



Ministry
of Justice

DATA FIRST



Ministry of Justice Data First Showcase

Ministry of Justice

20 March 2024



What is Data First?

- Ambitious data-linking, academic engagement and research programme led by the Ministry of Justice and **funded by ADR UK** (Administrative Data Research UK).
- Data First is unlocking **the potential of the wealth of data** MoJ already collects when people interact with the justice system.
- Administrative datasets from the courts, prison and probation services are linked to **enable analysis of our users' journeys, interactions and pathways** across the system and with a range of other public services, including education and employment.
- Enabling researchers across government and academia to access **datasets in an ethical and responsible way** via the ONS Secure Research Service and the SAIL Databank
- Working in **partnership with academics** to develop new evidence to shape government policies and drive real progress in **improving justice outcomes**.



Data First datasets

Criminal justice

- Ministry of Justice Data First magistrates' courts defendant
- Ministry of Justice Data First Crown Court defendant
- Ministry of Justice Data First prisoner custodial journey
- Ministry of Justice Data First probation
- *Ministry of Justice Data First Offender Assessment System (OASys) – In progress*

can link on
defendant and
on case

focus on the
defendant /
offender – can link
on person

NEW cross-justice system
linkage, can link on person
(including multiple parties
in civil & family cases)

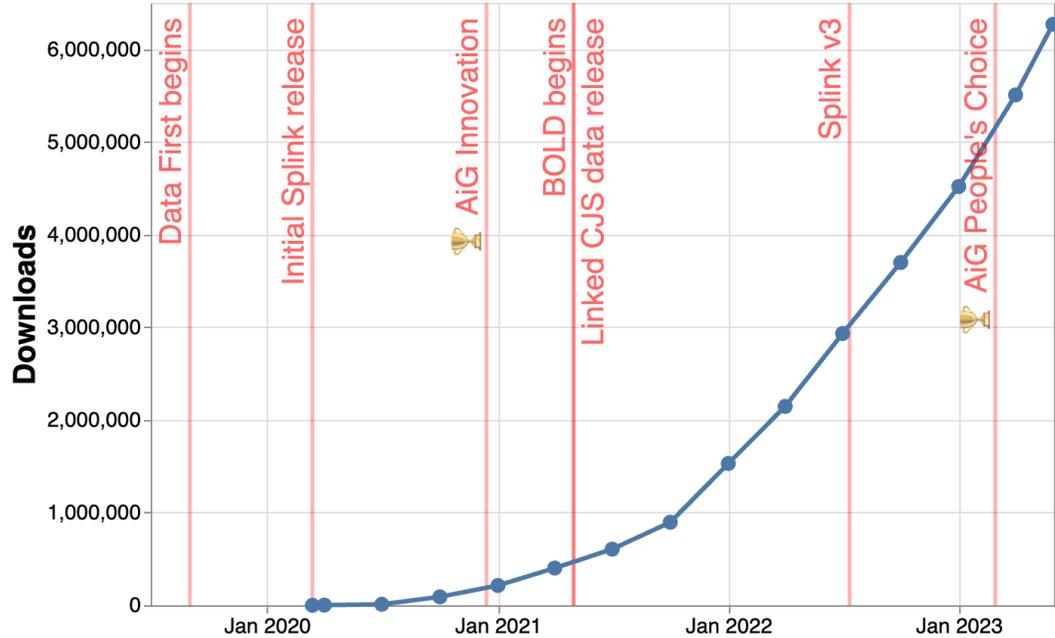
Other court jurisdictions

- Ministry of Justice Data First family court
- Ministry of Justice Data First civil court
- *Ministry of Justice Data First linking dataset(s)*

