



Delivering Better  
Outcomes by Linking Data,  
evidence and experiences

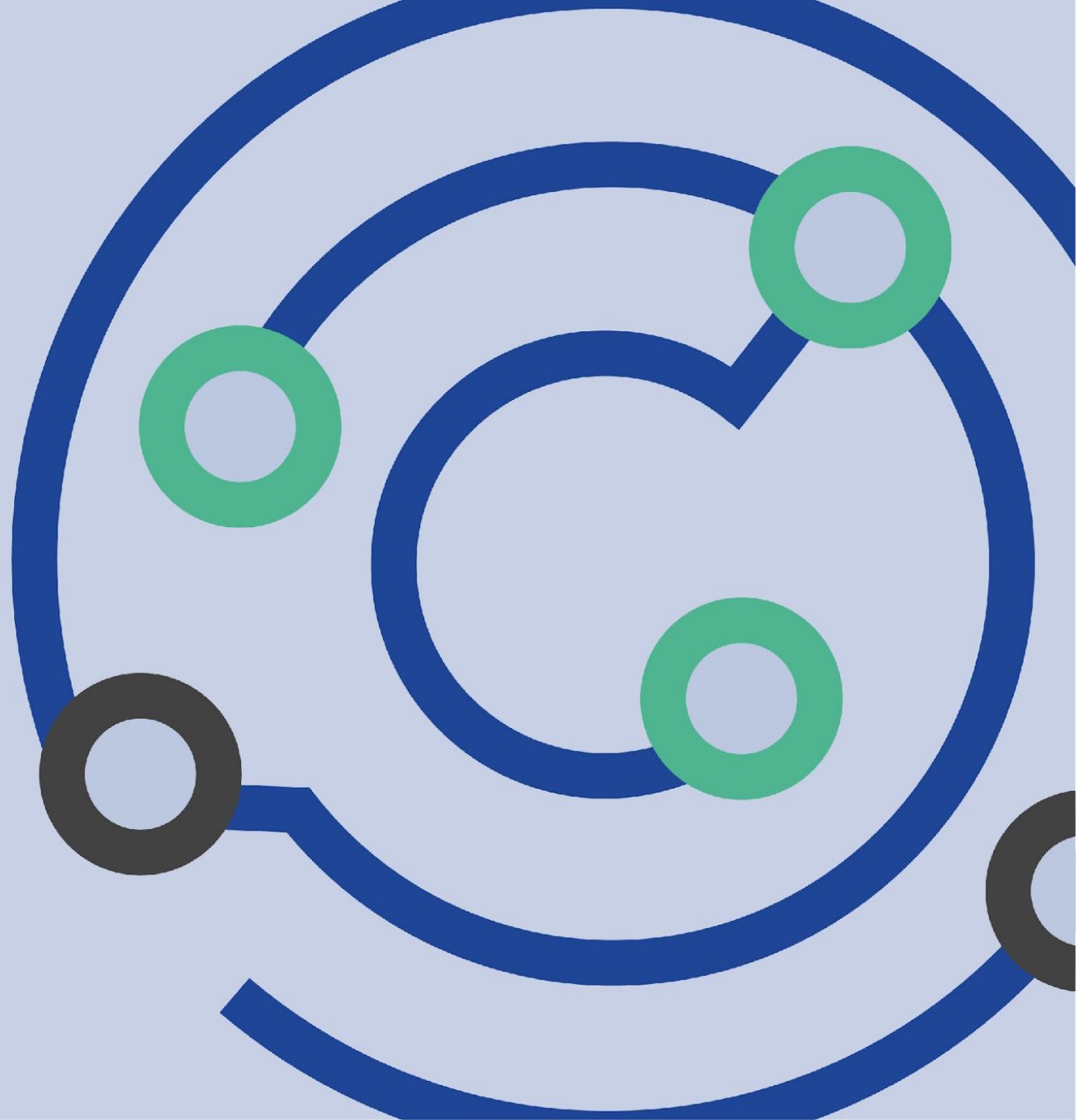
# BOLD Programme Overview

Presentation to Data Connect  
Launch Event

*Charlotte Owen – BOLD Head of Tech & Data, MoJ*  
*Maryam Ahmad – BOLD Pilots and Evaluation, MoJ*

May 2023

 HM Government



# Millions of vulnerable people in the UK could benefit from the right support – but the data relating to their needs and service provision are highly fragmented

