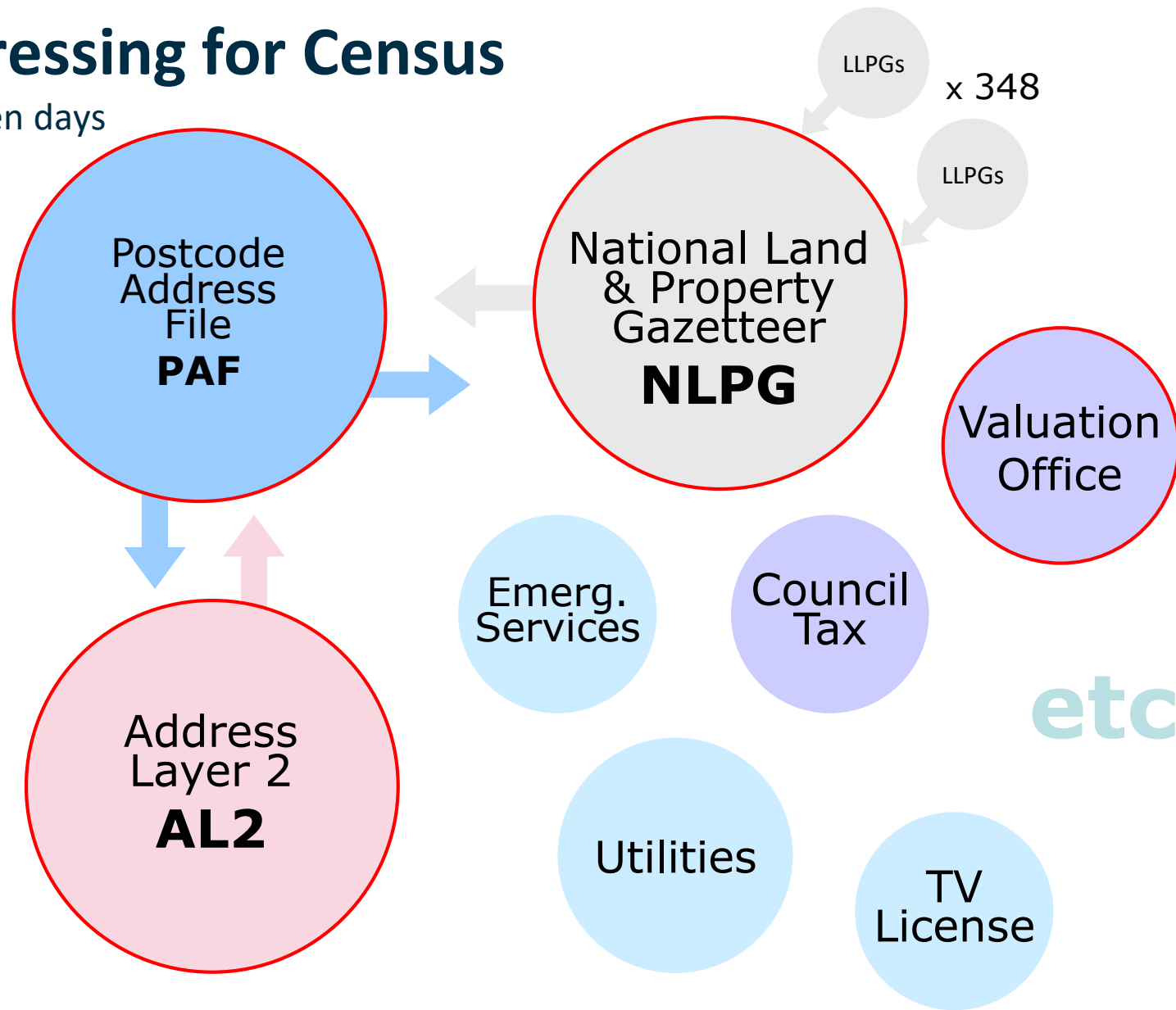


The requirement (tbc)

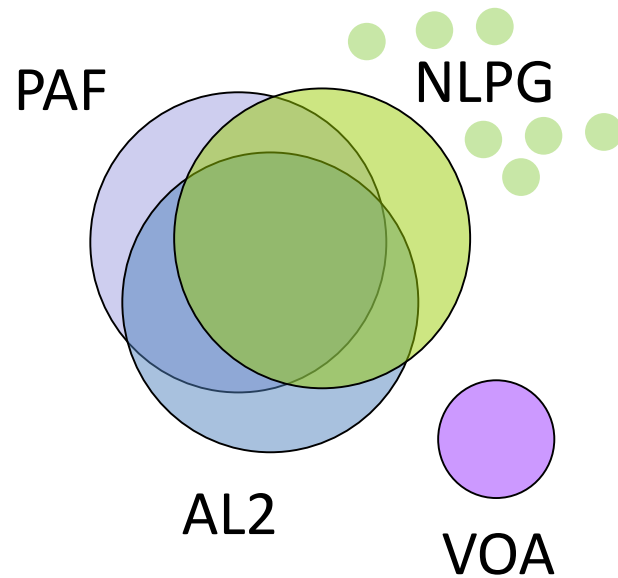
- A 'complete' household frame (complete)
>99% of household spaces (\neq addresses)
- Minimal over-coverage (properly classified)
duplicates / commercial / demolished etc < 2 or 3% ?
- A brilliant (integrated) communal frame (definitions ? ..)
- Residential, communal & business (& non postal)
- Up to date, correctly located etc etc And more

