Agenda

- **City Datastore:**
  
  *Paul Hodgson, GLA*

- **LODA Pilot:**
  
  *Andrew Collinge, GLA*

- **Digital Catapult:**
  
  *Lucie Burgess, Digital Catapult*

- **Smarter working using matched data:**
  
  *Ben Evans, LB of Newham*

- **London Ventures:**
  
  *Thomas Man, London Councils & Ian O’Donnell, LB of Ealing*

- **London Data Sharing Alliance?**:
  
  *Andrew Mobbs, LFB & Vivienne Avery, GLA*
City Datastore
Paul Hodgson
GIS & Infrastructure Manager GLA
View of Smart Cities in popular culture

'The Hunger Games' ©Lionsgate
New York ‘Hub & Spoke’ model
There will never be a single warehouse for all of London's data, so we need to connect..
There will never be a single warehouse for all of London’s data, so we need to connect..

- Open Source
- Cloud-based
- Open APIs
- Sharing knowledge with other cities

- secure sharing of catalogues &/or data
Open Data

City Data
- Personal data
- Commercial
- Early drafts
City DataStore

1. upload files
2. validate
3. share
4. metadata
5. event driven
6. break apart tables
   - Join
   - Aggregate
   - sub-set
   - auto-schema
7. search
City DataStore

1. upload files
2. validate
3. share
4. metadata
5. event driven
6. break apart tables
   • Join
   • Aggregate
   • sub-set
   • auto-schema
7. search

S3
City DataStore

1. upload files
2. validate
3. share
4. metadata
5. event driven
6. break apart tables
   • Join
   • Aggregate
   • sub-set
   • auto-schema)
7. search
City DataStore

1. upload files
2. validate
3. share
4. metadata
5. event driven
6. break apart tables
   • Join
   • Aggregate
   • sub-set
   • auto-schema)
7. search
City DataStore

1. upload files
2. validate
3. share
4. metadata
5. event driven
6. break apart tables
   - Join
   - Aggregate
   - sub-set
   - auto-schema
7. search

Event driven

split data

DynamoDb
City DataStore

1. upload files
2. validate
3. share
4. metadata
5. event driven
6. break apart tables
   • Join
   • Aggregate
   • sub-set
   • auto-schema)
7. search
City DataStore

1. upload files
2. validate
3. share
4. metadata
5. event driven
6. break apart tables
   • Join
   • Aggregate
   • sub-set
   • auto-schema
7. search

Future developments
- Automatic schema recognition
- Building up topics
How do the different programmes fit together?

SHARING CITIES
EU Horizon 2020

WITAN
(Innovate UK)

DATA STRATEGY

City DataStore

sensor system 1

sensor system 2

Advanced functions

harmonisation & aggregation

API

London Office of Technology & Innovation

Technical

Legal

Data Science

Project management

LODA

open data store
London Borough Data Partnership Meeting
13th June 2017

London Office of Data Analytics
Andrew Collinge
Assistant Director of GLA Intelligence Unit
A LONDON OFFICE OF DATA ANALYTICS

Using data to address urban challenges that we share because they cross administrative boundaries; and to drive collective innovation in public service delivery.

We need to focus on creating meaningful insight and measurable value.
The Hype Cycle

- Peak of Inflated Expectations
- Plateau of Productivity
- Slope of Enlightenment
- Trough of Disillusionment
- Technology Trigger

Gartner Group
LODA PILOT AIMS

Test the policy or service impact of data science

Show that data-sharing is possible and has tangible benefits

Develop data sharing protocols useful for the longer term

Identify barriers to collaborative working and develop solutions

Contribute to the development of a culture of data-sharing within London
LODA PILOT CHALLENGE SHORTLISTING

Build a ‘coalition of the willing’ (15 LBs attended)

Develop a long list of challenges, problem areas, ideas for action

Shortlist to a single pilot project

Identifying unlicensed Homes of Multiple Occupancy (HMOs)
LODA PILOT TIMELINE

April 2016
Consultation
Ideas for the LODA pilot were discussed at a meeting of the Borough Data Partnership