Millions of people in the UK are classified as vulnerable at some points in their lives. These are the times when they need the most support from HMG, and also when well-targeted support has the best chance of genuinely improving lives



**12 million**

victims of crime, many of whom don't receive any support services



**600,000**

people with unmet needs for alcohol and drug treatment



Over **250,000**

people at high risk of homelessness



**275,000**

people requiring substance misuse treatment

However, support for vulnerable people is variable. Understanding of their needs and service provision is highly fragmented across multiple departments, agencies and arms-length organisations



**25%**

of people leaving custody are released without access to suitable housing



**65%**

of people leaving custody are not adequately picked up by community-based treatment programmes



**40%**

of rough sleepers have a drug dependency which is not picked up by healthcare providers



**29%**

of police investigations are closed through victim attrition

... and this is just the tip of the vulnerability iceberg.

# Why are we doing this?

The BOLD Programme has been created to join up disparate sources of data so that policy-makers have access to a more complete evidence-base on which to make decisions. BOLD is aiming to enable:

- ***Better understanding of the impact of services delivered in one part of government on outcomes in another through linking disparate datasets together***
- ***More effective joining-up of local and national services through understanding how services interact***
- ***Respond to unmet needs by analysing where people may not be accessing the services they need***

*For example, numerous organisations are involved in supporting offenders to turn their lives around, but each hold their own data. If we could combine these data sources we could transform our understanding of what works in reducing reoffending*



## BOLD Programme Vision:

Enable better-evidenced, joined-up and more effective cross-government services to support and protect vulnerable adults at specific touchpoints in their interactions with government-funded services, supported by timely linked data and evidence.

# Delivering Better Outcomes through Linked Data, Evidence and Experiences - BOLD



## £19.7 million

A three-year programme funded by HMT Shared Outcomes Fund



## Cross-govt

BOLD is made up of 80 staff in 4 departments, who work together as one team



## Multi-disciplinary

BOLD is made up of hybrid teams of policy people, analysts, tech specialists and others



## Pilots

Four pilot projects being delivered over three years – driving system change



## SPLINK

Technology solution to link and analyse datasets across departments.



## Data Protection by design

Full compliance with data regulation and law and an emphasis on Data Ethics

# BOLD comprises three key components



## Pilots & Use-Case Projects

The 4 Pilot areas each run multiple data linking 'use-case' projects – to unlock multiple 'use-cases' (i.e. decisions that could be enhanced by linked data)

**Substance Misuse Pilot**

**Reducing Reoffending Pilot**

**Victim Pathways Pilot**

**Homelessness Pilot**



## Enabling Functions

Teams/capabilities for delivering the technical aspects of the BOLD programme, supporting the pilot teams

**Data Sharing**

**Data Engineering**

**Data Linking**

**Strategy & Delivery**

**Evaluation**



## Wider Impact Projects

Broader projects aimed at improving how the process of sharing and linking data is done across government

**Public & Data Subject Engagement**

**Information Governance & Ethics**

**Data Sharing Approach & Framework**

**Knowledge Management**

**Communications**



# **BOLD is focused on four demonstrator pilot areas**

## **Homelessness**

The Homelessness Pilot project aims to better understand why some people repeatedly become homeless, which services are most effective in preventing prison leavers from becoming homeless (and thereby decreasing their chances of reoffending), and what role drug treatment services can play in preventing homelessness.

## **Victims Pathway**

The Victim Pathways Pilot project aims to better understand how government and third sector services can most effectively support victims of crime to cope and recover and to confidently seek justice. The ultimate goal is to use this evidence to improve victims' experience of the criminal justice system and support them to achieve a positive justice outcome.

## **Substance Misuse**

The Substance Misuse Pilot project aims to better understand what factors drive successful recovery from addiction for particular groups, such as prison leavers or rough sleepers. As part of this pilot, Welsh Government aim to ensure that challenges unique to Wales are also addressed.

## **Reducing Reoffending**

The Reducing Reoffending Pilot project will support the Government's ambitions to tackle the root causes of reoffending by linking data to better understand the impact of specific interventions to help offenders turn their backs on crime, particularly in terms of their linked outcomes (employment, health, housing and family).



# The BOLD Programme has been designed to unlock specific 'use-cases'

## What is a use case?

A 'use case' is 'a policy, management or operational intervention where linked data is needed that targets a clear outcome for vulnerable populations'.

