



# SMART DATA LAKE

Richard Holland – London Borough of Waltham Forest  
*February 2019*



# Agenda

Who are AWS?

Smart Data Lake – Waltham Forest

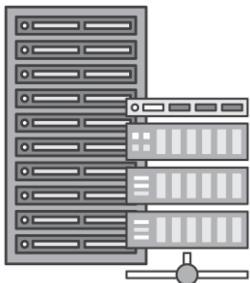
AWS Partner NGDATA – Enabling Resident DNA



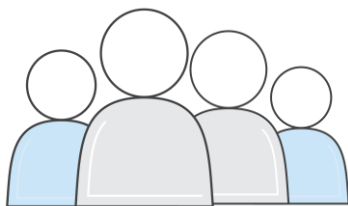




## On Premise



Equipment



Resources and  
Administration



Contracts



Cost

## AWS Cloud



No Up Front Expense  
Pay for what you Use



Agility



Scale Up and  
Down



Self-Service  
Infrastructure



## Who We Work With



**NETFLIX**



**UCAS**



**JustGiving™**

# LOCATION



# SUMMARY

- Waltham Forest is a predominantly **residential borough** with one of the smallest economies in London (9% 2016 vs. 6% in London). Self-employment rate higher than London (16.5% vs. 13.6%)
- **Strong business growth** in recent years but less impact on job growth as **mainly small businesses**
- **High population growth** throughout last decade driven by international migration (mainly from the EU) 275,800 in 2016
- **Young age** structure with more children and working-age residents compared to the UK average. Over 65,000 children and young people (about a quarter of population)
- Increasingly **diverse population** without a single majority group
- **House prices have rocketed** since the recession and the borough is among the ten local areas in the country where house price/earnings ratio has increased most since 1999.
- **High population churn** with more people moving to other areas in the UK than other way around, particularly families with young children. 42% of outflows are to other London boroughs. 74% of inflows are from other London boroughs, mainly neighbouring boroughs.
- Population is projected to continue to grow and is **increasingly ageing**
- Significant uncertainty around migration and the wider impacts of **Brexit**.
- Ranked as 35th most deprived local authority in England (out of 326) in 2015

# WALTHAM FOREST SERVICES

We have 2800 officers and 60 Councillors running the following services for 280k residents.

- Rubbish and Recycling
- Planning & Building Control
- Special Education Needs and Transport
- Adult Social Care
- Housing
- Enforcement and Licensing
- Arts, Part and Libraries
- Benefits and Money Advice
- Electoral Services
- Births, Deaths & Marriages
- Regeneration
- Children Services (Fostering, Adoption, Support)
- Telecare
- Commercial Trading Company (Pest Control, Weddings etc.)
- Highways & Parking
- CCTV
- Public Health

Our IT department consist of approximately 70 individuals supporting over 150 systems all in silos, more than 80 customer portals & websites, and one ChatBot. We have **loads** of data!

**BUT** It is very difficult to combine data together and produce reliable, real-time intelligence



# RESIDENTS AND STAFF EXPECT A CONNECTED EXPERIENCE

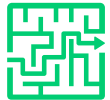


# SLOW & FRAGMENTED CUSTOMER DATA.

Most public-sector organisations struggle to offer exceptional experiences and to build long-term relationships with staff and residents due to:



DATA SILOS (IRRELEVANCE, REACTIVE, SEGMENT-BASED)



COMPLEX LEGACY ARCHITECTURES



LACK OF SHARED INTERNAL & EXTERNAL INSIGHT ACROSS THE ORGANISATION

# WHAT IS SELECTIVE LICENSING?

All houses within the private rented sector in an area that has been designated for selective licensing must be licensed

An area may be designated for selective licensing either

- (i) if the area is (or is likely to be) an area of low housing demand **or**
- (ii) The area is experiencing a significant and persistent problem caused by anti social behaviour and some or all of the private sector landlords are failing to take action to combat the problem that it will be appropriate for them to take

A designation can last for five years. It can be renewed

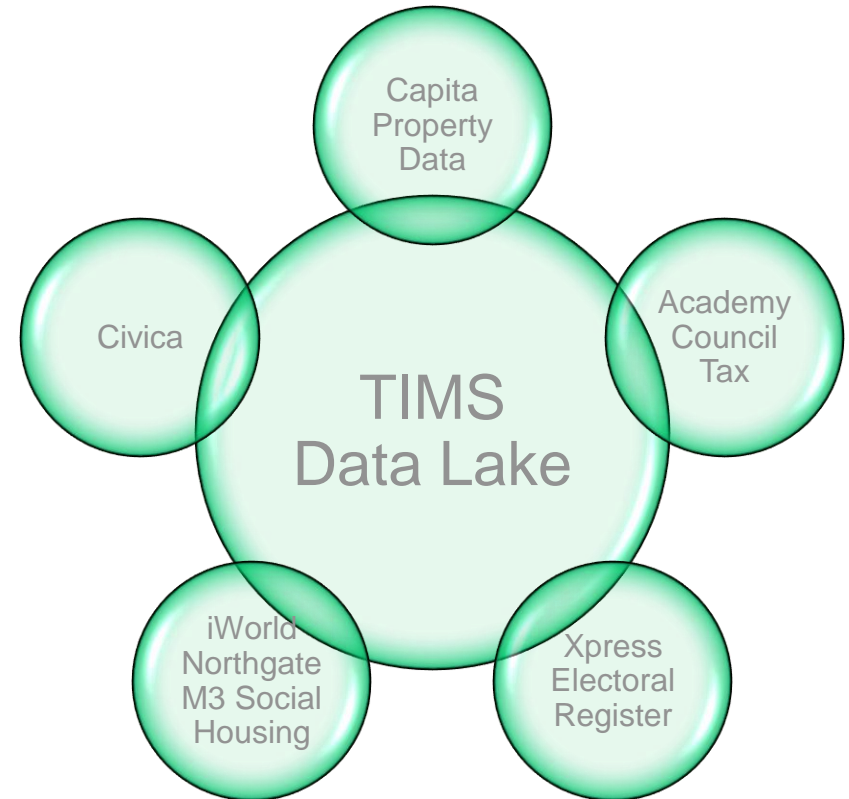


# WHY TENURE INTELLIGENCE MANAGEMENT?

- Targeted enforcement to alleviate demand on all enforcement services
  - e.g. Neighbourhood Services
- Reduced reputational risk though improved targeted warning letters and visits that avoid unnecessary contact with excluded tenures such as owner occupiers
- Build an evidence base to support the review and evaluation of the Property Licensing scheme (2019)
- Track and monitor problem behaviours (e.g. ASB, FPN, etc.) and guide enforcement and compliance
- Potential to support property level BAU enforcement and compliance activities including for example : Council Tax, Revenue and Benefits and Electoral Roll