Addressing for Census

The olden days



Addressing – Census 2011

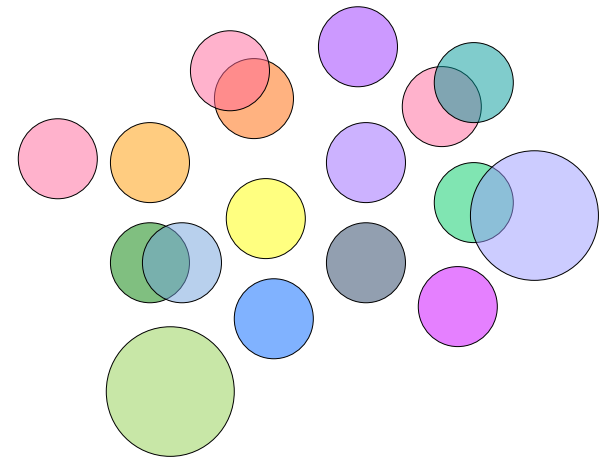


RULES



Residential address list

Very good



CLERICAL

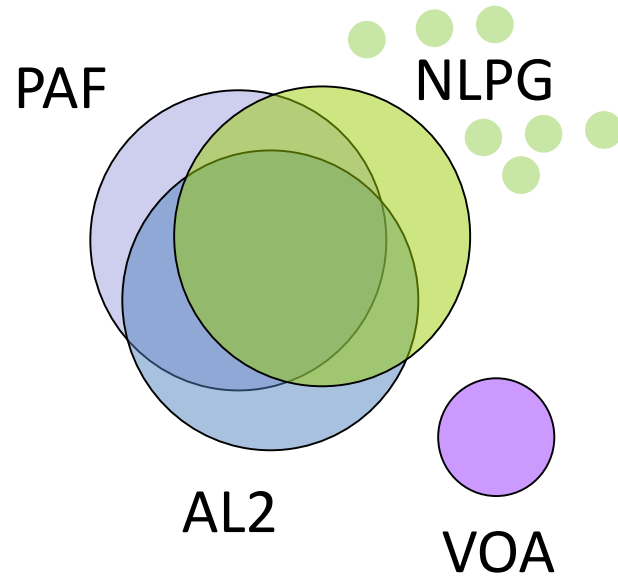


Communal address list

Just good enough

Why it's going
to be easy
this time

2011 Census

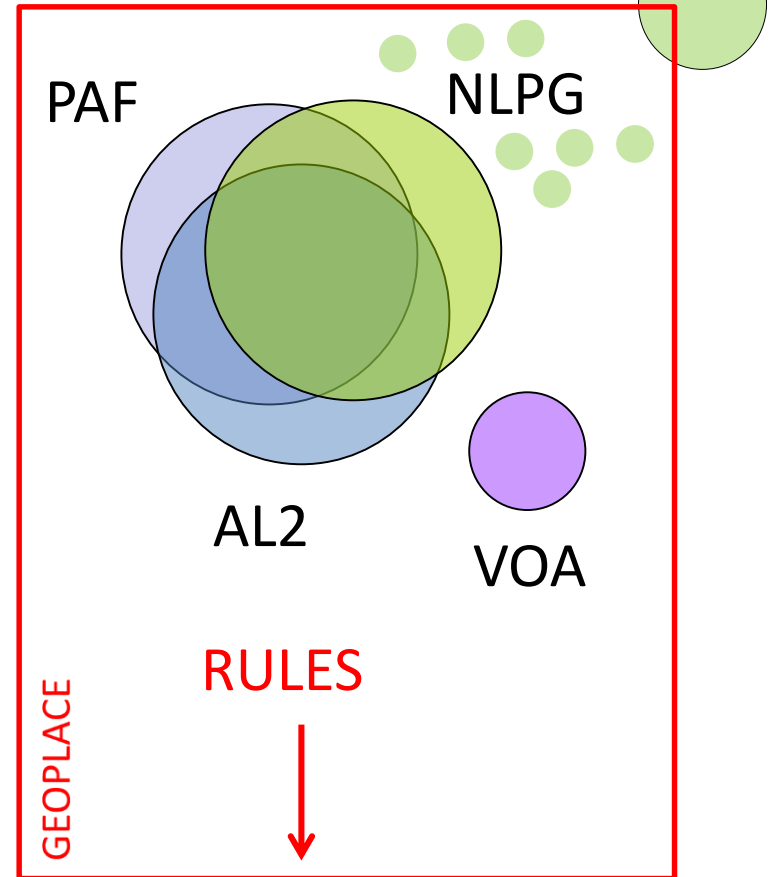


RULES



Residential address list

AddressBase



RULES



Residential address list

Communal address list

Why it's still
really hard....

The challenge Why it's hard this time

- We have an excellent starting point but addresses are complicated and change a lot. There will be error & error clusters itself in the areas we care about the most – **Very difficult to check quality**
- Extracting the right ones is difficult. Small errors can be significant – and cause trauma
- Communals are important and particularly challenging
- We plan to do MUCH more with addresses than post-out – huge opportunity but attribute thinking is new
- Addresses are complex so matching is really hard