Use cases fall into three categories:

- **Policy and research analysis** (national policy, stats etc)
- **Planning and management analysis** (local resource allocation and service planning)
- **Operations** (front line activity)

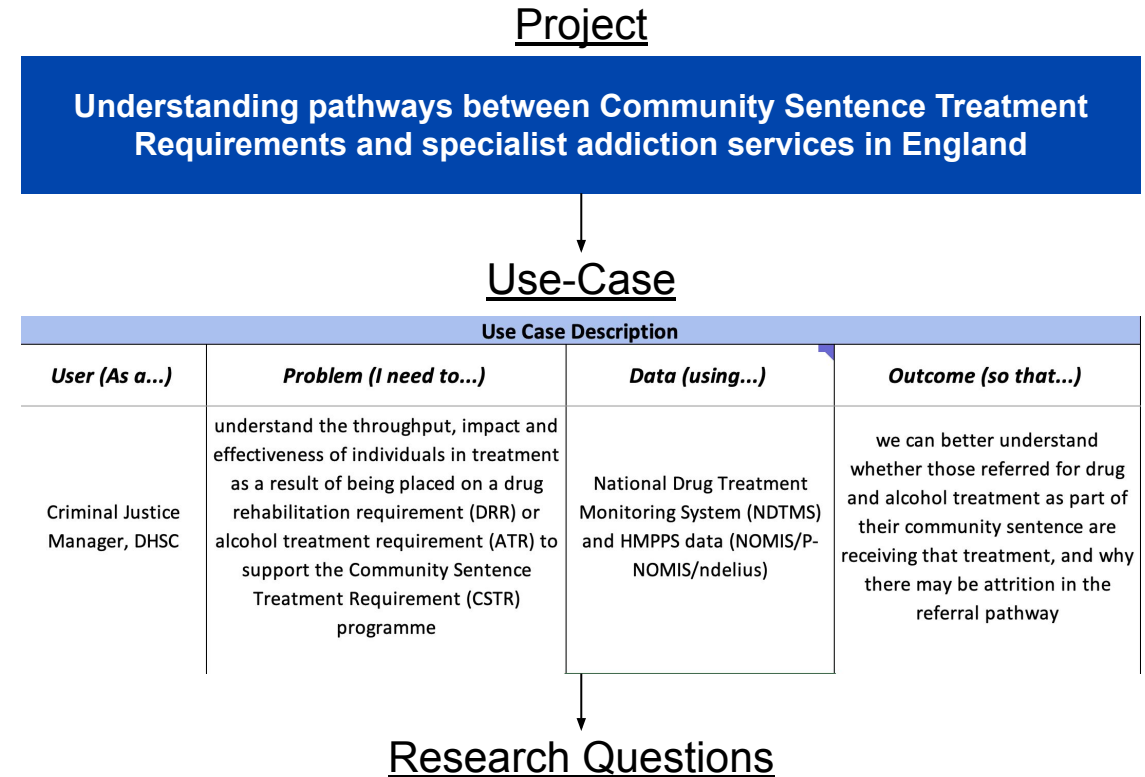
Does not require personal identifiers

Requires personal identifiers

## How data linking use cases were developed

- Use cases were sourced for each of the BOLD pilots through extensive stakeholder engagement and user research.
- We developed a non-exhaustive long-list of use cases, which were then prioritized based on an assessment of impact vs feasibility and form the basis of BOLD's various projects
- Of these use-cases, all but one\* are to support policy/research, and therefore do not require PII