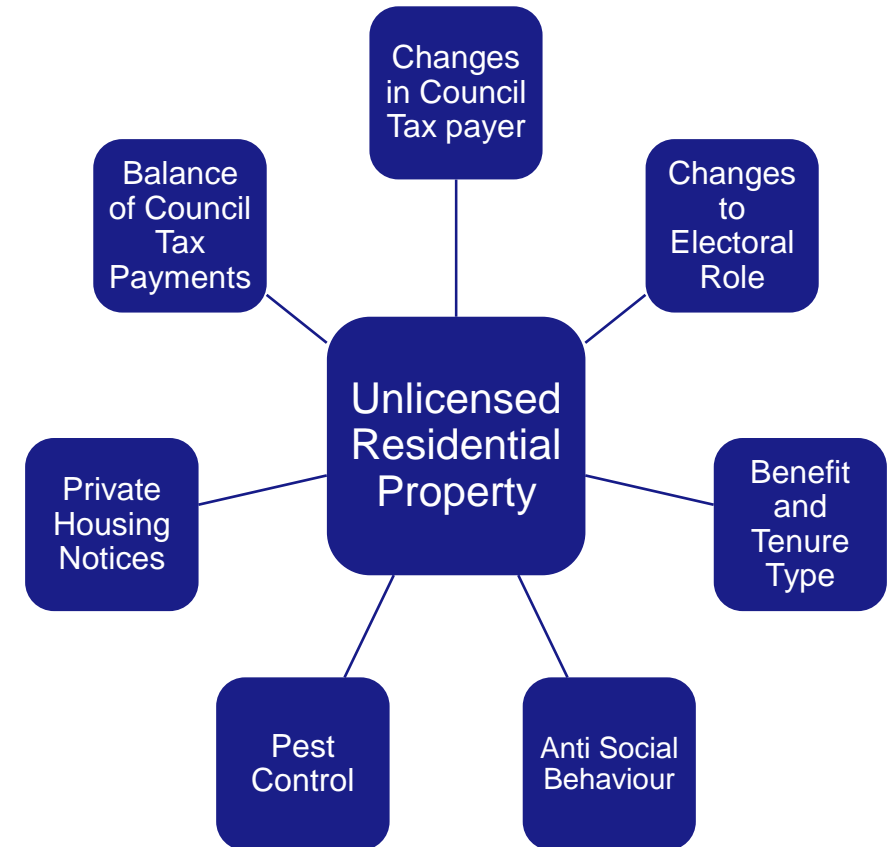
# HOW DOES SMART TIMS WORK?

- Build mappings and ingest data to Hive with Big Data Manager (“BDM”)
- Provision data in Informatica Enterprise Data Catalog (“EDC”)
- Profile and create sample Data Quality rules using Informatica Big Data Quality, show how this insight can be exposed and leveraged by Waltham Forest stakeholders via the Big Data, Data Quality software
- Combine the solution architecture, technical details, analytics, statistics etc. into a final presentation
- Provide high-level walkthrough of the data journey, ingestion using S3, reuse of components and APIs
- Overview and demonstration of Data Quality, creation of new business rule, running the plan and examination of the result
- Using a Predictive Algorithm tells us where the private landlords are and who is not performing.



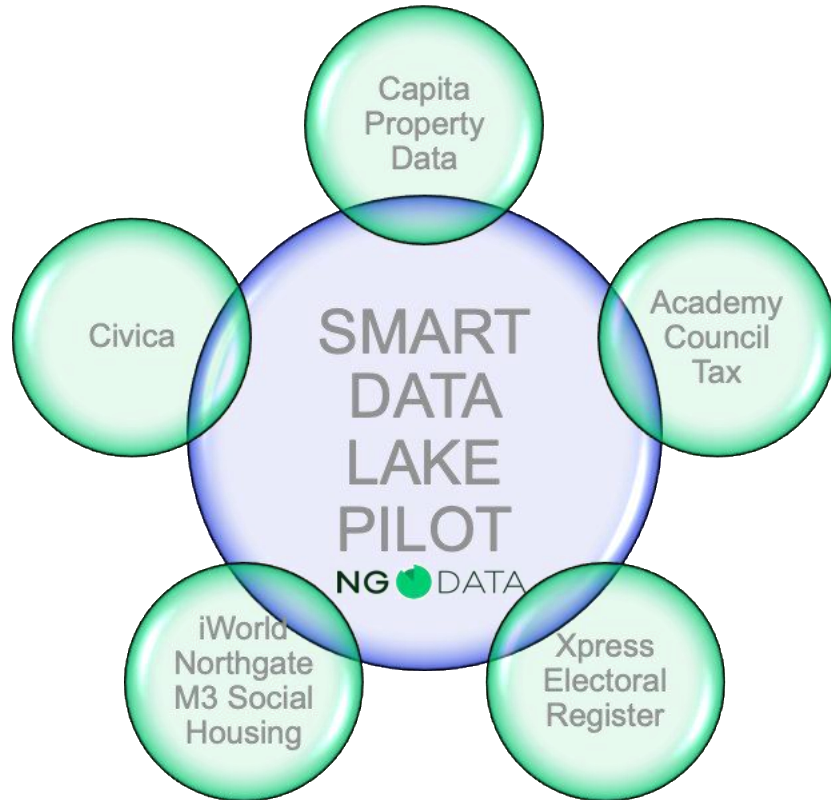
# PREDICTIVE TOOL

- Statistical regression used to 'weight' different variables.
- Regression model looks for links and changes between variables that may indicate unlicensed HMOs.
- Weighting transferred into algorithm to provide unlicensed property risk-rating (0-100).
- Measures relationship between variables and insight on cause/effect.
- The algorithm is embedded into the AWS and Informatica data lake environment and updated in real time.

