LEAP Data Integration Platform

LEAP

Lambeth Early Action Partnership

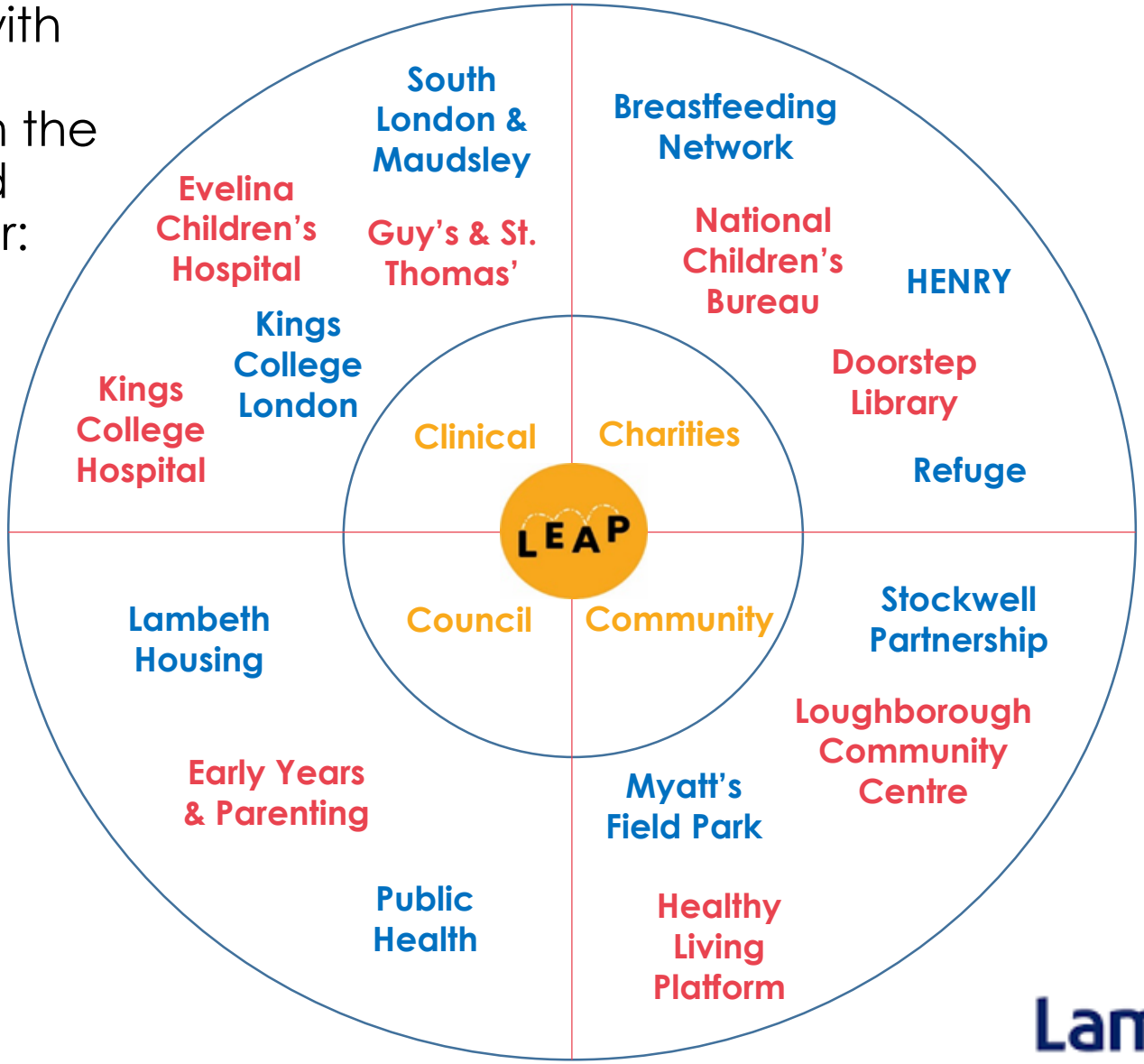
Update as at mid Jun 2021

Introduction to LEAP

- LEAP is an innovative programme created to better the lives of thousands of children in the Lambeth community, focusing on four wards: Stockwell, Coldharbour, Vassall and Tulse Hill.
- Our goal is to make Lambeth the best place in the world for a baby to be born and to grow up.
- Funded by the National Lottery Community Fund, and working with partners locally and nationally, LEAP is a £38m ten year project that aims to support the social, emotional, communication and language development of babies and children, their diet and nutrition as well as parents' wellbeing, their social networks and the strength of their communities and wider environment.