CLARENDON
HOUSE

FLATS 1-6

32-33
HAMPSHIRE
TERRACE

CLARENDON
HOUSE

FLATS 1-6

32-33
HAMPSHIRE
TERRACE



SIR

FL Flat 1

FL Flat 2

FL Flat 3

FL Flat 4

FL Flat 5

FL Flat 6

FL Flat 7



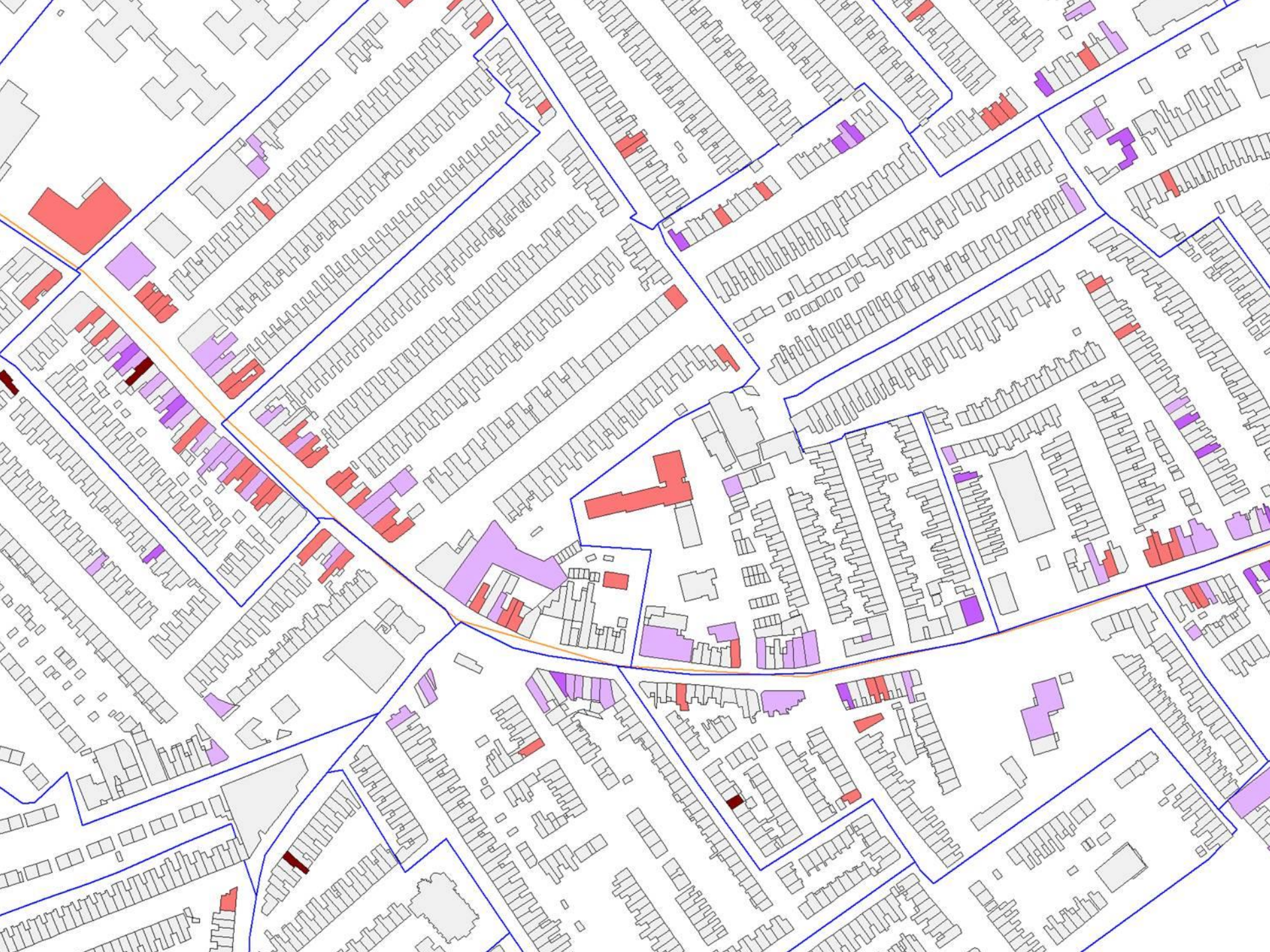
CLARENDON
HOUSE

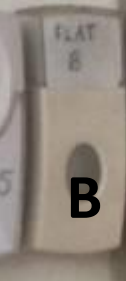
FLATS 1-~~6~~⁷

32-33
HAMPSHIRE
TERRACE









10

The challenge Why it's hard this time

- We have an excellent starting point but addresses are complicated and change a lot. There will be error & error clusters itself in the areas we care about the most – Very difficult to check quality
- Extracting the right ones is difficult. Small errors can be significant – and cause **trauma**
- Communals are important and particularly challenging
- We plan to do MUCH more with the register than post-out – huge opportunity but attribute thinking is new
- Addresses are complex so matching is really hard

DAILY ECHO

DAILY NEWSPAPER OF THE YEAR

EDF ENERGY
London & South of
England Media
Awards 2013



When news happens text pix and video to 80360. Start your message with BE then leave a space.

Search »

[News](#)
[Sport](#)
[Leisure](#)
[Video](#)
[Photos](#)
[Info](#)
[Your Say](#)
[Announcements](#)
[Jobs](#)
[Homes](#)
[Cars](#)
[Dating](#)
[Click2Find](#)
[Buy&Sell](#)
[Advertise](#)

[Districts](#)
[Education](#)
[Drink drive campaign](#)
[Crime and Court](#)
[Blogs](#)
[Opinion](#)
[Health](#)
[Features](#)
[Air Festival](#)
[Events](#)
[In Depth](#)
[Strange](#)

Echo » News »

NEWS



Send your news, pictures & videos

No common census as Moors Valley ticket machine gets form

By Harriet Marsh

10:30am Wednesday 9th March 2011 in News



Where's Wally? Maybe he's the "occupier" at the Moors Valley Car Park Hut



Email



Print this page



Comment(7)

CLICK2FIND

C T MOTORCYCLES

View details »

More Motorcycles

Pete
Spacagna

View details »

More Car Body Repair

20.0°C

0% chance of rain

weather forecast

FEATURED JOBS

HOUSEKEEPING STAFF Wareham
RESTAURANT STAFF Wareham
HOTEL RECEPTIONIST Wareham
Learning Support Assistants BOURNEMOUTH
Mac Operator - Page Layout Artist Ringwood
YARD PERSON/ VAN DRIVER Swanage

See 1,829 jobs in and around Bournemouth

MOST READ

1. OAP couple living on the streets after dispute over pension payments
2. Our grandson died after accidentally drinking drug-spiked Lucozade – we cry every day that we have lost him
3. Hit-and-run victim has bike stolen following crash
4. Poole Town legend Taffy Richardson: I wouldn't swap

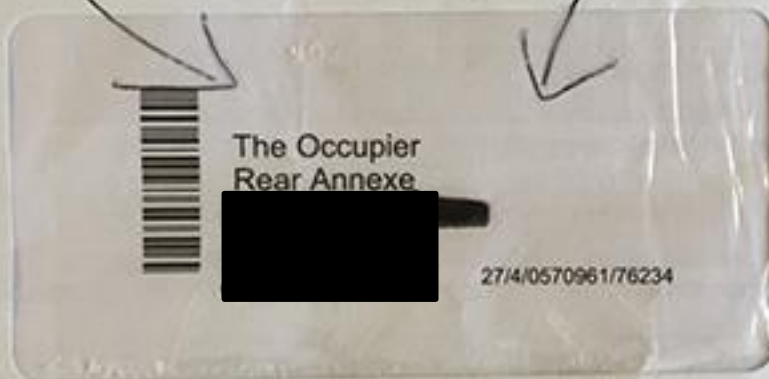




You've just wasted
my money on stamps
your [redacted]



RETURN TO THE IDIOTS



Stop wasting my
Council TAX. Let
me make it easy
4 you There is NO
NO NO NO NO NO
[redacted]

You Already sent
a foot soldier he
could not find a Rear Annexe
So be told

look on
Map

Rear Annexe
your [redacted]
you [redacted]

Census 2011
PO Box 658
FAREHAM
PO14 9LJ

NO REAR ANNEXE TO BE
Found however if you do find
it let me know. So we can
let it out and earn money, money
money.