June 2016
Workshop
15 boroughs attended a workshop to determine the subject of the pilot

July 2016
Start Pilot
ASI data Science were commissioned to work on the project

November 2016
Proof of concept achieved using Westminster CC data

January 2017
DSP Finalised
Data Sharing Protocols completed and signed

March 2017
Testing
6 participating boroughs testing the model outputs

June 2017
Evaluation
Publication of LODA evaluation including recommendations for Next steps
LODA PILOT PROGRESS TO DATE

1 month
Problem identification

5 months
Data gathering and preparation

1 month
Analysis / modelling

3 months
Mobilise front-line staff

Most significant effort here

Today
LODA PILOT FROM DATA MODEL TO SERVICE TRIAL

Household-level Borough datasets HMO RELEVANT

- Resident Count
- Council Tax Band
- Building Age
- Number of Floors
- Electoral Register

Added to GLA HMO RELEVANT data

Predictive Model (Distributed Random Forest)

Information on most predictive datasets

Prioritised address list for visits (more likely to contain HMOs)
12 Boroughs commit to moving forward with the ‘data ask’ and service pilot (inc. randomised control trials)

6 move forward into live pilots, for now…

On time, sufficient data quality
- City of Westminster
- Camden
- Royal Kingston
- London Borough of Barking & Dagenham

Behind schedule, sufficient data quality
- Lambeth
- Islington

Outstanding issues or blockers
- Brent
- Sutton
- London Borough of Bexley
- Lewisham
Biggest issue: matching and linking data to UPRN or similar unique identifier
- Data Maturity Assessment needed at start of project to save time/effort

Huge range of ‘housing features’ data (>40 to 5)
- Most boroughs do not have sufficient technology/capacity to meet requirements within project timeframe

Precise data requirements would have helped LBs to prioritise activities under resource constraints
- This rather than flexibility and creativity

Absence of data warehousing in LBs means significant effort and time needed to work across departments
DATA ISSUES

Data quantity: When it comes to machine learning, the more data the better. However, most boroughs struggled to provide a sufficient quantity and variety of data across all properties within the timescales of the pilot. Also, a lack of known HMOs in the borough meant the machine learning model had too few cases to train on to reliably predict other HMOs.

Data quality: Data submitted by most boroughs required significant cleaning, processing and merging. With every merge, as much as 10% of properties would be lost when records failed to match up.

Data availability: Data on Private Rental Sector properties, which could have helped filter out owner-occupied and other ineligible property types, was a critical missing piece of the puzzle.

Lack of precise data requirements: In some cases more precise and prescriptive requests for datasets could have helped boroughs prioritize what they provided.
TECHNICAL ISSUES

Lack of matching technologies in boroughs: The inability to accurately match and link datasets significantly influenced the quality and quantity of data individual boroughs were able to provide.

Absence of data warehousing: Boroughs with centralised business intelligence teams and data warehouses had an easier time pulling data from across the organisation.

In-house expertise: The range of technical expertise available in-house varied across boroughs. For example, in one case, a borough would have had to contract a supplier to extract data related to its housing benefits.
CAPACITY ISSUES

In Boroughs: Though every effort was made to minimise the burden on participants the pilot required a larger number of staff and resources to implement than anticipated.

In Nesta: Most staff time was spent on fielding highly specific technical, legal, and operational questions, and supporting the overall data acquisition process. This left less time to focus on risk mitigation, creative problem-solving and identifying opportunities for development and sharing of best-practice.

In ASI: As the project increased in complexity, our data science partner was challenged to provide on-going and in-depth guidance across all participating boroughs. This level of support was unexpected and difficult to meet on a consistent basis throughout the project.
LODA PILOT OPEN LEARNING EXERCISE

LODA Pilot Update

As announced in previous blogs, Nestle is working with the GLA and more than a dozen boroughs in London - and with local authorities, the Digital Catapult, and Sunderland Software City in the North East - to pilot data analytics projects that address public service challenges.