LEAP Partners

LEAP works with a range of partners from the statutory and charity sector:



LEAP Service Landscape

LEAP
Lambeth Early Action Partnership

Diet & Nutrition Strand

- Community Activity & Nutrition
- Pregnancy Information for Nutrition and Exercise
- Oral Health Packs
- Oral Health Supervised Toothbrushing
- LEAP Into Healthy Living
- Breastfeeding Peer Support
- Family Nutrition
- Environmental Health

Social & Emotional Development Strand

- PAIRS 1-2-1
- PAIRS Together Time
- PAIRS Circle of Security
- Empowering People Empowering Communities
- Overcrowded Housing
- DV Enhanced Caseworkers
- DV Groups
- Family Nurse Partnership
- Baby Steps

Communication & Language Development Strand

- Making it REAL
- Sharing REAL with Parents
- Babies' Next Steps
- Speech and Language Therapy (Evelina Award)
- Speech and Language Therapy (Chattertime)
- Doorstep Library
- Natural Thinkers

Systems Change Strand

- Parent Champions
- Family Engagement Workers
- Capital Programme
- Family Partnership Model

Cross Theme Strand

- Caseload Midwifery
- Group Pregnancy Care
- Maternity Pathway Coordinators

Outline of the problem

- Reporting systems for LEAP interventions were in silos. Individual providers sent anonymised and aggregated data to LEAP on a quarterly basis. This data couldn't be linked across LEAP's services.
- This created a number of challenges:
  - It prevented LEAP from building a full understanding of who accessed its services (and who did not) and patterns of engagement.
  - It inhibited the ability to evaluate the collective impact of LEAP services for beneficiaries.
  - Most critically, it did not enable accurate reporting on unique beneficiaries to the Funder (i.e. overall reach figures).
- The data integration platform seeks to help solve these problems.
- An ITT for a strategic lead and an organisation to develop and maintain the platform were issued and awarded respectively to Fotheringham Associates and Lambeth Council.