Your Unbelievable IDIOTS.
I can't believe you've been designated
this job. Google my [REDACTED]
house

ENV CRL 01

The challenge Why it's hard this time

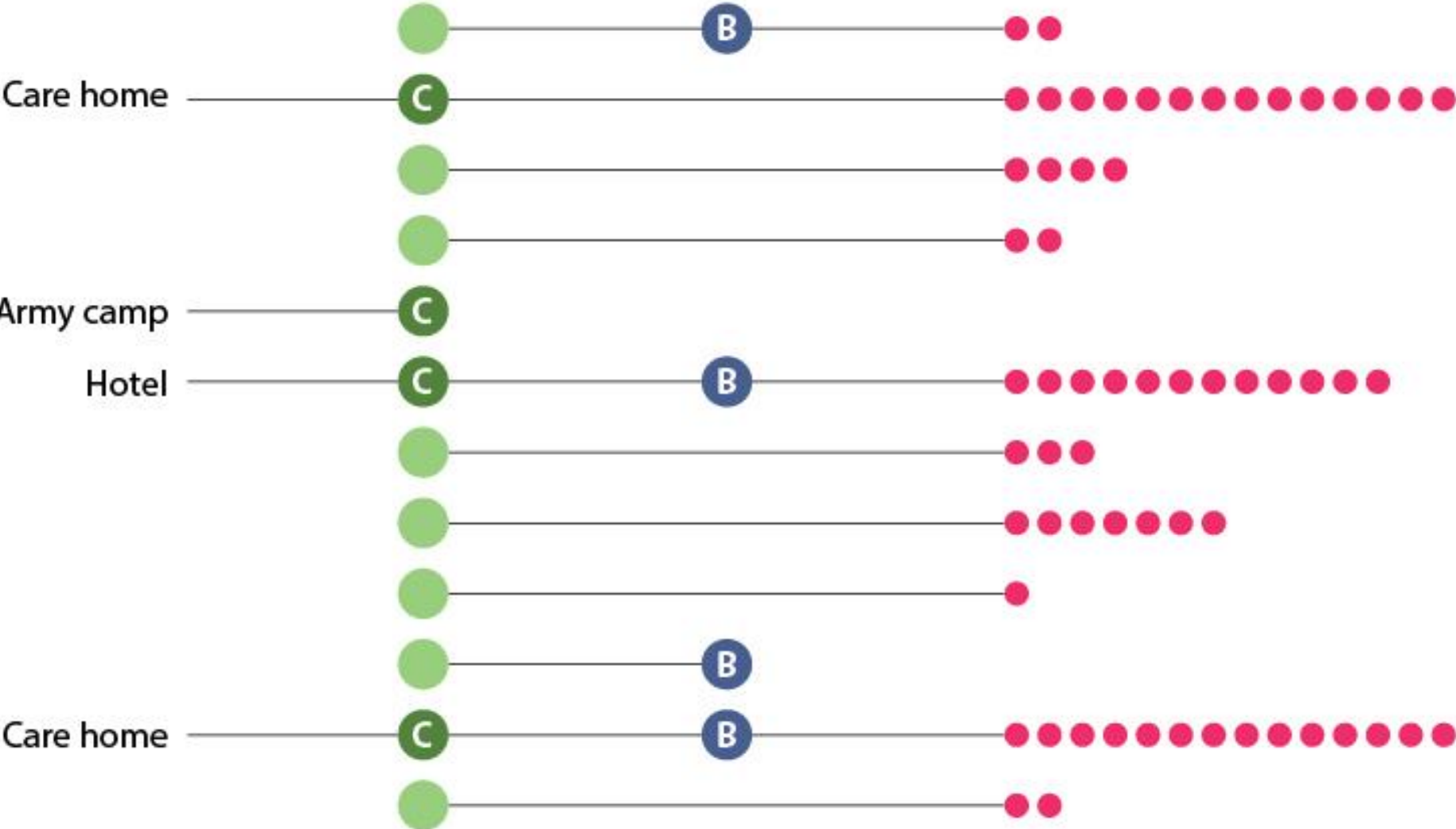
- We have an excellent starting point but addresses are complicated and change a lot. There will be error & error clusters itself in the areas we care about the most – Very difficult to check quality
- Extracting the right ones is difficult. Small errors can be significant – and cause trauma
- **Communals** are important and particularly challenging
- We plan to do MUCH more with addresses than post-out – huge opportunity but attribute thinking is new
- Addresses are complex so matching is really hard