Linked to other government department's data

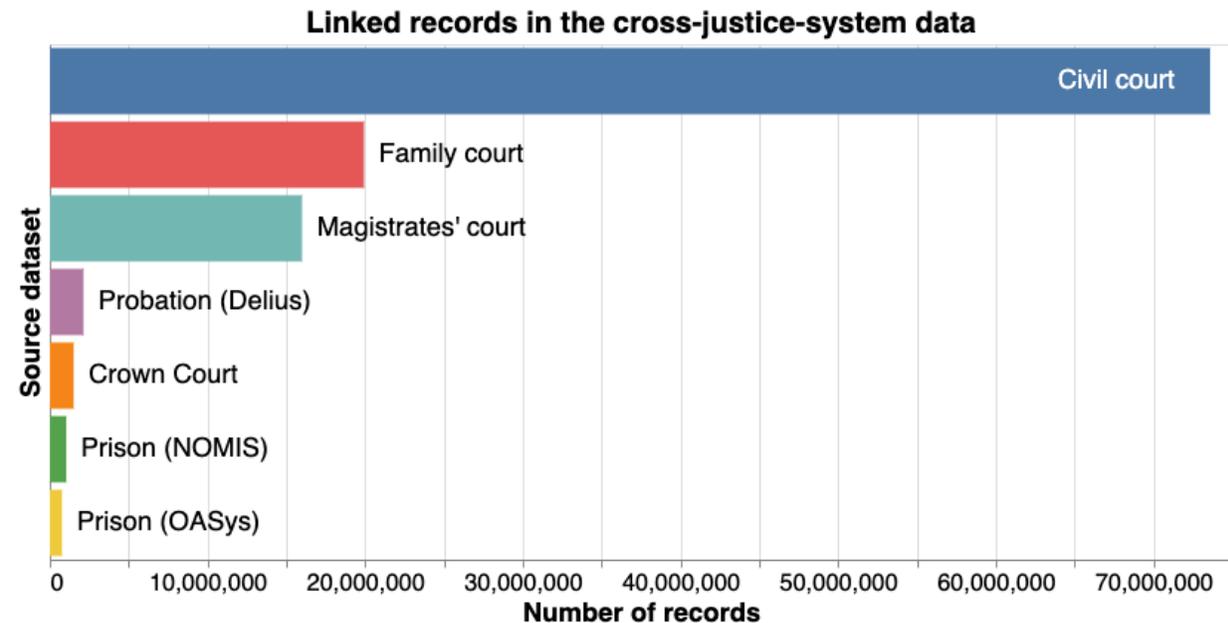
- MoJ-DfE data share joining the Police National Computer (PNC) to the National Pupil Database (NPD) } Separate linkage

Splink: Transforming data-linking methodology and capabilities



- **6.3 million downloads** of our Splink software (100,000/week)
- Recognition in the **Analysis in Government Awards**
- Underpins data linkage work at MoJ and beyond. The recommended data linkage tool in the **ONS Data Linkage Hub** (as used for Census 2021)

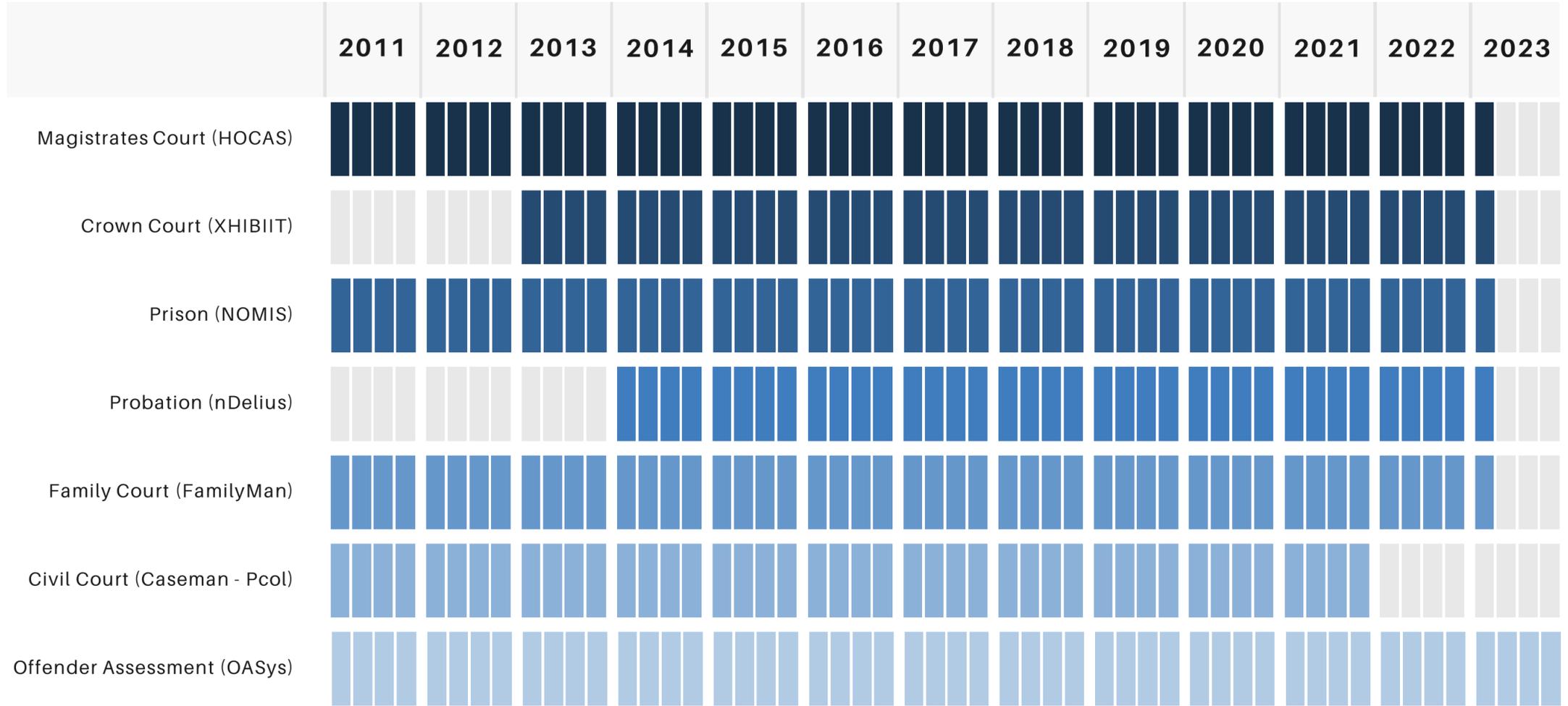
- Over **110 million person records** (without unique IDs) linked across the justice system
- Identified individuals and their journeys through and between justice services
- **This has never been possible before**