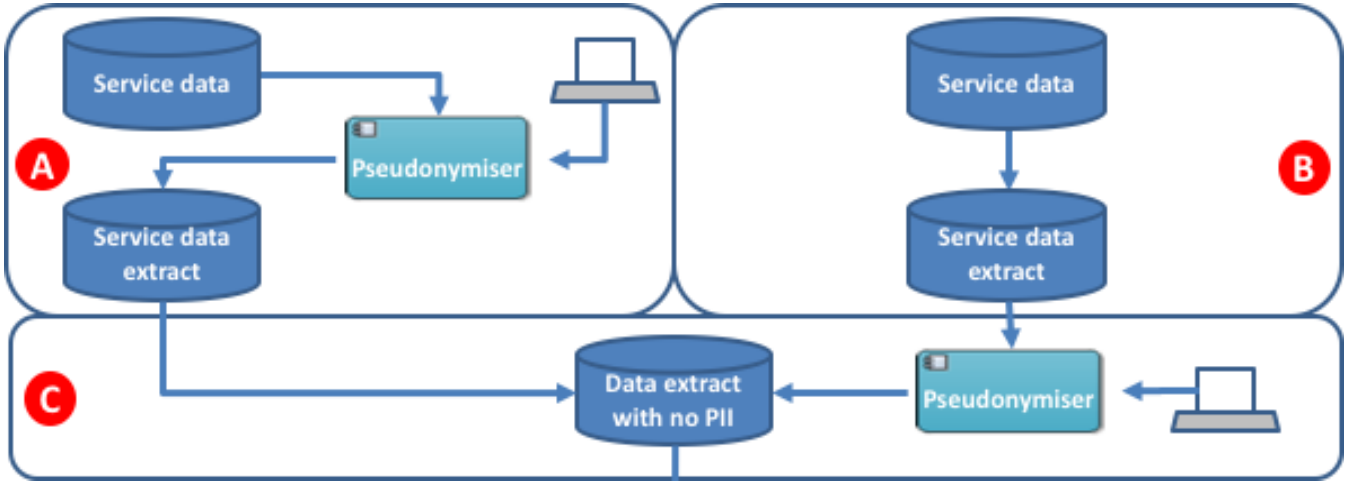
Solving the unique beneficiary problem

- A key challenge within the project was defining an approach to uniquely identifying beneficiaries.
- Use of standard identifiers such as names, postcodes, NHS numbers, mobile numbers were considered.
- Not all services have NHS numbers, so the following was decided upon:
  - For child beneficiaries = Using a key that consists of parent email address*, child dob, child gender and part of their first name**
  - For other beneficiaries = Key is the Email address*
- This type of data is classed as Personal Identifiable Information (PII) and therefore has to be protected.
- To overcome this a pseudonymisation approach was undertaken
  - Pseudonymisation is a technique where we swap identifiable data for non-identifiable data via an algorithm which provides consistent results even from different locations.

* According to ONS 2018 figures 99% of age range 16-34 have an email address (email addresses have to be unique)
** Same gender multiple birth children could cause an issue but the recorded numbers of these in the Lambeth borough according to the ONS are very small (<1%)