Lists of
communals

Linked to
Address
Index

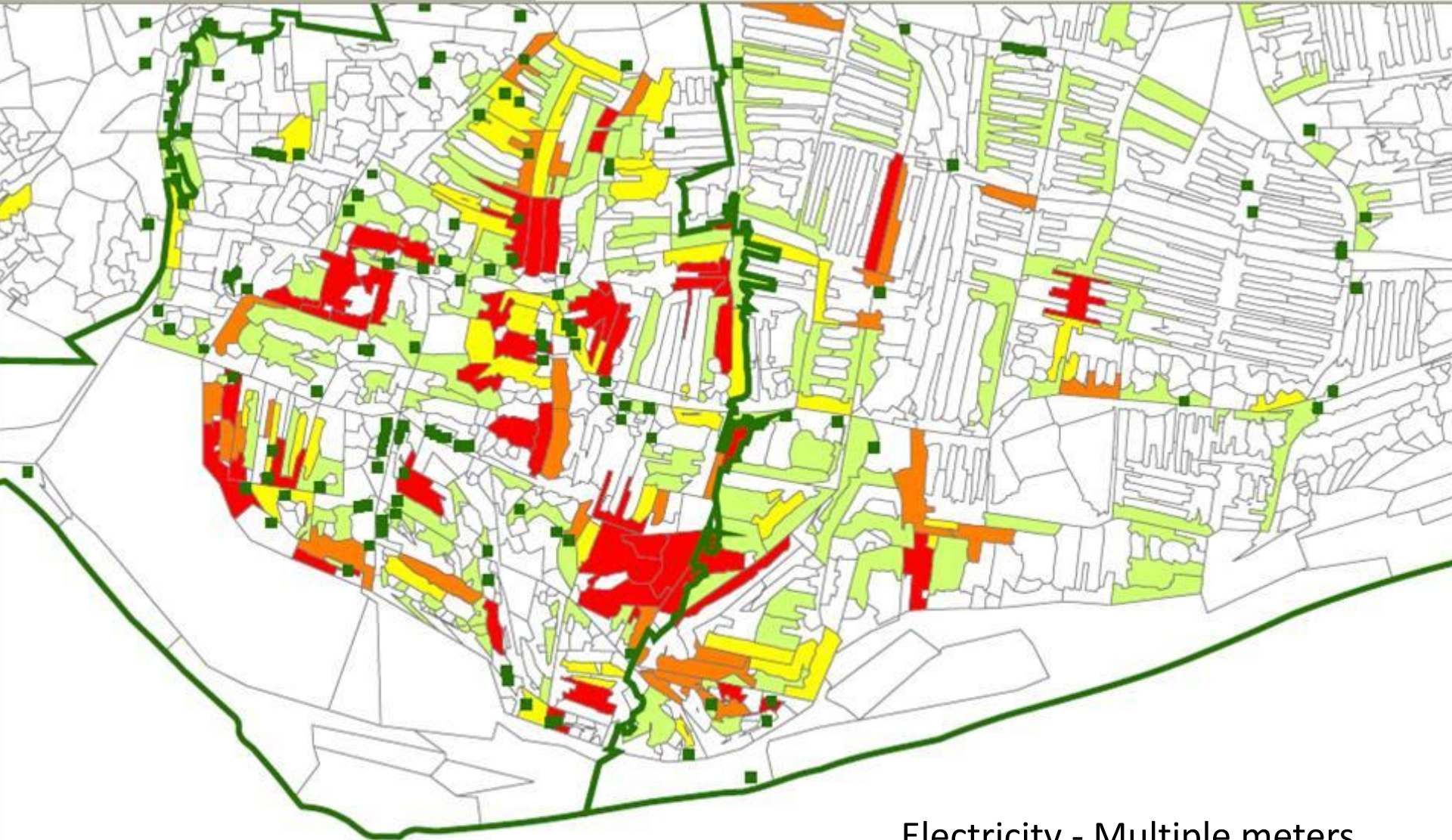
Linked to
Business
Index

Compared to
counts from
admin data



The challenge Why it's hard this time

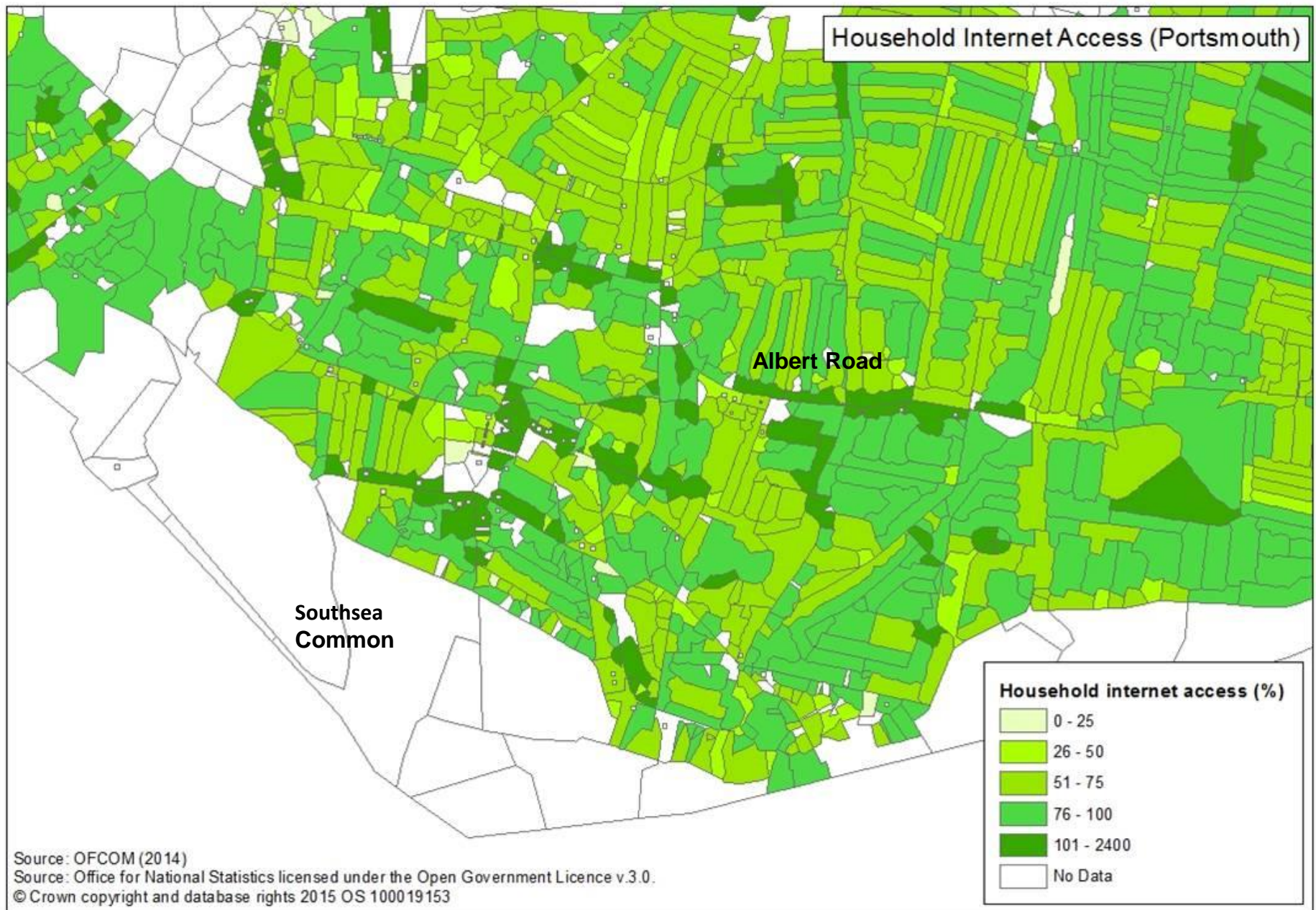
- We have an excellent starting point but addresses are complicated and change a lot. There will be error & error clusters itself in the areas we care about the most – Very difficult to check quality
- Extracting the right ones is difficult. Small errors can be significant – and cause trauma
- Communals are important and particularly challenging
- We plan to do MUCH more with addresses than post-out – huge opportunity but **attribute** thinking is new
- Addresses are complex so matching is really hard



Electricity - Multiple meters



Electoral register - Multiple names



... also directly relevant topics such as internet access ... now updated opendata at postcode

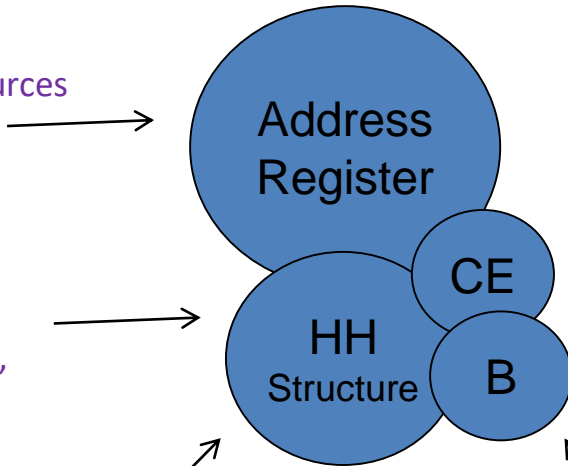
Conceptually – all subject to ethical and privacy discussion !