This post provides a brief update on the latest developments.

What can you do with data?

For each pilot the first objective has been to identify a public service challenge for which there is: 1) a big problem to solve, 2) good data available, and 3) a chance of making an impact.

The data will be used to make recommendations for how the boroughs can learn from each other and the results of the pilots.

The Data Science – Identifying HMOs

Working with the City of Westminster our data science partners at the GLA have developed a predictive model for identifying unlicensed HMOs. The model takes the local authority’s own data on properties in the private rented sector to identify likely addresses for unlicensed HMOs. The database of properties that can then be passed back to the local authority as a prior to follow-up.

The next step is to feed data into the other pilot boroughs, and to that end and we are currently working with a first pilot borough (Bexley, Camden, Haringey, and Lambeth) to supply housing and environmental health data to the GLA. Our solution will help the local authorities to get a better understanding of the challenges they are facing.

This is a exciting project which we think will help boroughs to more efficiently prioritize their investigations of HMOs. We hope this data-based approach will have real tangible impacts on HMO registrations in London, putting pressure on unlicensed landlords.

The Process – Enabling Information Sharing

With help from five boroughs, we have drafted an Information Sharing Protocol for the pilot to ensure we use shared data securely, legally, and ethically. It is also clear that we are going to demonstrate that.

Traditionally, the complexity of government and the pilot's scope is daunting, but the pilot is showing the way forward to.

The Issue – Where are the unlicensed HMOs?

The issue that is thought to have the most potential was identified as unlicensed HMOs, houses of multiple occupancy (HMOs) by at least three people who are not from one household – but who share facilities such as a bathroom and kitchen. HMO licences only apply to properties in London.

The London Office of Data Analytics Pilot: two weeks of showing and telling to focus the data science and sharpen the overall approach

By Laura Armstrong

It’s been 2 weeks since our kick-off workshop for the London Office of Data Analytics (LODA) pilot programme, a joint venture between the GLA and Nestlé, with involvement from nearly half of the London boroughs. The broader context shows a real sense of growing intent and purpose around data sharing for impact. In the team as we set out for the first pilot meeting – a show and tell event that we want to offer to the boroughs – and having good data availability and good data quality are key to this.

But as our own exercise, the straightway goals are to find actionable insights that save money on public services, and in the process show that pooling up data from multiple boroughs can lead to solutions benefitting Londoners that wouldn’t be possible otherwise. The consensus view is that unlicensed HMOs (houses of multiple occupancy) is an issue that is important to the boroughs and well suited to a data-driven approach that will lead to those much desired practical, identifiable outcomes.

The task of working out a statistical approach to the problem has now begun.

Three lessons on City Data Analytics from Mike Flowers

By Eddie Copeland, Director of Government Innovation, Nestle

I’ve previously written about plans by the GLA, Andrew Collings, and Nestle to run a pilot for a London Office of Data Analytics, inspired by the Mayor’s Office of Data Analytics (MODA) in New York City.

Last week saw significant progress when 15 London boroughs came together for a workshop with the GLA to select the public service challenge that could be tackled with data. This is shortly publish details of the six shortlisted challenges discussed (which covered areas from health to waste management, and from housing to social care), and the conclusions we reached.

The boroughs were asked to vote on each according to the extent that it would likely be:

- a significant money-saver
- have good data available, and
- lead to actionable insights.