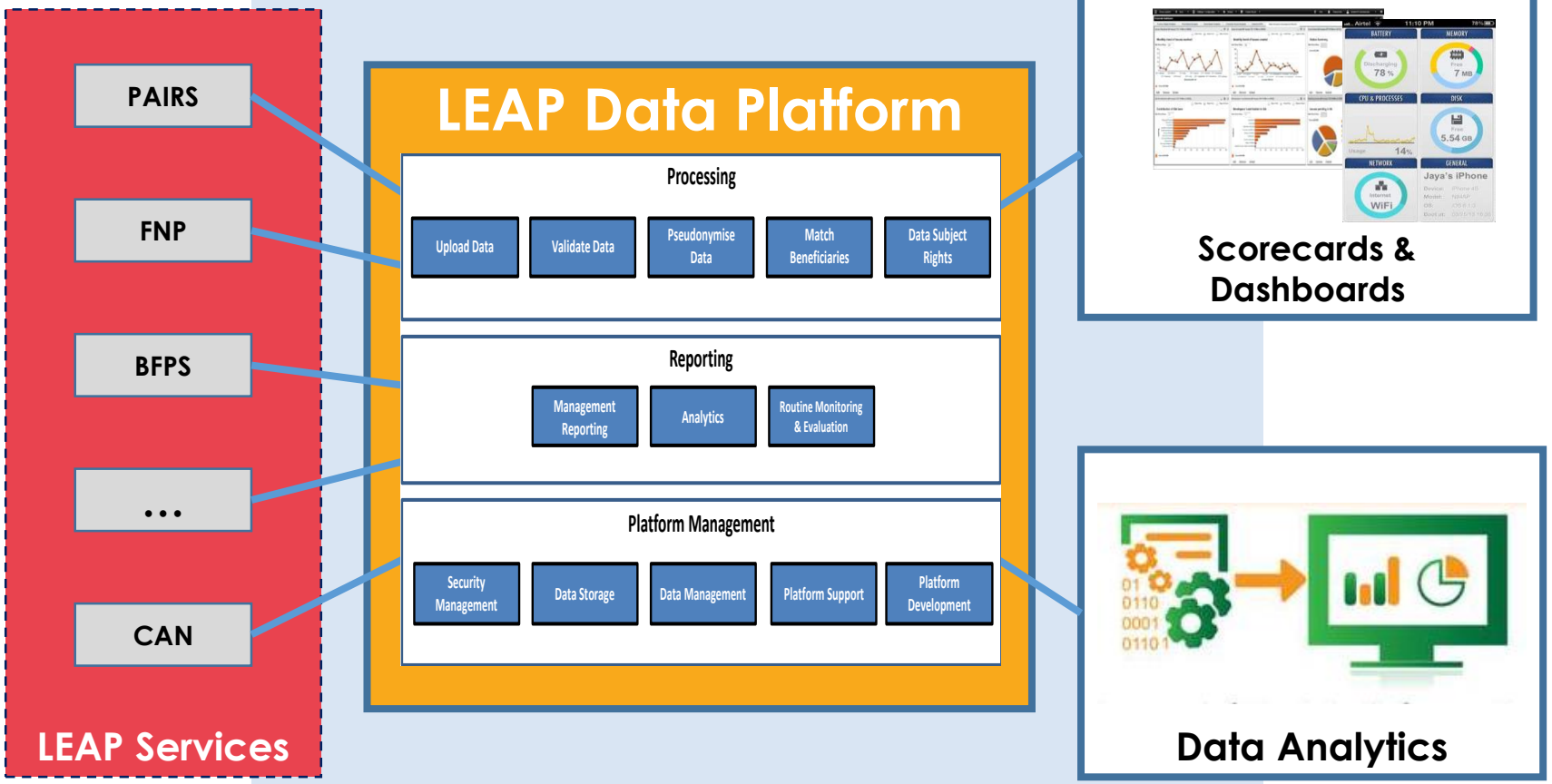
Two approaches to pseudonymisation

- Discussions on data sharing with NHS Trusts have indicated that data will only be shared if pseudonymised at source
- The majority of non-NHS services are expected to provide data in the clear
- Two approaches are needed, but both must use the same method & algorithm



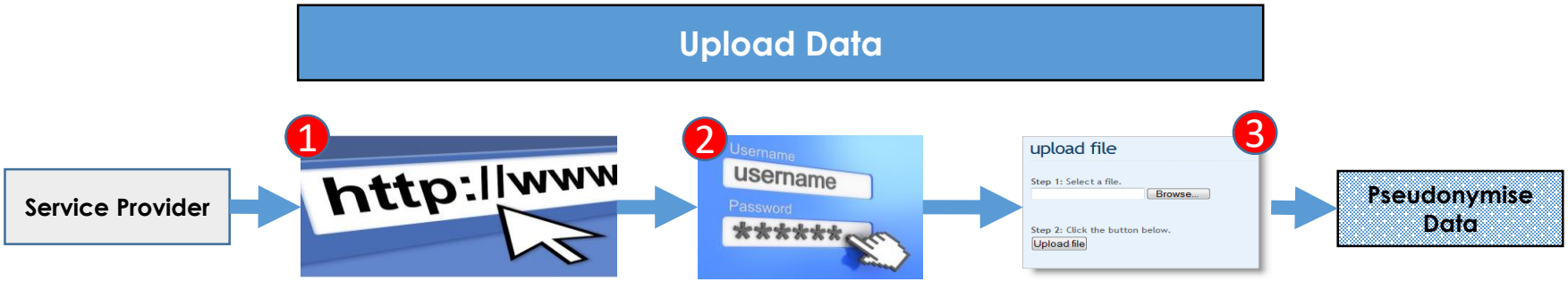
A. Services where pseudonymisation at source is possible
B. Services where pseudonymisation at source is not possible
C. Restricted LEAP environment to apply pseudonymisation

Data Platform Overview



Scope of Data Platform Service Provider

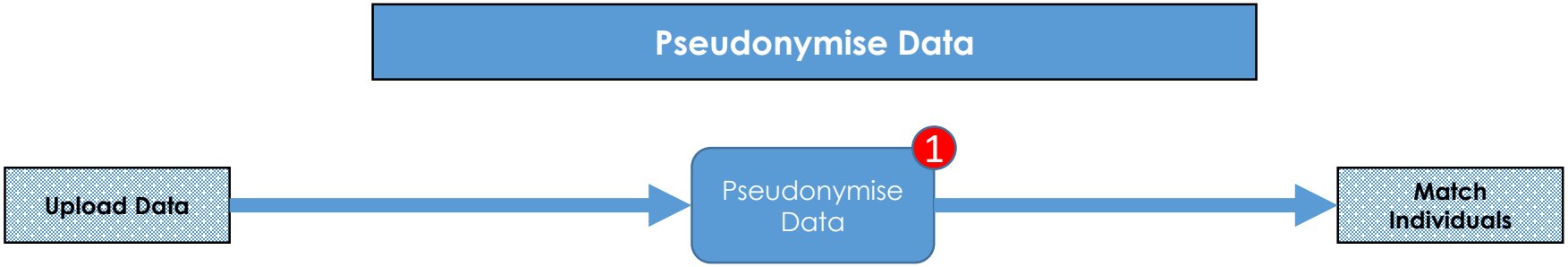
Processing – Data Upload



1. The Service Provider uses a browser to go to a specified URL.
2. The Service Provider enter credentials and uses factor 2 authentication.
3. The Service Provider selects their file to be uploaded from their environment.