## Example of how a project is built upon use-cases



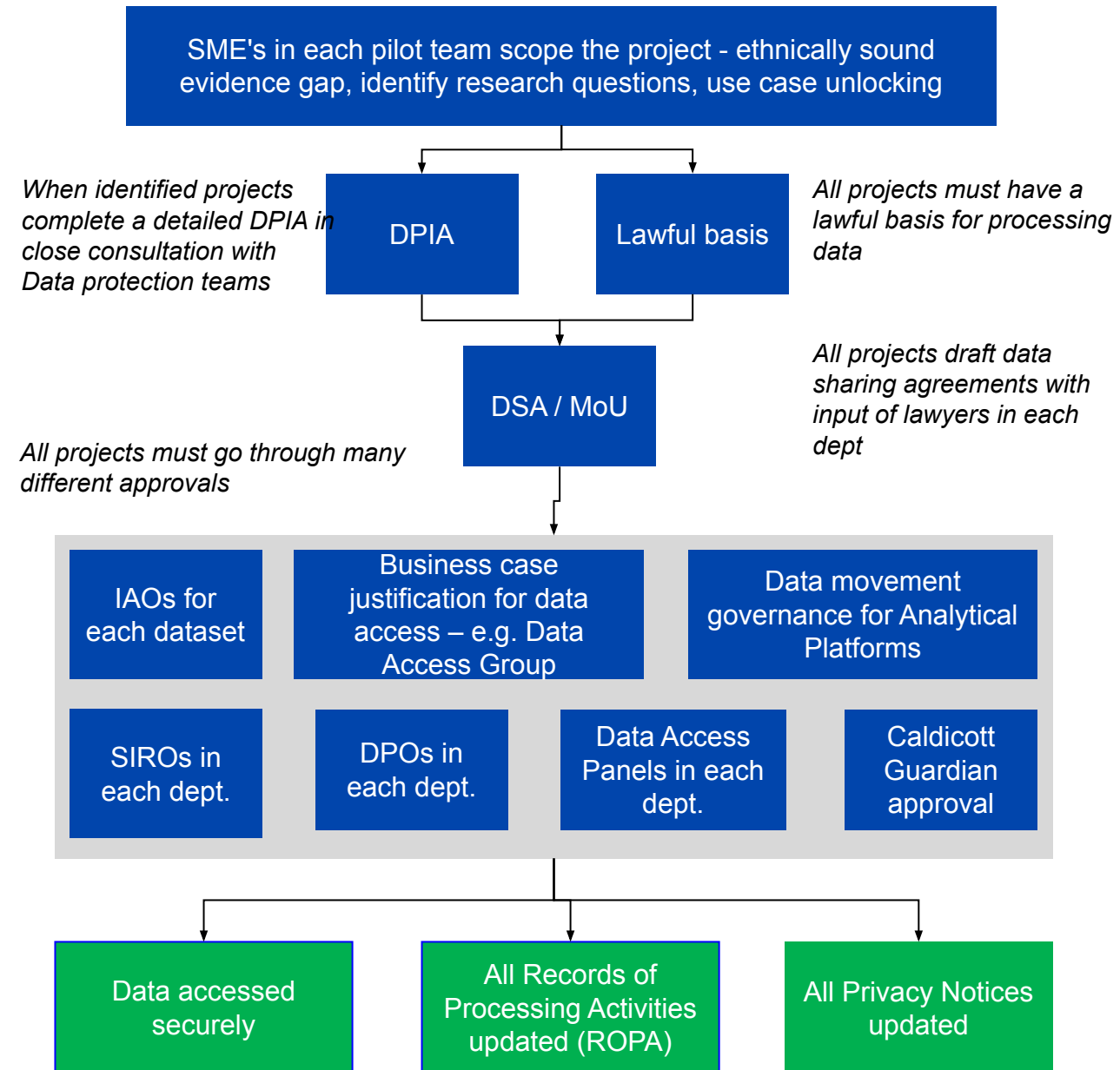
- E.g. What is the impact of CSTRs on the successful completion of community orders?
- E.g. What is the time lag between sentence start date and DRR/ATR start date?
- E.g. What is the latest status of treatment for those on sentenced to DRR/ATR (i.e. successful completion vs dropout)?

# BOLD's Approach to Data Sharing

## All data sharing undertaken through BOLD is fully compliant with the ICO's Code of Practice for Data Sharing

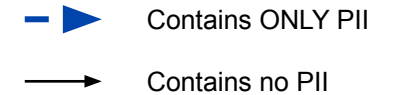
- Personal information is only shared where there is lawful basis under UK GDPR
- Data Protection Impact Assessments (DPIAs) and Data Sharing Agreements (DSAs) as minimum requirements
- A thorough set of governance processes
- Privacy notices