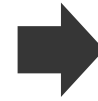




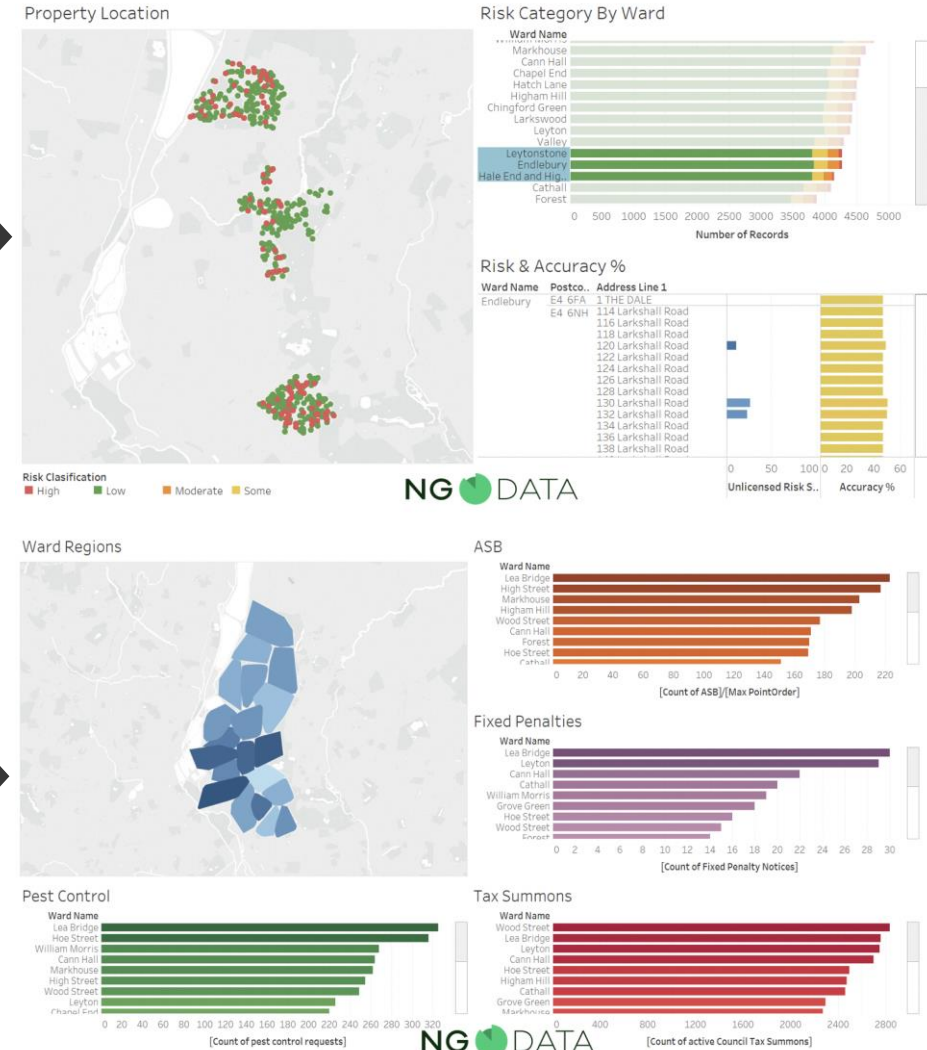
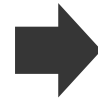
# PILOTING THE HUB APPROACH