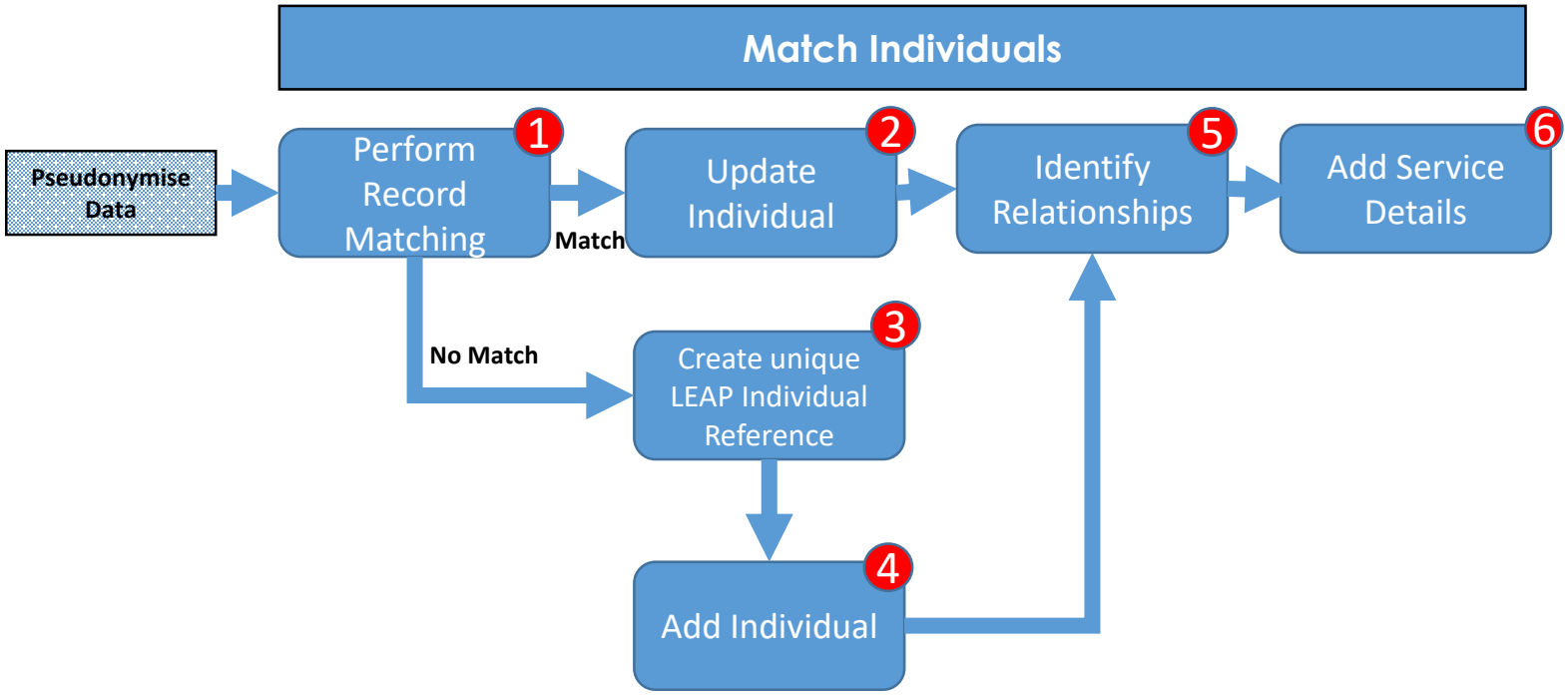
Uploaded data is mapped to the data platform requirements and validated. Some services require extra processing to fill in gaps in the data, this ensures a standard input into the data processing stage.