Deliberations were made easier thanks to the presence of Mike Flowers, Chief Analytics Officer of Enginio (a leading US data analytics company) and the creator of the MODA model. Mike advised the boroughs to consider three additional factors when making their assessments:

1. Keep it simple

Given the external pressure on public finances, local authorities are understandably tempted to tackle the most expensive and important problems first, but if it doesn’t address areas such as adult and child social care – the biggest line items of local authority expenditure? Mike’s advice: wait before you run. One significant challenge with social care issues is that they entail using a lot of personal data. Ensuring that all the right protections have been put in place, clear and established to, and comments received can take a huge amount of time. At best, that delays the start of any data initiative. At worst, there’s a risk of inanimately stumbling into another data breach, spiralling people, and wasting the whole data analytics agenda back years.
LODA PILOT WILL DELIVER…

- Open source, reusable data model (across LBs and other regions)
- Full toolkit (Data Sharing Agreement; Privacy Impact Assessment)
- Actionable intelligence for a front line service in participating boroughs
- Proof of concept – demonstrate to boroughs that a LODA is feasible and worthwhile
- Evaluation report…
LODA PILOT EVALUATION

• Evaluation of results (did the data science work?)
  - Was the pilot successful in terms of driving new practice, savings and other (e.g. public health) outcomes

• Review of the process
  - Full costs and benefits, estimation of time spent across key project activities (e.g. data sharing)

• Recommendations on future operating model
Other observations at this stage

1. **80:20 rule**
   - Culture, organisational capacity, co-ordination
   - Data science and supporting technology/data infrastructure

2. **A LODA is more complex than a MODA because of the operating environment**
   - So what can we achieve and how can we achieve it?
   - Boroughs are very important as data and ‘problem’ owners
OPERATING MODEL
ADAPTING THE LONDON OPEN MODEL

OFFICE OF DATA ANALYTICS
DATA SCIENCE
LEGAL
TECHNICAL
STANDARDS

Data Flow

Public/Private City Datastore

LONDON BOROUGHS
LONDON COUNCILS
LFEPA
MOPAC
TFL
ODPC

LONDON BOROUGHS
LONDON GOV
PUBLIC
BUSINESS
CHARITIES
ACADEMIC INSTITUTIONS
Provide additional (or initial) Data Science expertise
- data-driven policy decisions and tools
- new products and services (for front-line staff)

Facilitate data sharing
- technical/legal support (e.g. let’s do GDPR once only)
- develop and promote open shared standards for data management and use
- share/exchange data via London Datastore or a secure City Datastore

Moving from Borough Data Partnership to a Data Academy (management, visualisation, analysis – see San Francisco)

Programme and Project management (identify “good data projects” >2 / partner / yr)

Impactful Questions | Accessible data | Actionable insights
NEW FORMS OF DATA, NEW RELATIONSHIPS, NEW LINES OF INQUIRY…. NEW RULES
OPERATING MODELS

LODA

- Project management
- Data Science
- Technical
- Legal
OPERATING MODELS RESOURCES WE COULD CALL ON

- **Project management**
  - GLA IU
  - Existing posts
  - Contract funded posts

- **Data Science**
  - Boroughs
  - Alan Turing inst.
  - Urban Dynamics Lab at UCL
  - Large tech/analysis companies
  - Specialist SMEs
    - Mastodon C
    - ASI
    - Social Finance
  - TfL legal

- **Technical**
  - City DataStore (secure data sharing)
  - Commercial (new city data) datasets
  - GLA virtual PCs

- **Legal**
  - Call-off contract with legal firm
  - Existing data sharing agreements
    - Anti-fraud hub (led by Ealing)
    - LODA pilot
    - LfB / Information Commissioner’s Office templates

- **Gtr. Manchester**
  - Existing data sharing agreements

- **IU managers**
  - Contract funded posts

- **GLA IU**
  - Existing posts
  - Contract funded posts

- **LODA**
OPERATING MODEL OPTIONS

A. Virtual team
- 100% existing resources from GLA, Boroughs (& possibly universities)
- Would lead to greater sharing than at present
- Project-by-project agreements
- Limited initial impact

B. Core LODA team
- Recruit small dedicated team (e.g. 2/3 staff)
- Grow organically by demand & funding (possibly through savings)

C. Consultants
- Pay by results
- ‘free’ offers
- Project-by-project agreements

D. Big Bang
- Large initial investment (go straight to Amsterdam-sized team of 14 + staff)
- Capacity to tackle large challenges
OPERATING MODEL OPTIONS