Analytics Model



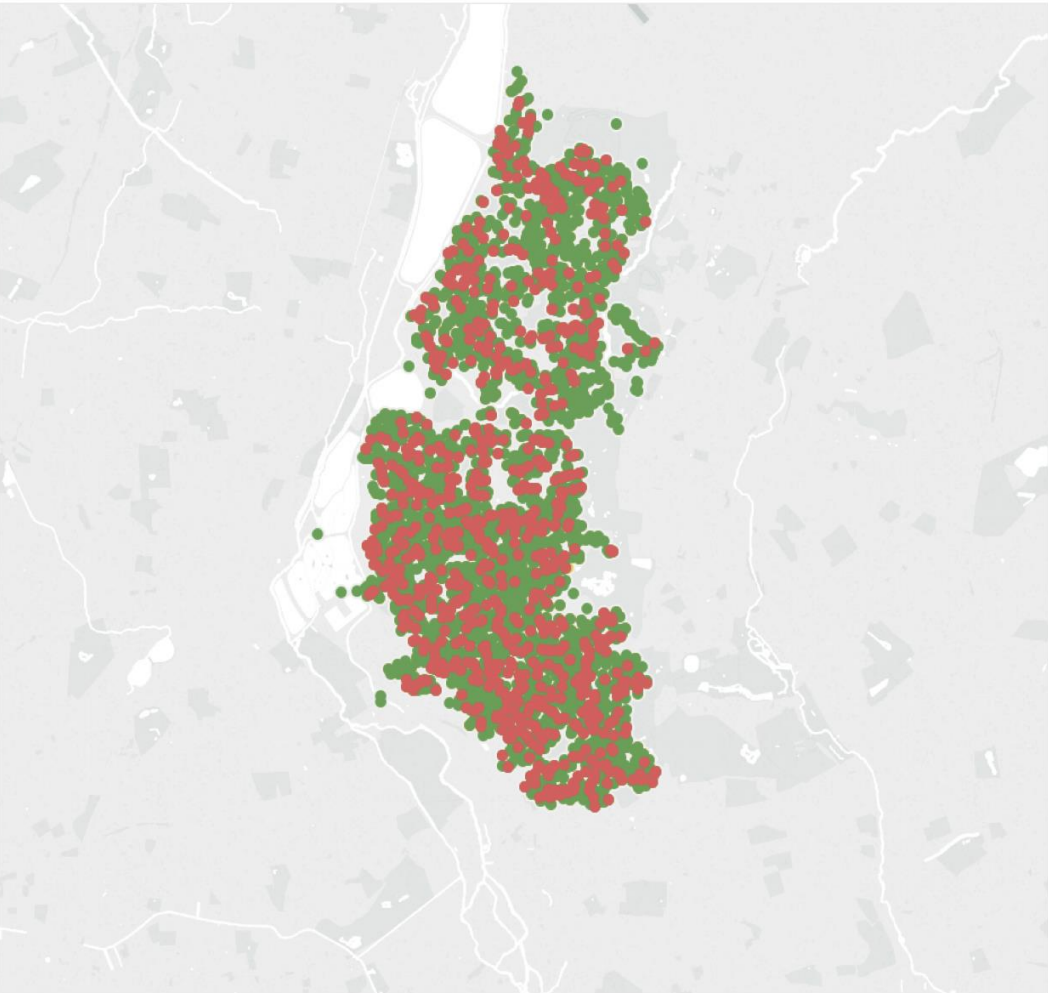
Reporting Extracts



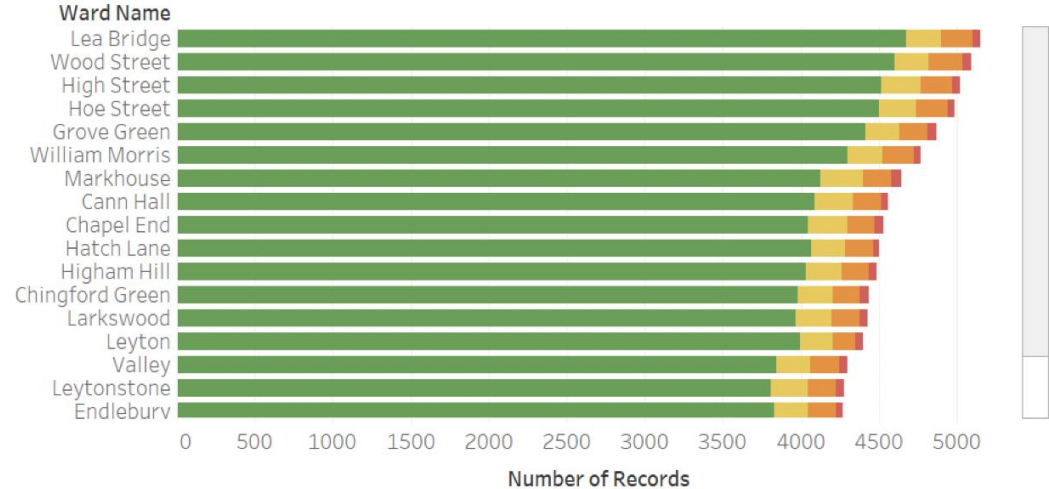
5 Weeks to integrate systems, understand quality, capture data, blend into new data sets, prototype models and build reports.