A probabilistic address frame

New definition / schema

Geoplace

And other CE sources



Admin data

HH structure, churn,
names, house
prices, phone
numbers

Activity data

Energy, utilities,
broadband, health,
house sales

Business Reg

Business structure,
type, churn

2011

Census

HH structure,
churn, names

Potentially

Other

Shape / pattern
recognition
Survey paradata

LOCAL
INTELLIGENCE

Probability of

- Existence of address
- type - HH/B/CE
- HH Size / structure
- Change / churn
- Hard to countness / category
- (multivariate >> categorisation
- Eg possible holiday home, carehome, student accommodation

Inform field planning / targetting
Intelligent stratification
Prioritise follow up (address level)
Inform estimation & modelling

The challenge Why it's hard this time

- We have an excellent starting point but addresses are complicated and change a lot. There will be error & error clusters itself in the areas we care about the most – Very difficult to check quality
- Extracting the right ones is difficult. Small errors can be significant – and cause trauma
- Communals are important and particularly challenging
- We plan to do MUCH more with addresses than post-out – huge opportunity but attribute thinking is new
- Addresses are complex so **matching is really hard**
-

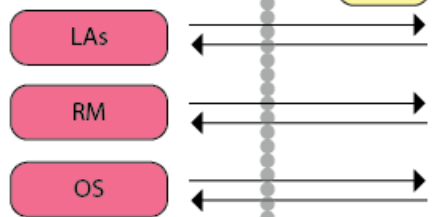
The bigger picture

– addresses as a linking
mechanism across government

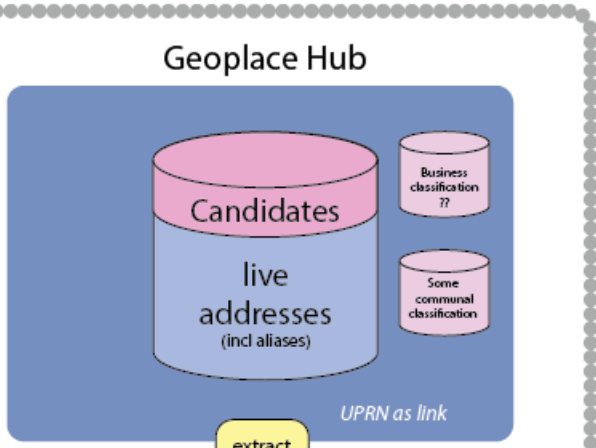
National Addressing ?

v 1.1 AC, MJ ONS 16/6/15

Address data suppliers have live open access to Hub
Update candidate addresses
progress information etc



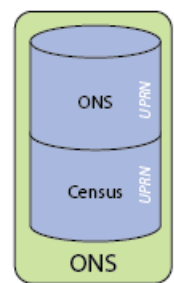
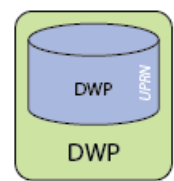
geoplace functions & clerical



CLOUD

Geoplace Hub acts as staging for candidate addresses, progress, QA etc
Geoplace & Govt Hub may be the same thing with different permissions / security
Assume Geoplace hub is also on cloud

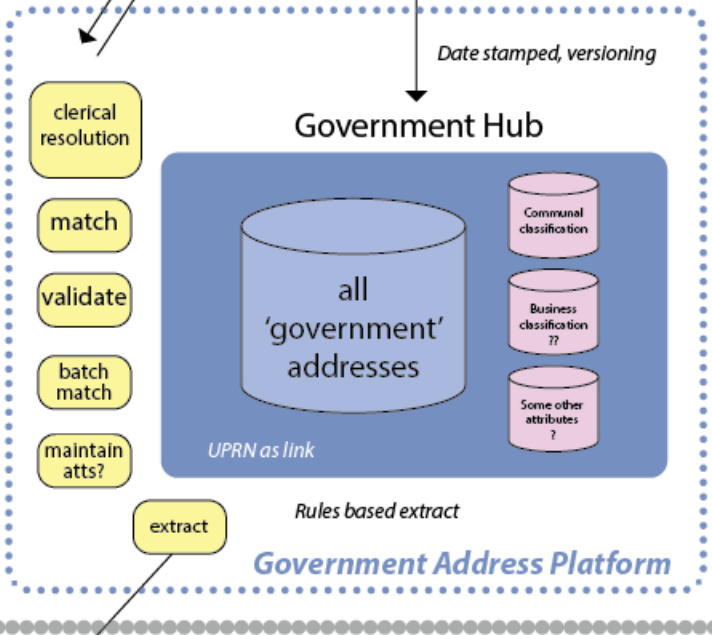
Departmental data



Individual address lookup or verify or batch match of addresses

In every case the return is UPRN

Feedback of error, unmatched, candidates etc (law, ethics?)

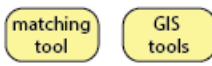


Daily rules based extract

Date stamped, versioning

Verify
Land Registry
Businesses ??
Buildings ???
How does this all work ??
HOW OPEN ??
Commercial view on data - free
Citizen view on data - free
Public verification ? / Wiki
UK ?

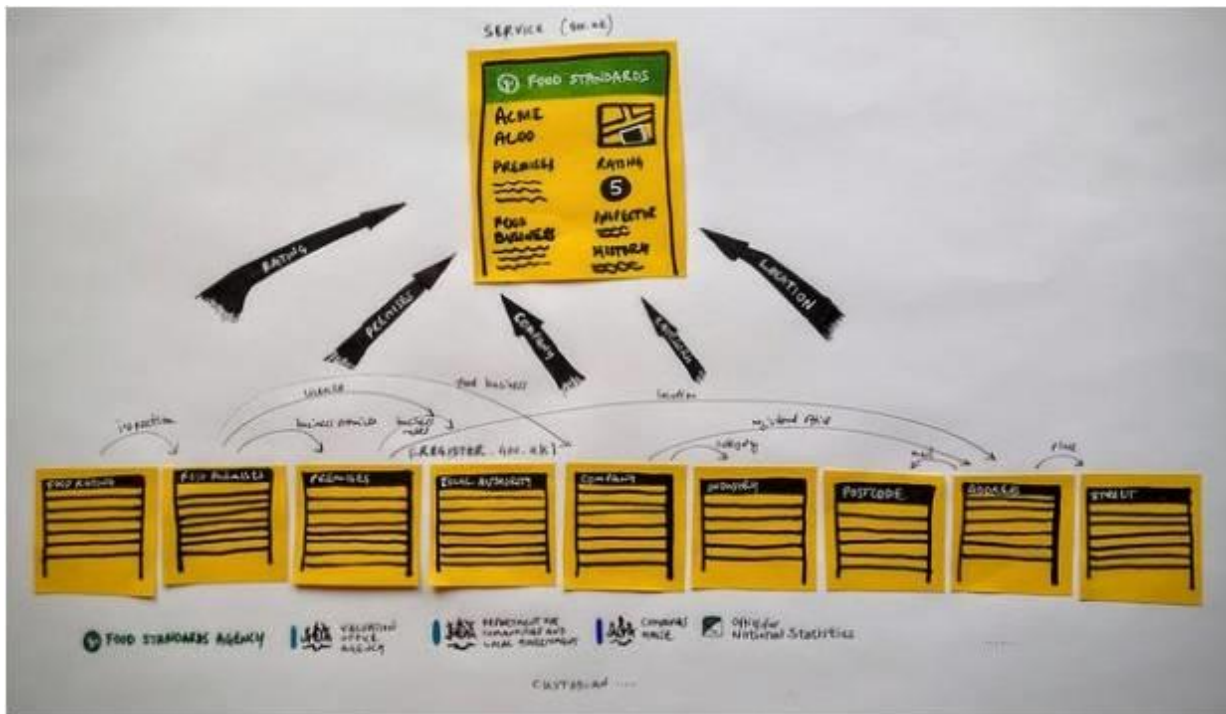
May be value in distributed..