A. Virtual team
- 100% existing resources from GLA, Boroughs (& possibly universities)
- Would lead to greater sharing than at present
- Project-by-project agreements
- Limited initial impact

B. Core LODA team
- Recruit small dedicated team (e.g. 2/3 staff)
- Grow organically by demand & funding (possibly through savings)

C. Consultants
- Pay by results
- ‘free’ offers
- Project-by-project agreements

D. Big Bang
- Large initial investment (go straight to Amsterdam-sized team of 14 + staff)
- Capacity to tackle large challenges

Increased upfront cost / risk
OPERATING MODELS OPTIONS

• Information Scheme
  - All boroughs contribute an annual sum

• Commissioning Model
  - Project-by-project commissioning with different groups of boroughs involved in different projects

• Pay by Results
  - Partner organisations are commissioned on a pay by results model
EXISTING ACTIVITIES AND PIPELINE
DATA-LED CHALLENGES SCHOOL ROLL PROJECTIONS

• Bespoke demographic and school roll projection service for Boroughs
• Boroughs provide local intelligence to GLA
• GLA combines national and local data in its state-of-the-art projection models
• Outputs inform local school place planning and sites such as the Schools Atlas
DATA-LED CHALLENGES HOUSING BENEFIT

- Mobility of Housing Benefit claimants as part of Cabinet Office Data Science Accelerator programme
- Applied Data Science techniques to claimant characteristics date from DWP data to identify spatial and temporal characteristics
- Identity any trends over time and space
DATA-LED CHALLENGES AIR QUALITY

**Project 1** - Set up a central Air Quality Data Store

**Project 2** - Work with London’s tech sector to create a first generation of apps and websites

**Project 3** - Carry out analytics to identify the areas and times of day/week when interventions would have the greatest impact.

**Project 4** - Carry out research and develop guidelines for how a multi-layered network of sensors can be used to maximum benefit.
Smarter working using matched data

Ben Evans, Data Warehouse Programme Manager
LB of Newham
Data Warehouse and Business Intelligence Programme

Smarter working using matched data
About Newham

- Population: 343,015

- It is one of the most ethnically diverse places in the UK with no single ethnic group having a majority

- Deprivation is high in Newham but improving. Ranked 25th in IMD 2015, down from 2nd in 2010

- Directly Elected Mayor: Sir Robin Wales
A genuine single view of Newham’s people and properties
Challenges

- Buy-in from Senior Leadership
- Technical
- Information Governance
- Skills
- Data Quality
Using The Data Warehouse

- Strategic
- Analytical
- Operational
<table>
<thead>
<tr>
<th><strong>Person Details</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtual Id</td>
<td>1381424</td>
</tr>
<tr>
<td>Customer Name</td>
<td>Bill Gates</td>
</tr>
<tr>
<td>Date Of Birth</td>
<td>28/10/1955 (60)</td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
</tr>
<tr>
<td>UPRN</td>
<td>000046067795</td>
</tr>
<tr>
<td>Postcode</td>
<td>SW1A 2AA</td>
</tr>
<tr>
<td>Address</td>
<td>10 Downing Street</td>
</tr>
<tr>
<td>NINO</td>
<td>WE542542D</td>
</tr>
<tr>
<td>NHS Number</td>
<td>442453444</td>
</tr>
<tr>
<td>Marital Status</td>
<td>Single</td>
</tr>
<tr>
<td>Mobile Number(s)</td>
<td>07803777777 (01/05/2015)</td>
</tr>
<tr>
<td>Telephone Number(s)</td>
<td>0208555666 (14/11/2014) 0208223333 (06/06/2012)</td>
</tr>
<tr>
<td>Email Addresses(s)</td>
<td><a href="mailto:Bill.Gates@microsoft.com">Bill.Gates@microsoft.com</a> (22/03/2015) <a href="mailto:newham@gmail.com">newham@gmail.com</a> (01/05/2015)</td>
</tr>
<tr>
<td>First Known Date</td>
<td>07/10/1996</td>
</tr>
<tr>
<td>Last Known Date</td>
<td>22/03/2015</td>
</tr>
<tr>
<td>Most Likely Property Ranking Value Max10</td>
<td>10</td>
</tr>
<tr>
<td>Most Likely Property Ranking Value</td>
<td>1330</td>
</tr>
</tbody>
</table>