# RISK BASED PROFILE OF LANDLORDS

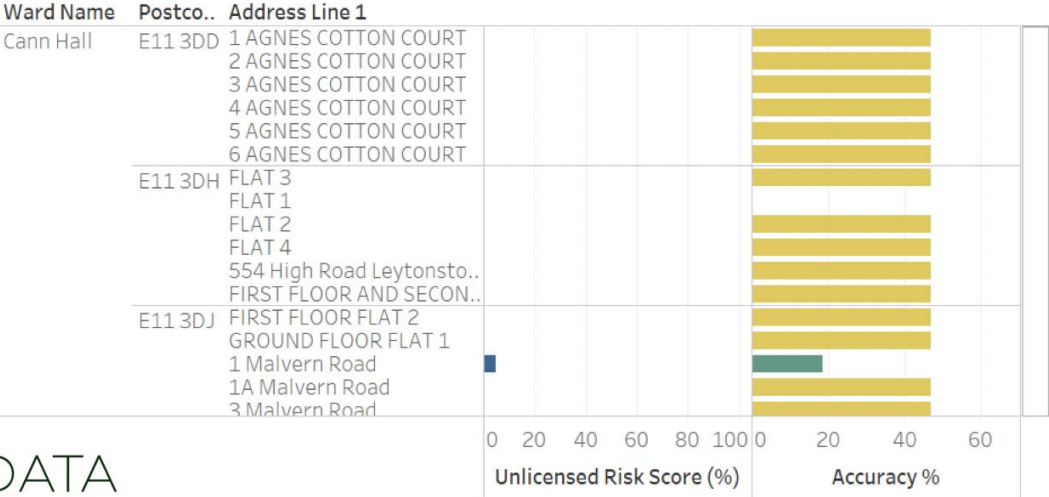
Property Location



Risk Category By Ward



Risk & Accuracy %

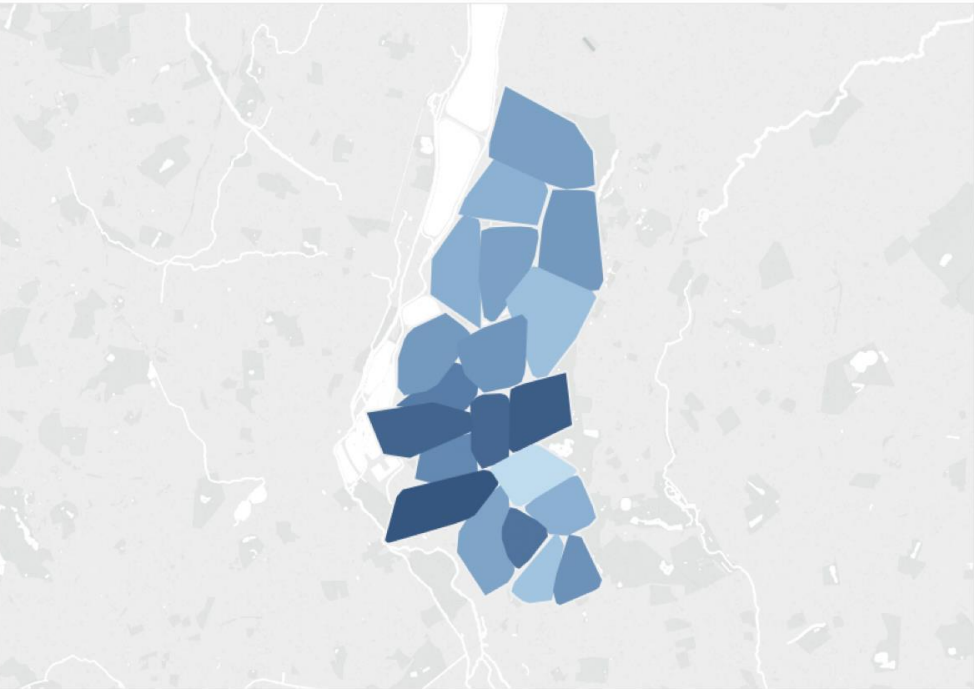


Risk Classification  
■ High ■ Low ■ Moderate ■ Some

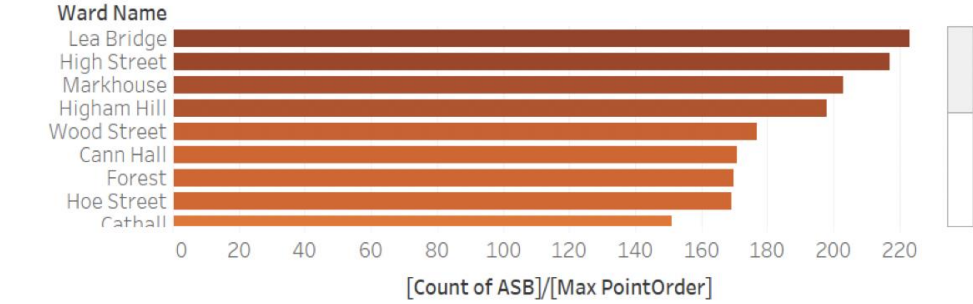
# TOTAL PER WARD



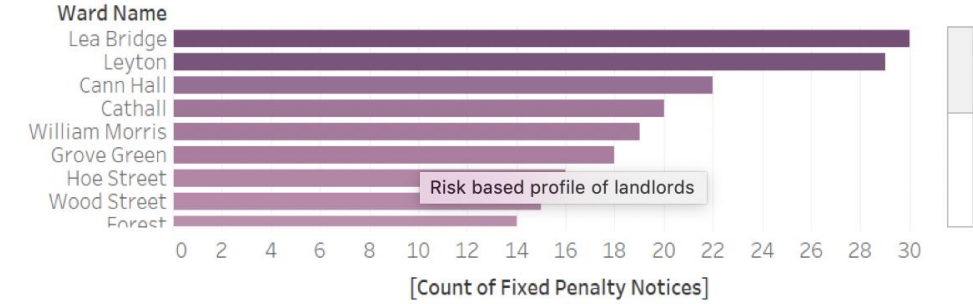
Ward Regions



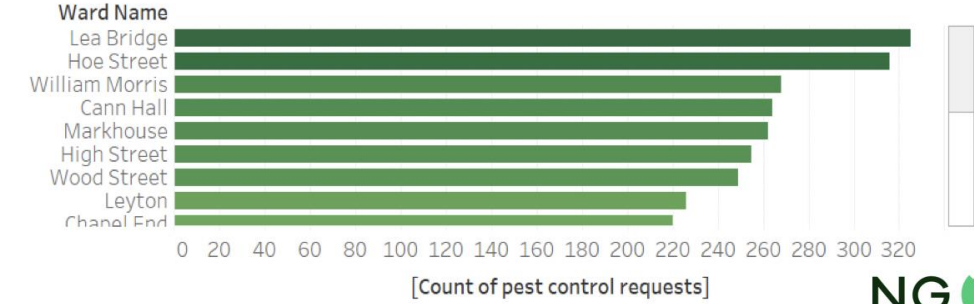
ASB



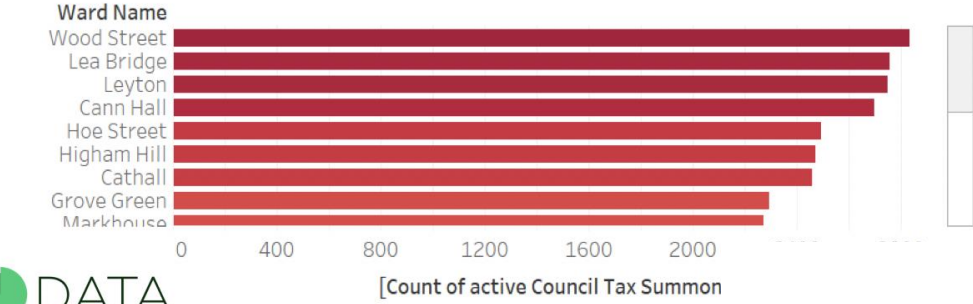
Fixed Penalties



Pest Control



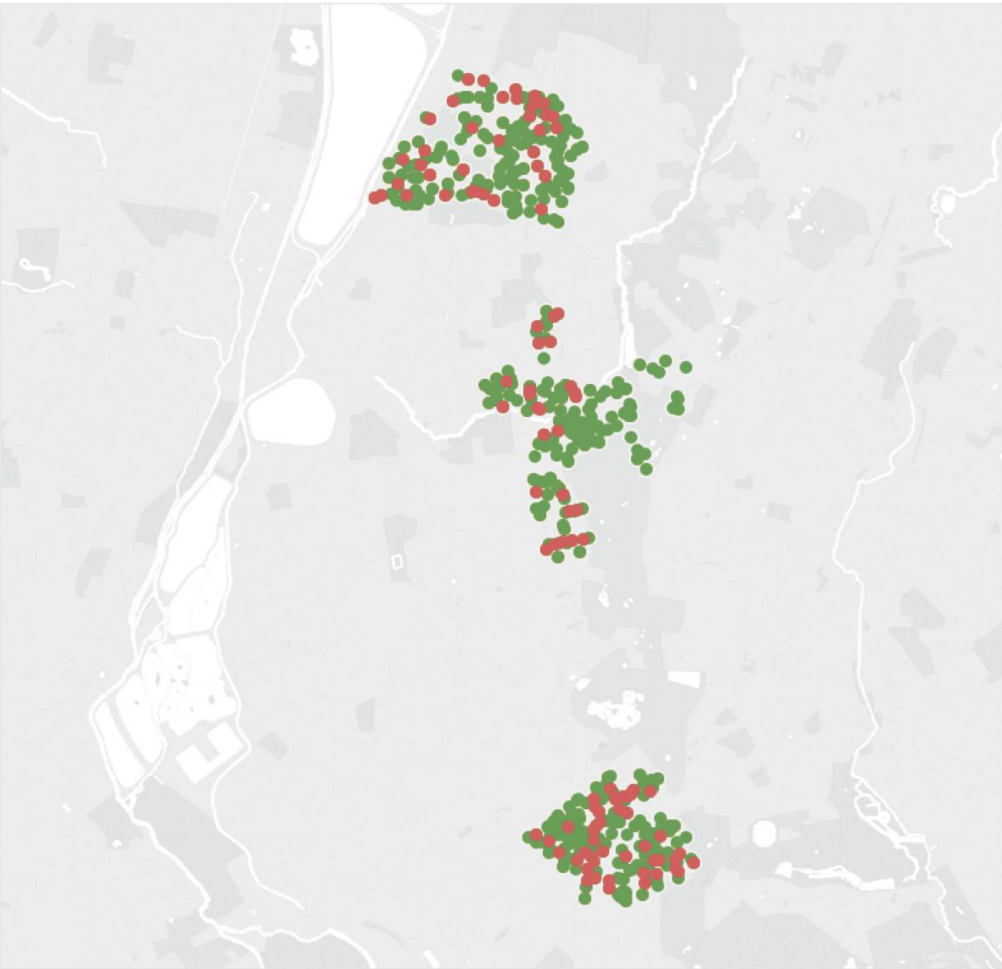
Tax Summons





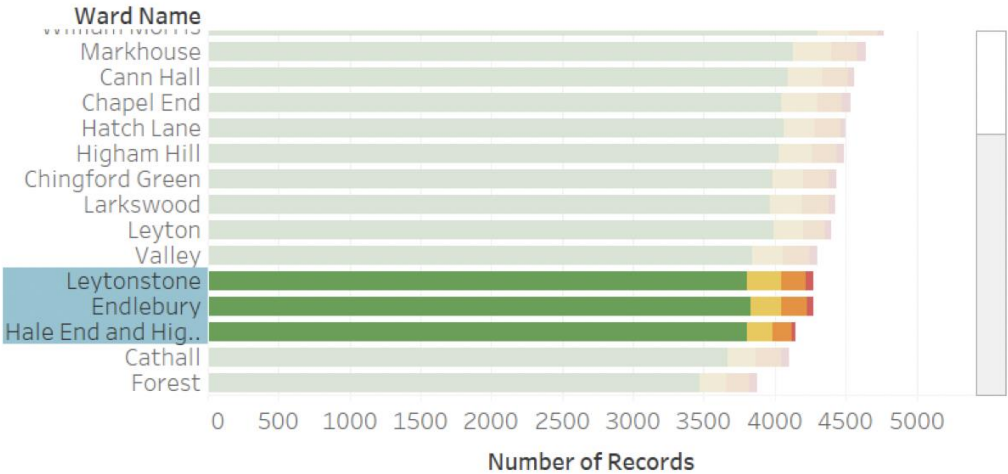