## A typical governance process for a BOLD Data share

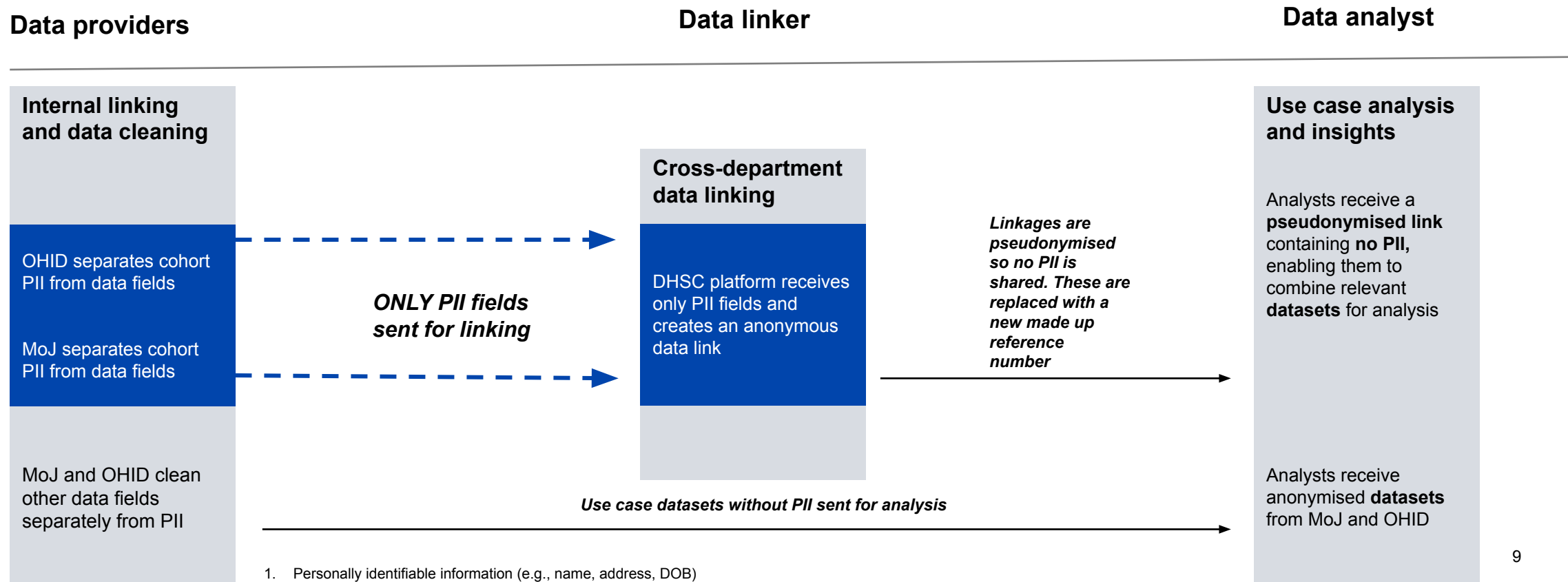




# To preserve privacy when linking data, BOLD separates and secures identifying information



An example combination of OHID and MoJ data to demonstrate the separation of PII<sup>1</sup> before linkage and analysis



1. Personally identifiable information (e.g., name, address, DOB)

# BOLD Data Linking with [Splink](#)

Linking records between (or within) datasets can be challenging due to a lack of unique identifiers such as a National Insurance or NHS number.

With Splink, we compare records using the available PII and estimate the **probability of a match** based on the **similarity pattern** of these identifiers.

## Example

ID	first_name	surname	dob	city	email
A	lucas	smith	1984-01-02	London	lucas.smith@hotmail.com
B	lucas	smith	1983-02-12	Manchester	lucas.smith@hotmail.com
	✓	✓	✗	✗	✓

Splink uses many comparisons like this to build a simple statistical model. That model then tells us that if **first name**, **surname** and **email** are a match, we can be almost 100% confident that they are the same person.



- Award-winning software package
- Open-source – downloaded 3 million times globally
- Recommended ONS Tool

## Use case analysis and insights from linked data

### Linked data to drive evidence-based policy

Across all our pilots, we will be analysing linked data to better understand the effectiveness of government programmes/initiatives to help inform policy decisions.

The data linked through BOLD will be used as a resource for evaluating government policies/programmes, and may also form national stats.

### Linked data to support operational decision making

We have one project looking to use linked data to support operational decision making

The reducing reoffending pilot includes an operational project which aims to help probation staff manage offenders after release from prison, by linking information on an offenders' needs and possible risks without the need for manual collection.

We will **not** be using predictive analytics to make decisions about individuals in any of our projects.

# Engaging the public, transparency and ethics

## BOLD's Approach to Public Engagement & Transparency

- ✓ Engaging data subjects
- ✓ Engaging organisations outside of government
- ✓ Public Engagement Advisory Group
- ✓ Proactive external comms

## BOLD's Approach to Data Ethics

- ✓ Data Protection by design and default
- ✓ BOLD follows CDEI's Components of Responsible Innovation model
- ✓ Ethical assessments

# Questions