**Troubled Families**

- Social Care Overall Flag: True
- Social Care Section 47: True
- Social Care Child Protection Plan: True
- Employment Out Of Work Benefits: True
- Health Overall Flag: True
- Health Adult Mental Health And Parent: True
Predictive Analytics

Using statistical and machine learning techniques such as regression, classification, probabilistic modelling to target our resources effectively.
Outcomes

We have saved or generated an estimated £1.2m in the first year.

**Transactional**

- Single Person Discount Fraud
  Estimated £35k per annum

- Freedom Pass Fraud
  Estimated £100k saving p/a

- New Homes Bonus
  Estimated £280k income

**Transformational**

- PRS Landlord licensing
  £700k income p/a

- CYPS Triage
  £500k saving p/a

- Homeless applications
  Estimated £20k saving p/a
Building on our success

Grow the current system

Push the boundaries

Share our knowledge
London Borough Data Partnership Meeting
13th June 2017

London Ventures
Thomas Man, Head of Capital Ambition
Ian O’Donnell, Executive Director of Corporate Resources, LB of Ealing
This report is intended for the sole use of London Ventures. EY and London Councils shall have no responsibility whatsoever to any third party in respect of the contents of this report. It should not be provided to any third party without EY or London Councils prior written consent.

London Borough Data Partnership

Tuesday 13 June 2017
London Ventures is a partnership between London Councils and EY. The programme brings innovative solutions to local government to transform services, save money, and ultimately deliver improved benefits for Londoners.
General Ventures

SOCIAL CARE & WELLBEING
- xantura
- alcove
- MyCognition
- circle
- visbuzz

COMMUNITY
- Spacehive
- Commonplace

BACK OFFICE
- blueprism
- Cerno
- FISCAL TECHNOLOGIES
- London Counter Fraud Hub
- Quadnet
Targeted Ventures

Our approach:

Choose a priority area

Understand the problem

Find the solution

Develop the solution

Build and Launch

Launch event

Concept portfolio

Dragons den

Improve outcomes in the priority area

Support the delivery of 5+ concepts

Timelines for targeted ventures cycle 1:

Phase 1: Oct 16 – Feb 17

Phase 2: Feb 17 – Apr 17

Phase 3: Apr 17 – Jun 17

Phase 4: Jun 17 – …

Support concepts and solutions to implementation

- We have £100,000 seed funding to support and nurture initiatives
- We are committed to delivering better outcomes for Local Authorities and Londoners through innovation

- We have a trusted brand in the local authority market, LV provide endorsement and access to the LA market
- We have strategic relationships with local authorities who can sponsor programmes of work and initiatives
We’d love to support innovation in your local authority

Please contact us to find out more…

londonventures@uk.ey.com

@LdnVentures

Search ‘London Ventures’
London Counter Fraud Hub

Ian O’Donnell
Executive Director of Corporate Resources
London Borough of Ealing

13th June 2017
Agenda

• Drivers for change

• Challenges

• Solutions

• Data & Data Analytics

• Lessons Learnt
Drivers for Change

- Fraud losses
- Ongoing funding cuts to local government
- National counter fraud strategy for councils
- Opportunity to harness new technologies

Fraud detected in 2016 (£61.8 million)
Challenges

A collaboration between all 33 London boroughs using the latest data analytics technology to prevent and detect fraud, aiming to save £60 million+ per annum.
Solutions

**Collaboration** – Multi level stakeholder engagement. Obtained backing from London Councils through Capital Ambition project. Used London local authority professional networks.