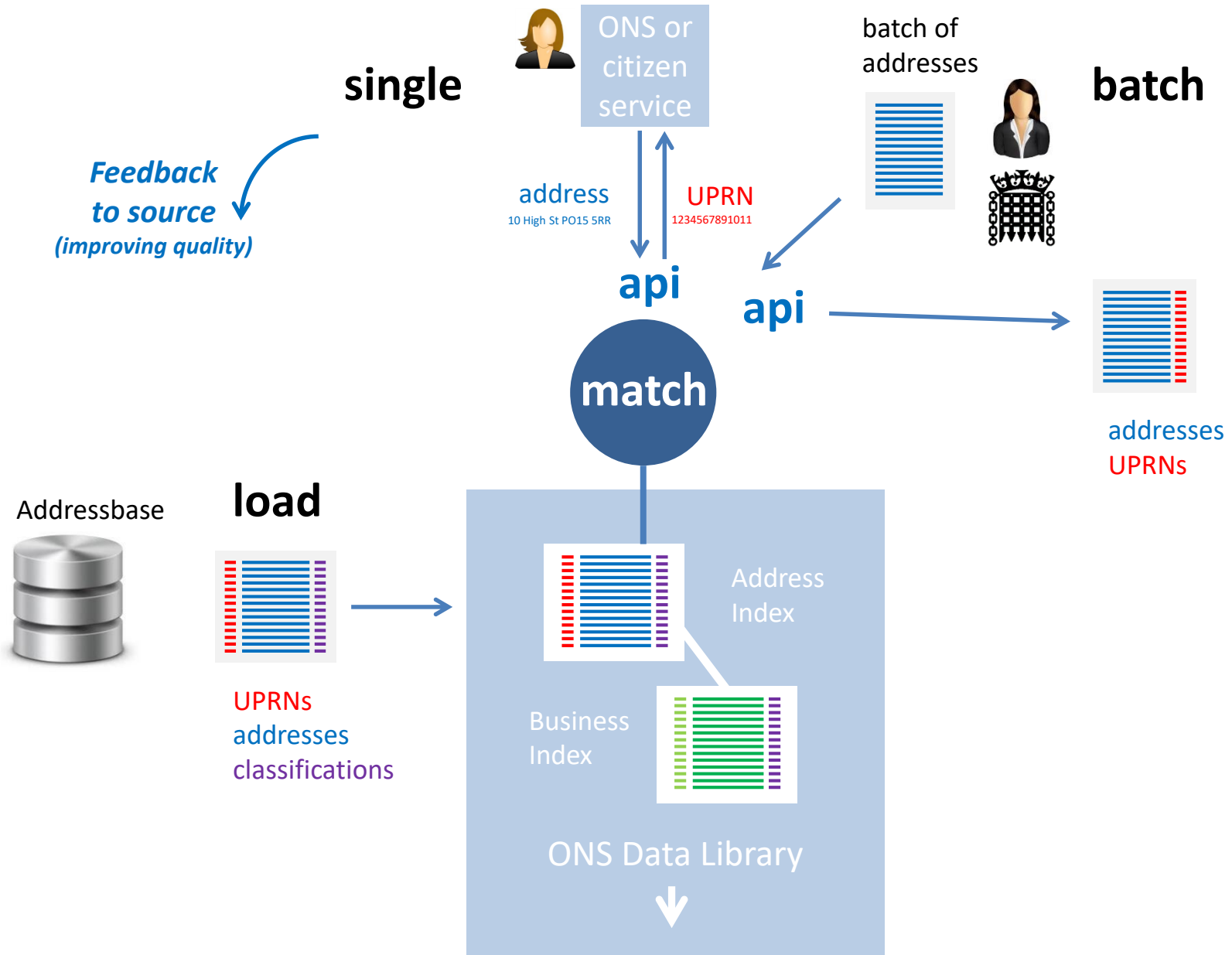
Sample
Business or Social Survey Frame
Census Frame etc

To be confirmed:
Roles & responsibilities
Security & environments
WHAT CAN BE OPEN ?
How integrated is the business register
Communal interests
Operational vs 'Statistical' purposes