Processing Pseudonymised Data



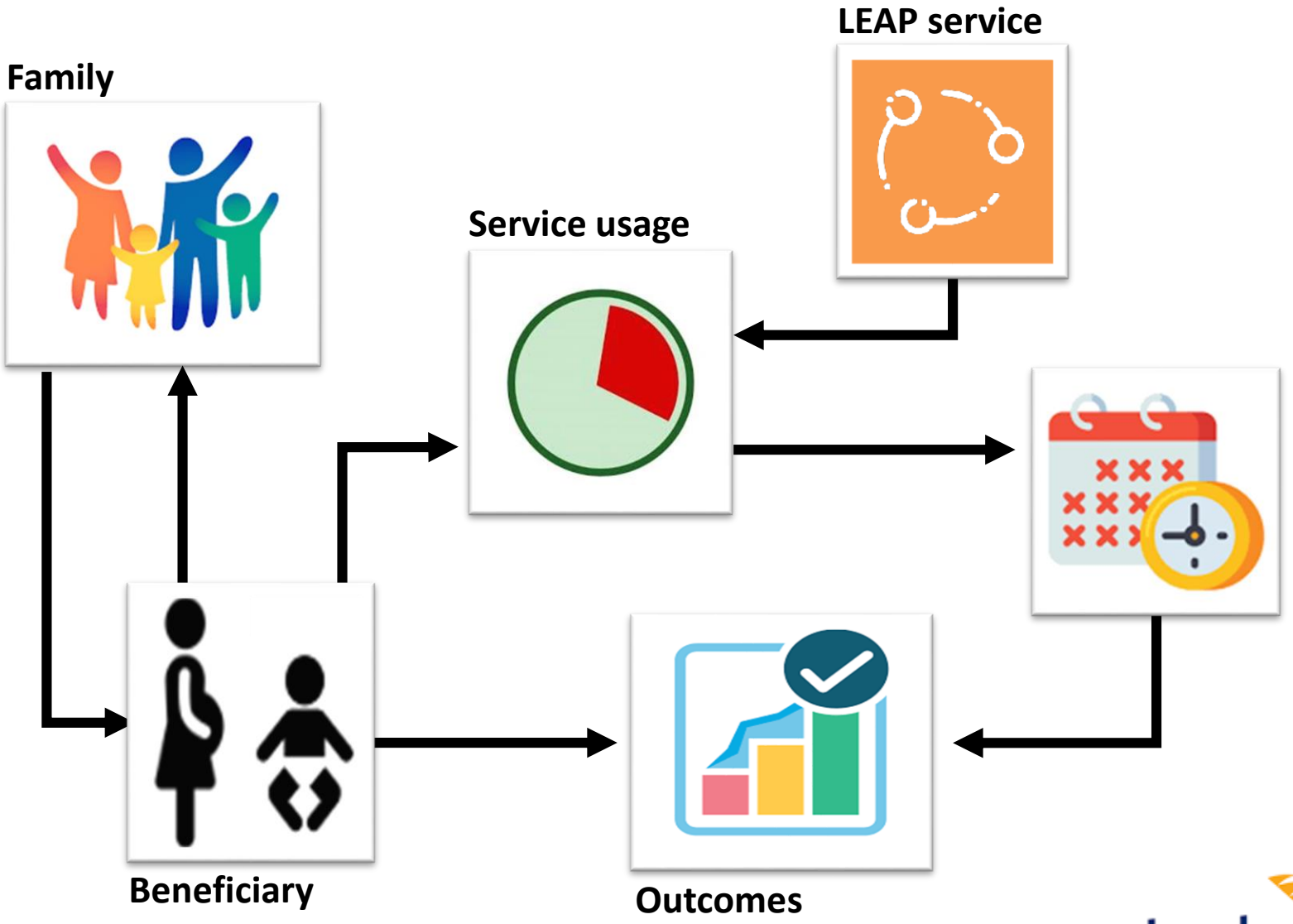
1. Pseudonymise Data
   - All useful 'unique identifier candidates' are used
     a) NHS Number
     b) For Child: primary carer email address, child date of birth, gender, first 3 characters of first name
     c) For Adult: email address
     d) Mobile phone number

Processing – Matching Beneficiaries



1. Perform record matching on main reporting data set
2. Match – update individual on main data set, including enriching with further key data
3. Create unique individual reference
4. Assumes unique individual, so new individual record is created
5. Identify any relationships to other individuals
6. Add service specific details