**Data Sharing** – Shared legal advice commissioned on data protection issues, and data-related agreements necessary to manage risk included in contract.

**Funding** – Small grant from DCLG used to develop concept further and conduct procurement. Private sector risk capital identified as primary funding source, leveraged through payment by results commercial model over 9 year term.

**Market** – Held informal dialogue with market to gauge interest, plant ideas with suppliers, and shape model using Capital Ambition / EY.
Solutions

**Evolving Solution** – Payment by results drives investment in innovation and long term transfer from detection to prevention.

**Analytics** – Procured advanced data analytics capability. Ability to design and run enquiries is built into solution, enabling evaluation of identity and entitlement at point of contact.

**Value For Money** – Competitive tendering process and pilot period testing the product thoroughly.
Use of Data & Data Analytics

The solution performs both complex data matching and sophisticated risk analysis of the data. Therefore, the cases that are presented to end users are not just the result of data matching, they have also been analytically assessed for risk. This approach is extremely effective at reducing false positives, as the solution is able to analyse the most complete set of data available for an entity before deciding whether that entity poses a risk or not.

High level data flow and how the Hub’s analytics engine generates data for the end user

1) Council data
2) Third party commercial and public sector data
   - GRO Deceased data
   - Land Registry data
   - Transactional data
   - Credit Reference Agency data
   - Companies House data
   - Ordnance Survey (OS)
3) CIFAS data
   (the world’s largest network of fraud data including confirmed identity frauds and compromised bank details)
4) Intelligence
   Live Amberhill data
   (the Home Office database which captures details about forged identity documents)

Alert & Case Management Outcomes

Entity Resolution → Build All Links → Build Relevant Networks → Mask & Enhance Networks → Score Entities & Networks

Alert Management & Case Management

Key is to move from detecting to preventing fraud
Data Flow

**FIREWALLED HUB, SECURE HOSTED DATA STORE**

<table>
<thead>
<tr>
<th>STAGE</th>
<th>DATA FLOW</th>
<th>SYSTEMS USING THE DATA</th>
<th>WHO HOLDS THE DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATA COLLATION AND AUTOMATED ANALYSIS</td>
<td>RISK SCORING AND CASE GENERATION</td>
<td>INVESTIGATION, REPORTING AND ANALYSIS</td>
<td>SANCTIONS, REDRESS AND COLLABORATION</td>
</tr>
<tr>
<td>Data is securely transferred to the Hub’s analytics engine.</td>
<td>Data is risk assessed and cases are generated in order of priority.</td>
<td>Cases are reviewed by end users, outcomes and feedback are captured.</td>
<td>Data securely transferred to other agencies as needed</td>
</tr>
<tr>
<td>It is cleansed and transformed before undergoing entity resolution and social network analysis.</td>
<td>Data is created for reporting and strategic analysis.</td>
<td>Data is securely transferred to LAs as required.</td>
<td>Data is accessed through:</td>
</tr>
<tr>
<td>Data is not accessed by any end users.</td>
<td>Data is not accessed by any end users.</td>
<td></td>
<td>- Hub investigation interfaces</td>
</tr>
<tr>
<td>Secure Hub infrastructure</td>
<td>Secure Hub infrastructure</td>
<td></td>
<td>- Hub or LA reporting interface</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Hub or LA case management</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Data is accessed through other agency systems</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Other agency infrastructure</td>
</tr>
</tbody>
</table>
Case Types

Our Hub Analytics Engine will generate cases of suspected fraud and error AND service enquiries.

**Simple Cases**
- Reviewed by Councils
- Payment by Results
- Councils
- Payment Trigger: Recovery Initiated

**Complex Cases**
- Investigated by Hub
- Fixed Fee + Payment by Results
- Hub (2a)
- Payment Trigger: Investigation & Recoverability Proven by Hub
- Investigated by Councils
- Fixed Fee
- Councils (2b)
- Payment Trigger: Case Accepted
Progress To Date

5 Pilot Authorities are taking part in the pilot

NFI and NNDR data provided for POC

Testing commercial and operating model principles

Governance structure set up

Clearly defined and measurable performance targets

Ensuring there is operational readiness and available resources to deal with the case plans
A London Data Sharing Alliance?