# DRILL DOWN PER WARD

Property Location

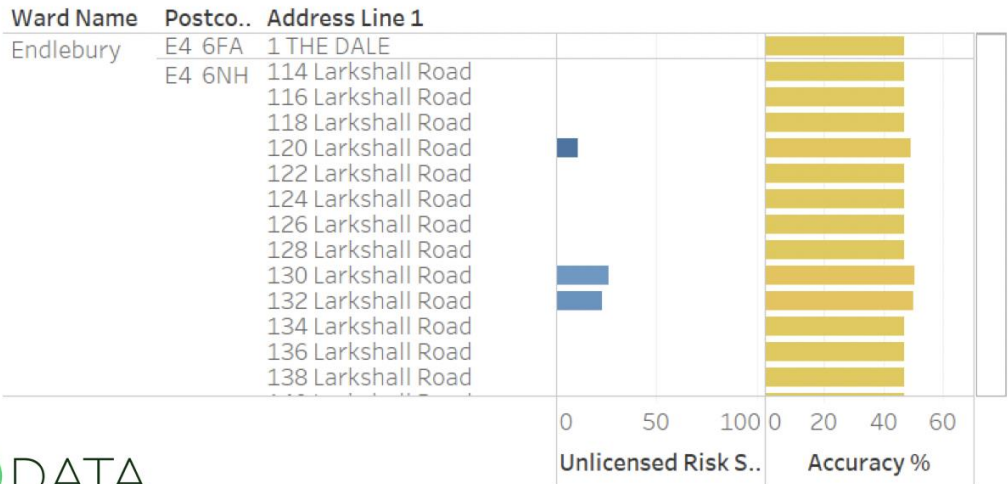


Risk Clasification  
■ High ■ Low ■ Moderate ■ Some

Risk Category By Ward



Risk & Accuracy %



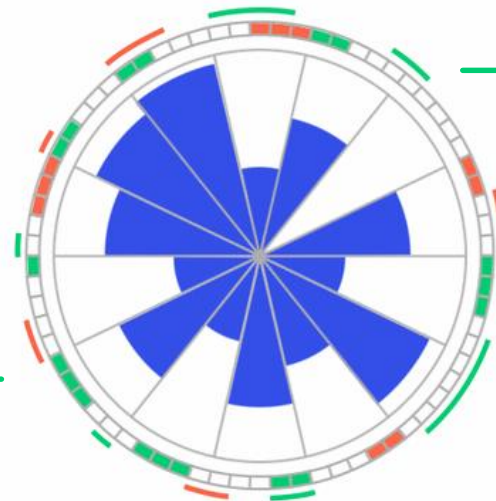
# FLEXIBLE DATA CREATION & INTEGRATION.

Quickly map and manage all council and third party data via a central integration hub to enable publish / subscribe for all data creators and consumers from systems to staff. Reduced IT costs with connections developed once per system and enable rapid access to data for both existing teams and new consumers.