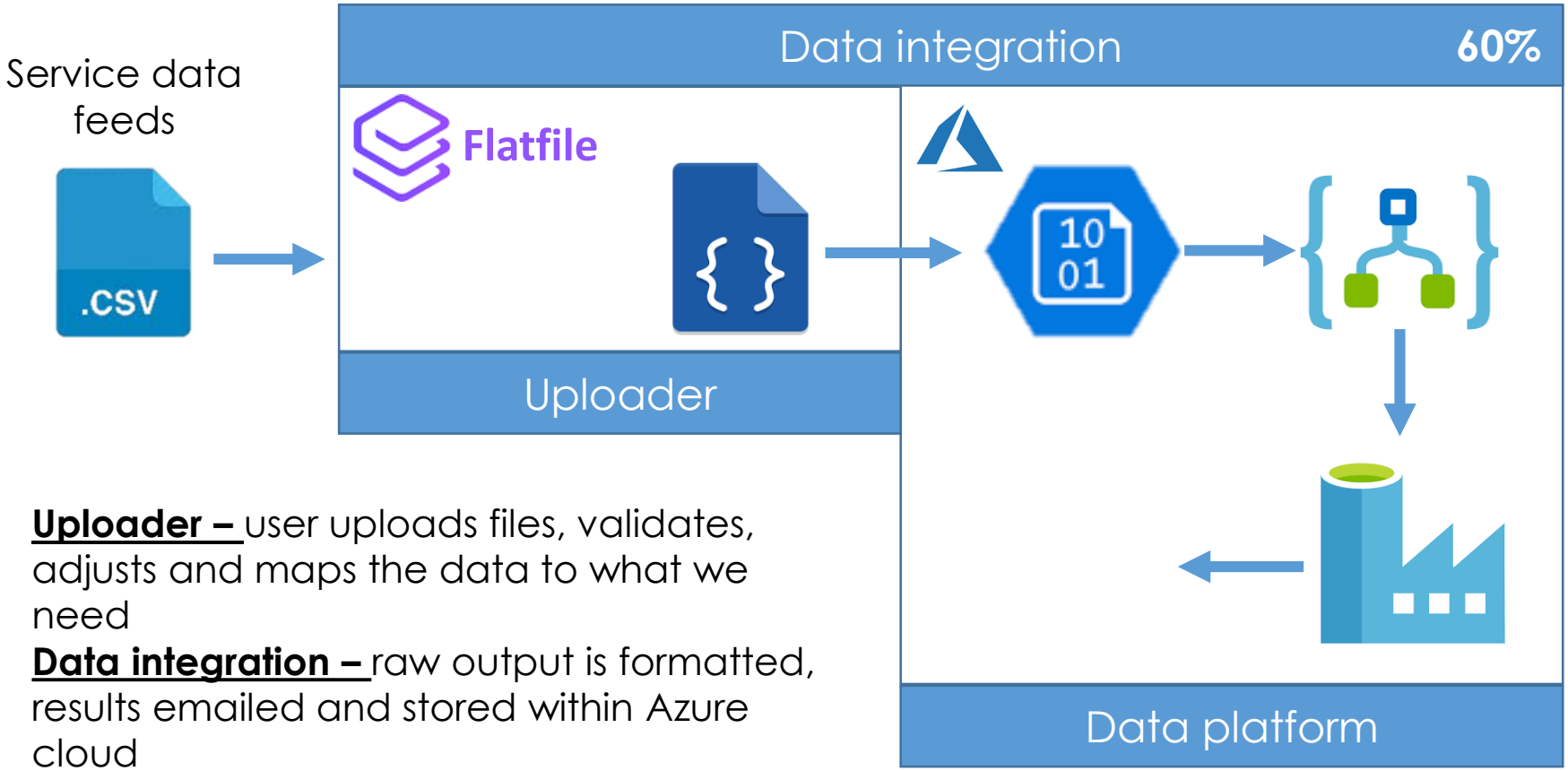
Target Data Model



Progress To Date

- Defined an aligned standard dataset across the programme relating to defining reach.
- Information Governance agreement has been achieved with all three NHS Trusts.
- Data sharing agreements are now in place with the majority of service providers, including all relevant services within NHS Trusts.
- Creation of solutions for both pseudonymisation at source (via a desktop application or SQL Server plug-in) and at destination via the platform.
- Key relationship with Lambeth Data, Analytics & Insight team has being developed:
  - The team will manage and support the data platform in production.
  - Knowledge transfer from the developer is well underway, with the team already taking part in configuring the new staging environment.
  - Service Level Agreement between the team and LEAP has been established.
- Next two slides have more detail on the platform build and service onboarding process.