Andrew Mobbs, BI Manager, LFB
Vivienne Avery, Demography & Policy Manager, GLA
London wide data sharing

Existing example of SafeStats crime and disorder intelligence depository

Impact of the GDPR on existing data sharing arrangements

What about data sharing for other objectives?

Can we work together to deliver a London-wide data sharing platform?
What is Safestats?

Secure data repository hosting and visualising multi-agency crime and community safety data

Available on an authorised-only basis to professionals working on the reduction of crime

Operating since 2001

Holds data from

- London Ambulance Service
- British Transport Police
- Metropolitan Police Service
- London Fire Brigade
- Transport for London
- Hospital Emergency departments
Current Safestats data sharing

**Disclosure MoU**
- Allows GLA to receive data from ‘Disclosing Bodies’ and act as a depository
- In line with statutory duties in crime and disorder legislation
- These duties set a framework for the receipt of the data

**Receipt MoU**
- Allows users to receive data via the GLA from disclosing bodies
- For purpose of reducing crime and disorder
- Allows both GLA and disclosing bodies to carry out those duties
Safestats and the GDPR

- Current MoUs date back to 2008 and have generally worked well
- Reliance on particular legislation can be restrictive e.g. policy making on public health and alcohol usage
- As part of a rebuilding project considering how Safestats will address the GDPR – General Data Protection Regulation
  - Designed to strengthen data protection for EU citizens
  - Comes into force in May 2018

*Should we develop a new Safestats solution to GDPR*

or

*Would London benefit from a broader approach to data-sharing across the region?*
How will you share data under GDPR?

ICO’s Overview of the GDPR

• **Individuals right to be informed [about]**
  - Any recipient or categories of recipients of the personal data
  - The source the personal data originates from and whether it came from publicly accessible sources

• **Individuals right of rectification/erasure/restriction**
  - If you have disclosed the personal data in question to third parties, you must inform them of the rectification where possible. You must also inform the individuals about the third parties to whom the data has been disclosed where appropriate.
  - If you have disclosed the personal data in question to third parties, you must inform them about the erasure of the personal data, unless it is impossible or involves disproportionate effort to do so.
  - If you have disclosed the personal data in question to third parties, you must inform them about the restriction on the processing of the personal data, unless it is impossible or involves disproportionate effort to do so.

• **Accountability and governance**
  - Implement appropriate technical and organisational measures that ensure and demonstrate that you comply.
  - Maintain relevant documentation on processing activities.
  - Implement measures that meet the principles of data protection by design and data protection by default. Measures could include: … Allowing individuals to monitor processing

ICO’s draft guidance on consent

• **you must identify yourself, and also name any third parties who will be relying on consent.**

• **Name your organisation and any third parties who will be relying on consent – even precisely defined categories of third-party organisations will not be acceptable under the GDPR.**
Can we solve together?

- Inspired by examples of other counties working within one data sharing framework
  - Whole Essex Information Sharing Framework ([https://weisf.essex.gov.uk](https://weisf.essex.gov.uk))
  - Lancashire and Cumbria Information Sharing Gateway ([www.informationsharinggateway.org.uk](http://www.informationsharinggateway.org.uk))

- Common framework

- Agreed set of templates for data sharing protocols, agreements, PIAs, etc

- Can we go further a have a common secure platform?
London data sharing alliance

- Common framework with agreed set of templates for data sharing protocols, agreements, PIAs, etc
- Organisation signs up to the data sharing partnership
- Securely upload their data to the platform
- Assign which other organisations can access the data
- Recipients authenticate and confirm purpose
- Transactions and downloads fully audited and logged
- Individuals can view their records and who/when data has been transferred [gov.uk verify?] 
- Notification process for rectification/ erasure/ restriction