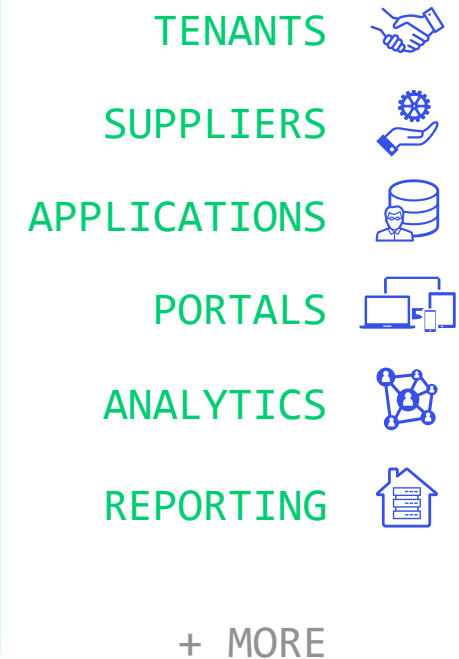
## DATA PUBLISHERS



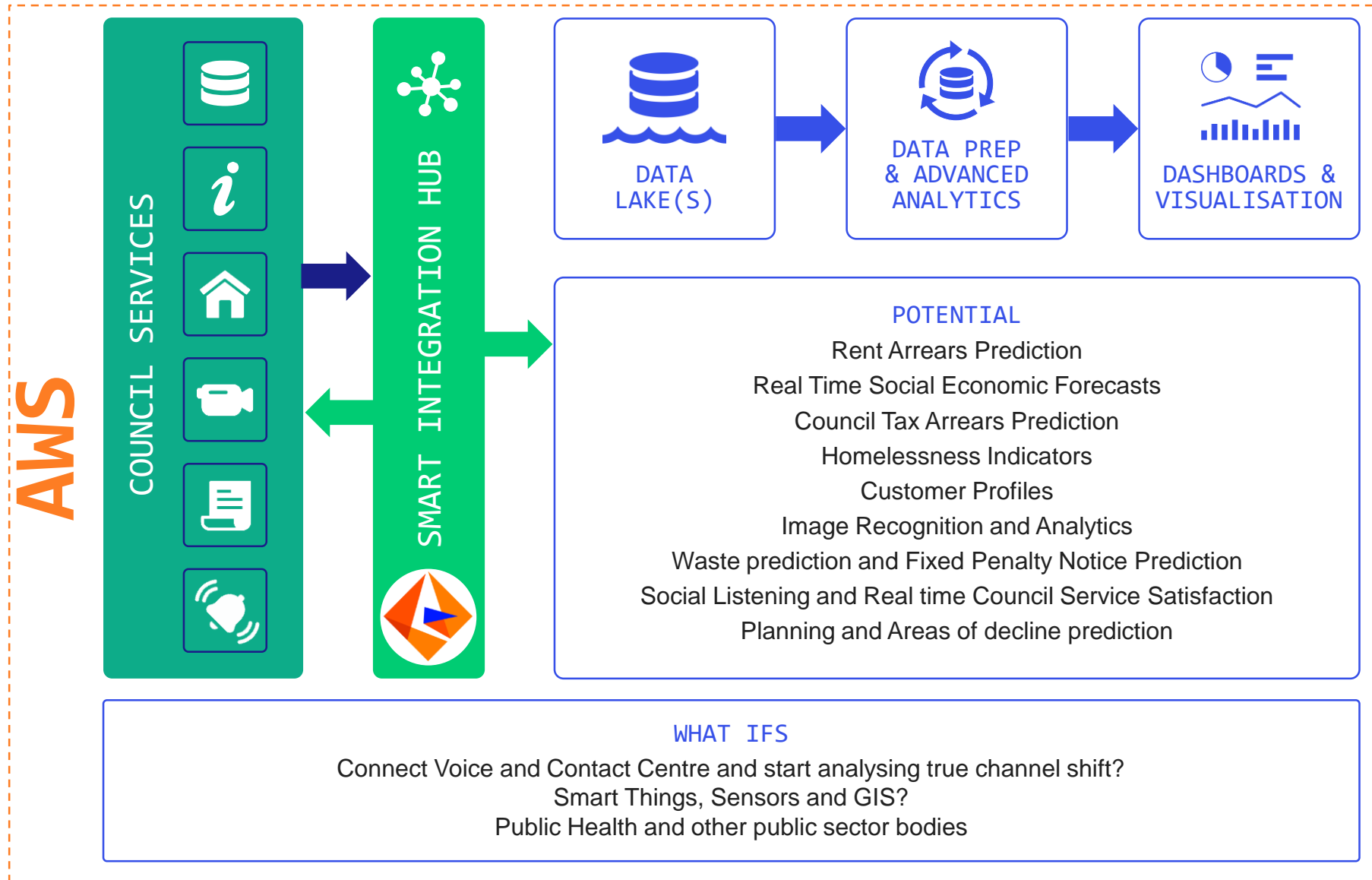
## LIVE DATA / DNA



## DATA SUBSCRIBERS



# SMART INTEGRATIONS - POTENTIAL





# EACH EXPERIENCE NEEDS TO BE DIFFERENT



New Resident



Home Owner

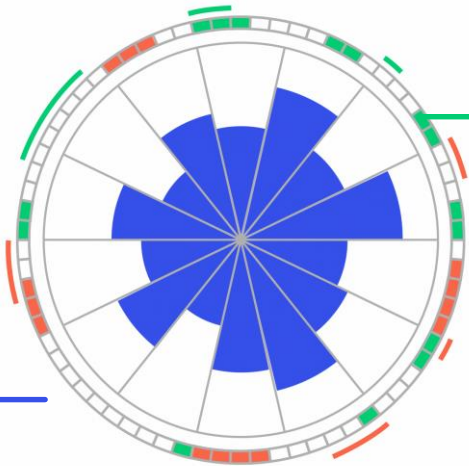


Vulnerable Resident

# DNA – NEW RESIDENT

Age	Address	Home Owner
Tenure	Tenancy	Education Status
Council Tax Status	School Registration	Library Status
Further Education Registration	ASB	Education Service Action
Contact Preference	School Meals	Service Preference
Children’s Services	Adoption	New Resident

## RESIDENT DATA / DNA

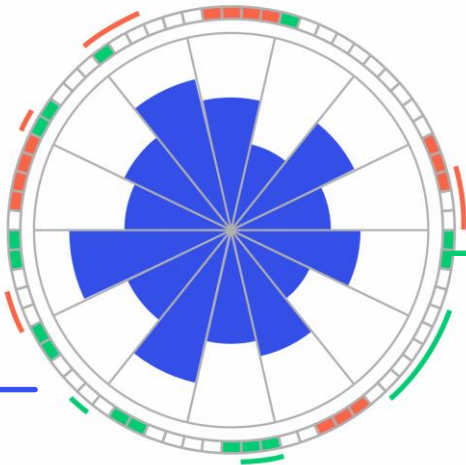


.....

# DNA – COMPLEX HOUSEHOLDS

Age	Address	Home Owner
Household ID	Size of Property	Address History
Council Tax Status	Tenancy	Property History
Planning Request	Planning History	PCN Count
Landlord	Multi-Property Landlord	Adult Social Care
PIP	Blue Badge	Freedom Pass

## RESIDENT DATA / DNA



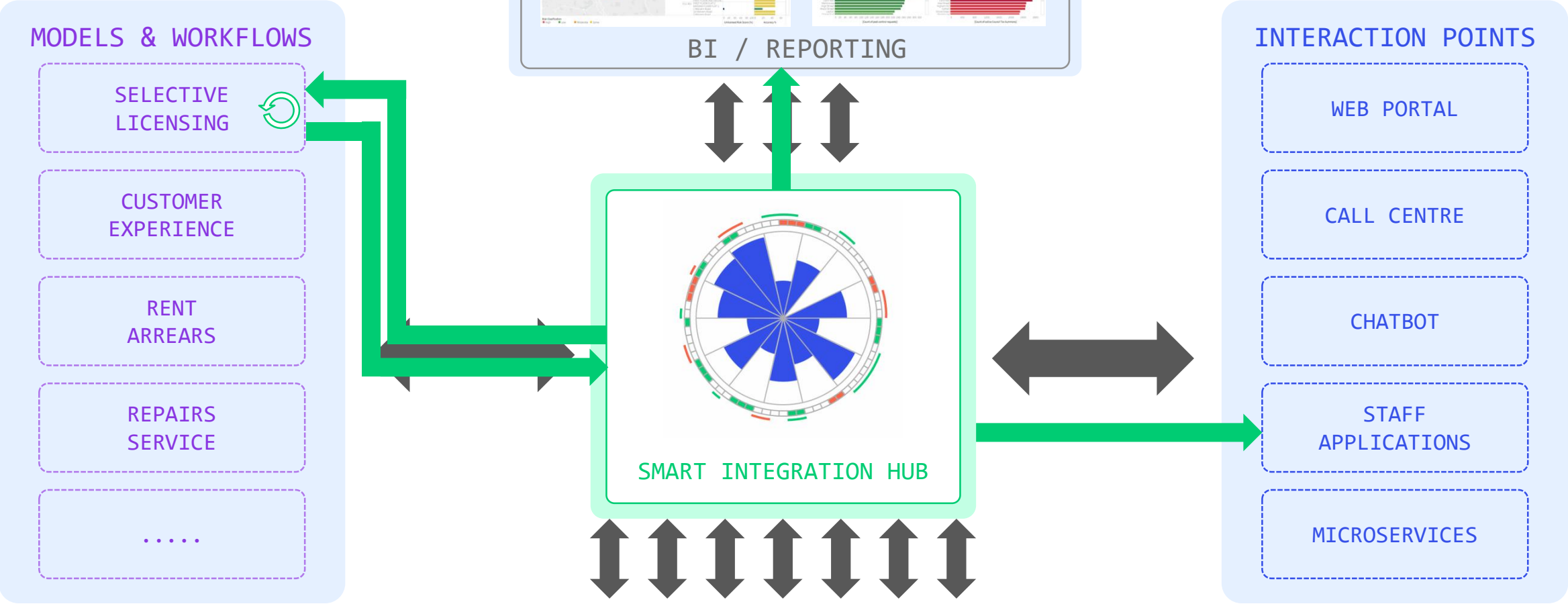
.....