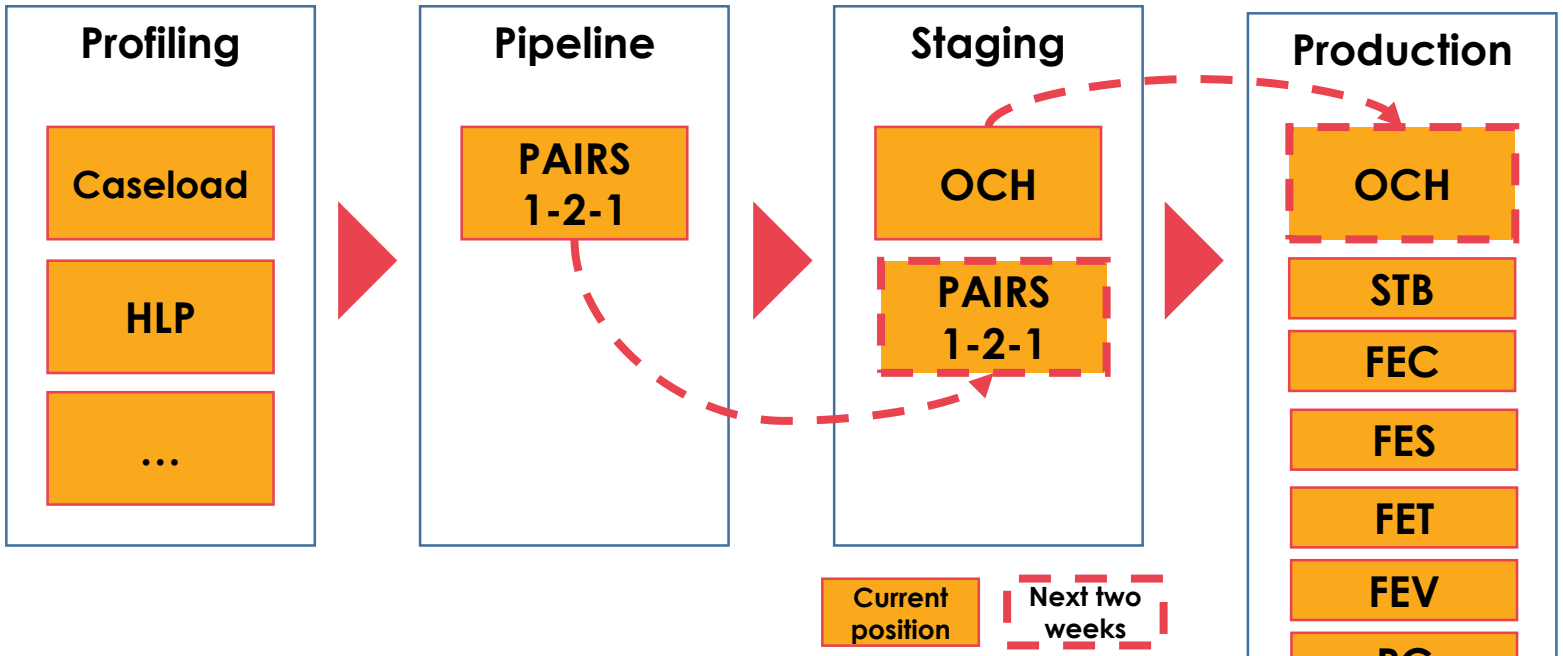
Developing the Platform and Progress



- **Uploader –** user uploads files, validates, adjusts and maps the data to what we need
- **Data integration –** raw output is formatted, results emailed and stored within Azure cloud
- **Data platform –** imports, pseudonymises, matches, updates database and makes data available for reporting

Service files can now be processed through the end to end process

Onboarding Pipeline



- **Profiling –** Checking the quality of data, feedback to services, corrections and re-check
- **Pipeline –** Services that have been through profiling and are ready for onboarding
- **Staging –** An environment that has a live setup but which can easily but scrubbed and the data re-loaded, used for checking data load is an expected
- **Production –** live database with ability to create reach figures

Next Steps

- Complete onboarding of last few user (demographic) datasets.
- Continue to work through engagement data specifics for remaining services that are not onboarded.
- Identify and resolve any data quality issues.
- Finalise different types of outcome measures and related codes, ensuring the proposed approach still works with each new measure.
- Update Engagement Reporting to show both service and programme level performance.
- Develop programme level dashboards for outcome data, and eventually for the Programme Level Theory of Change.
- Upload administrative datasets following discussions with IG leads in statutory organisations and NHS Trusts.
- LEAP team and Evaluation Partner to obtain direct access to platform data held in the Lambeth Council cloud.