Government Digital Service (GDS) Vision for Registers



Address Matching - Beta



Searching and matching – **what we want**

correct **match rate**

virtually **zero false positives**

balance between **automatic & clerical**

flexibility of match tuned but not limited

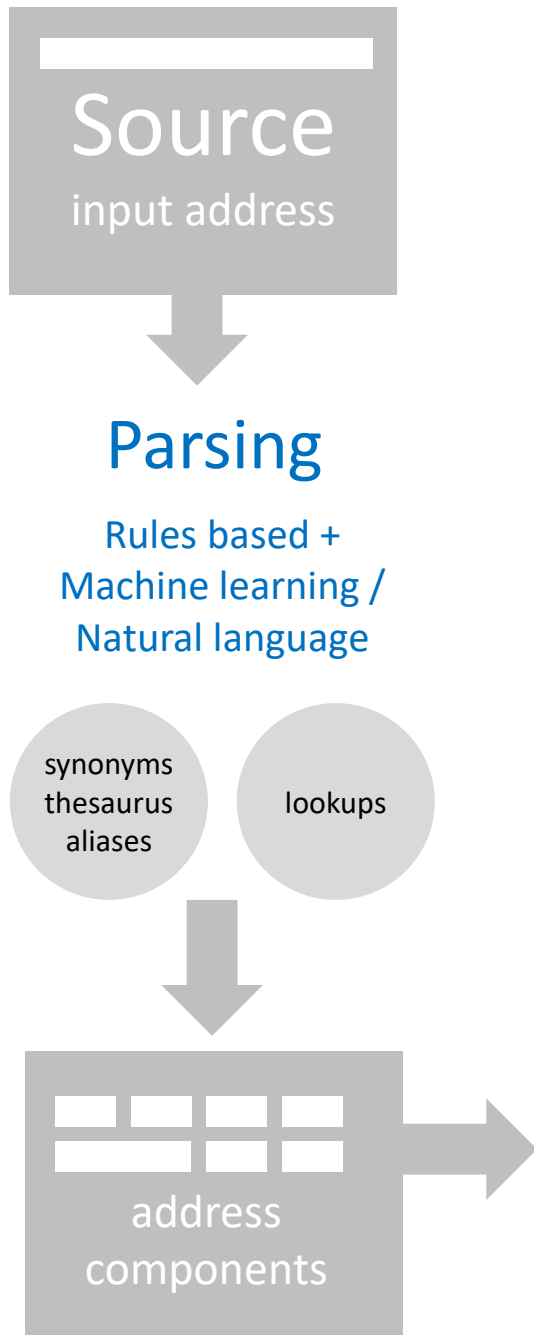
fast

scalable

accessible via **api**

non proprietary code -> **open**

how we are going to do it



Avenue Cars Limited
1st Floor
St. William of York House
22-24 First Road,
Street, Somerset
ZE10DW

Source

input address

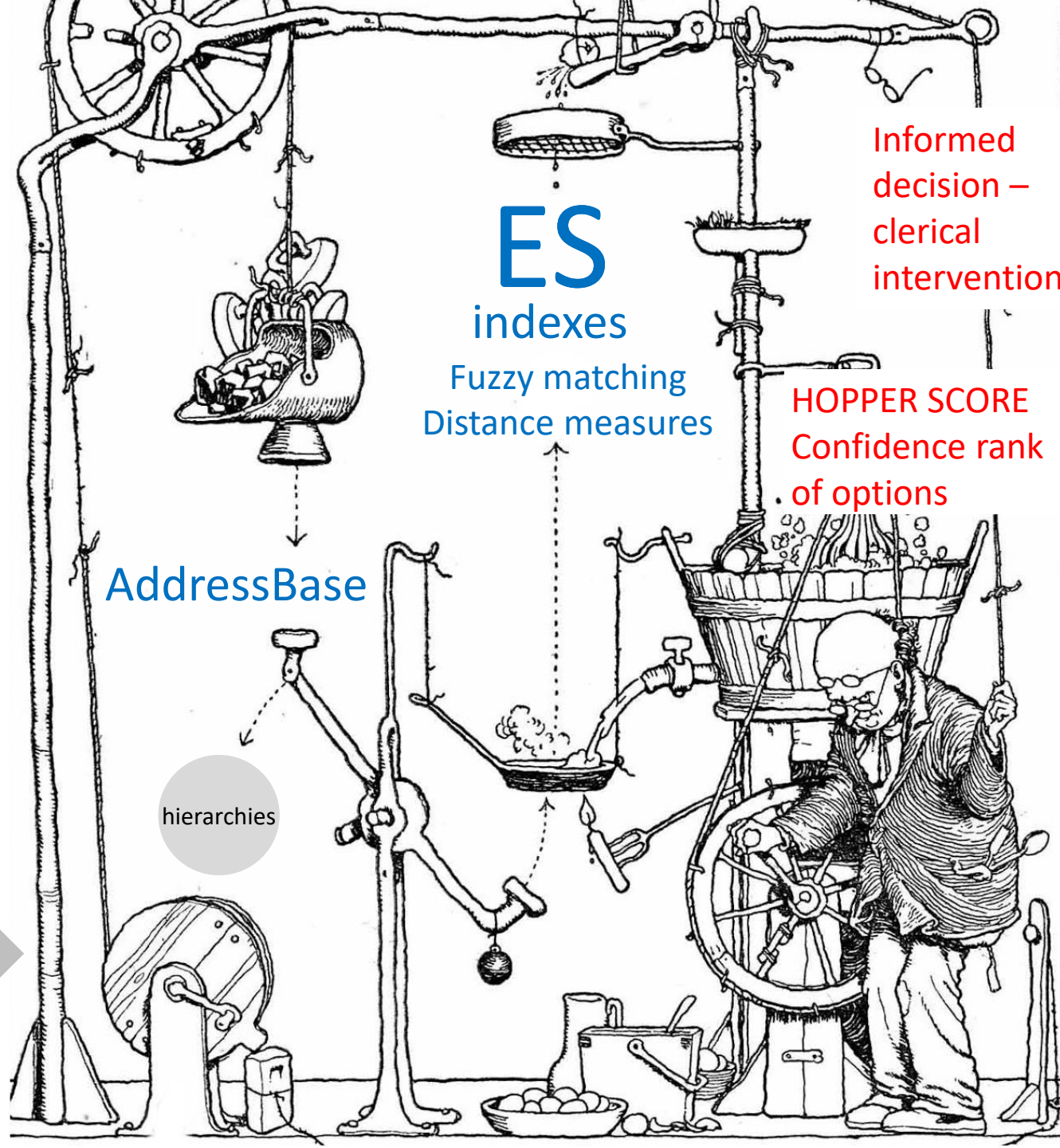
Parsing

Rules based +
Machine learning /
Natural language

synonyms
thesaurus
aliases

lookups


address
components



Summary

(and some conclusions)

Summary

- AddressBase at the core – need to confirm & ensure quality
- No separate national address register (except temp / operational)
 - it is all about improving the national source
 - increased use of source >>> linking >>> feedback to improve the national hub
- Relationship with you is absolutely fundamental
 - Want to agree approach and share lists much earlier than before
 - Coding of LLPGs the key – we want to extract direct from AB
 - **Will be writing out soon - regional events next year** – but want to work out a plan with your help & support
 - Support to the case for funding local addressing
- How do you want to work together ?????
- Matching Service Talking to GDS, OS, HMRC, BEIS, DWP ... etc. LAs ??
-  Love to share and talk about addresses and matching

Questions?

(& come and talk to us)

addresses@ons.gov.uk