# REQUIREMENTS TO EVOLVE

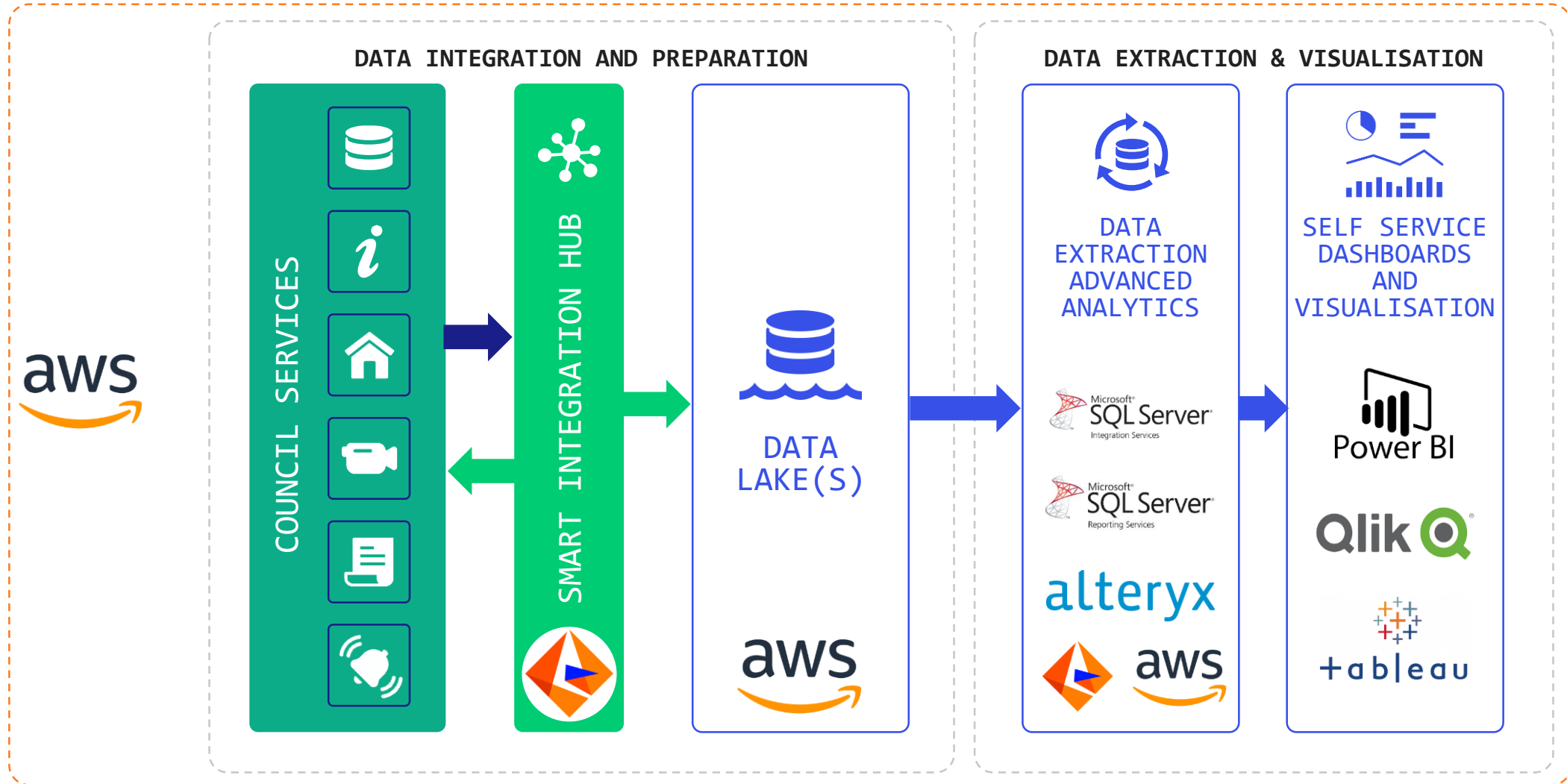




# REPORTING



# SMART INTEGRATIONS - POTENTIAL



# SUMMARY



## ✓ POC DELIVERED IN 5 WEEKS



FUTURE TECHNICAL  
ARCHITECTURE



TIMS ANALYTICS  
AND REPORTING



EVALUTATED QUALITY  
ON DATA SOURCES



## ✓ CONSUMABLE ANALYTICS & OUTCOMES



RAPID ANALYTICS &  
REPORTING DEVELOPMENT



OWNERSHIP OF  
ANALYTICS & MODELS



EASE OF DATA  
DISCOVERY



## ✓ SCALABLE FUTURE PROOF SOLUTION



CONFIGURATION  
NOT CODE



INTEGRATION &  
DATA QUALITY



NEXT GENERATION  
PLATFORM



## ✓ STARTING POINT FOR 2019 DIGITAL VISION



~4 TO 6 WEEKS  
MOVE TO PRODUCTION



OPTIMISES INVESTMENT  
=> PROCESS & ANALYTICS



ENABLES DIGITAL PROJECTS  
ACROSS BOROUGH



**Waltham Forest**

NG  DATA

**aws**