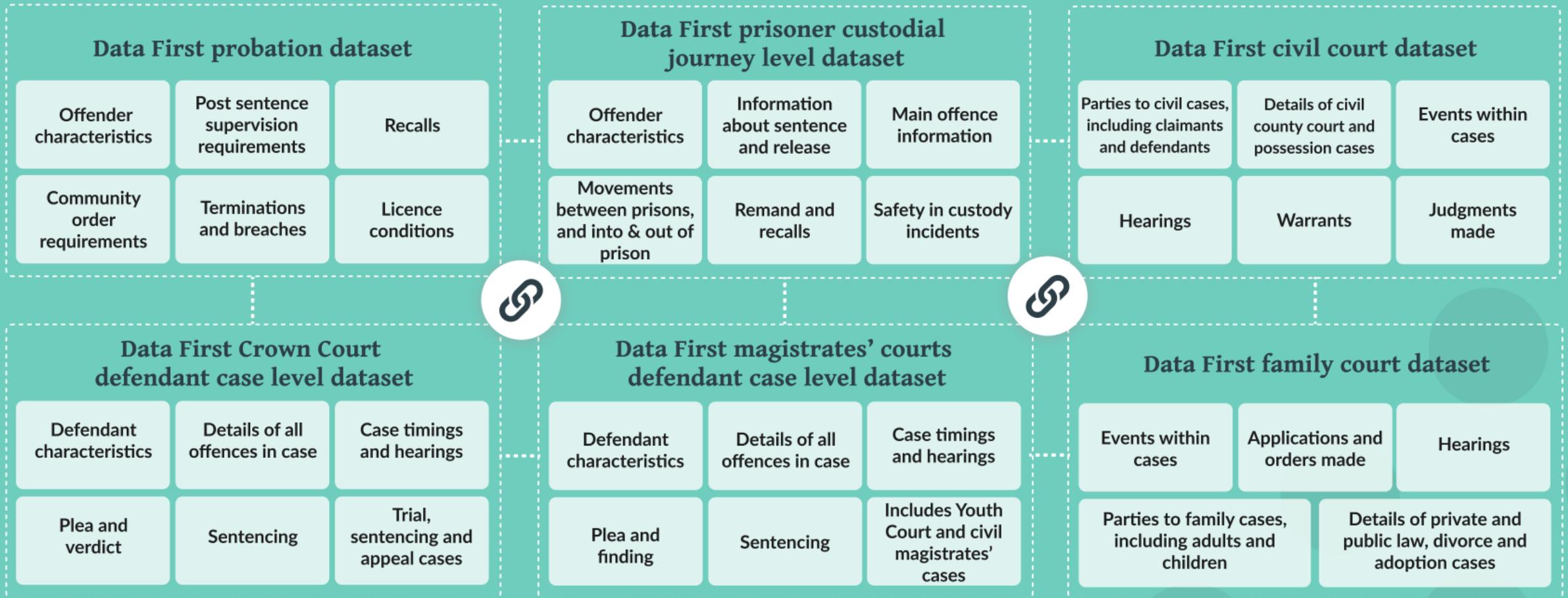
Dataset Overview – Cross Justice



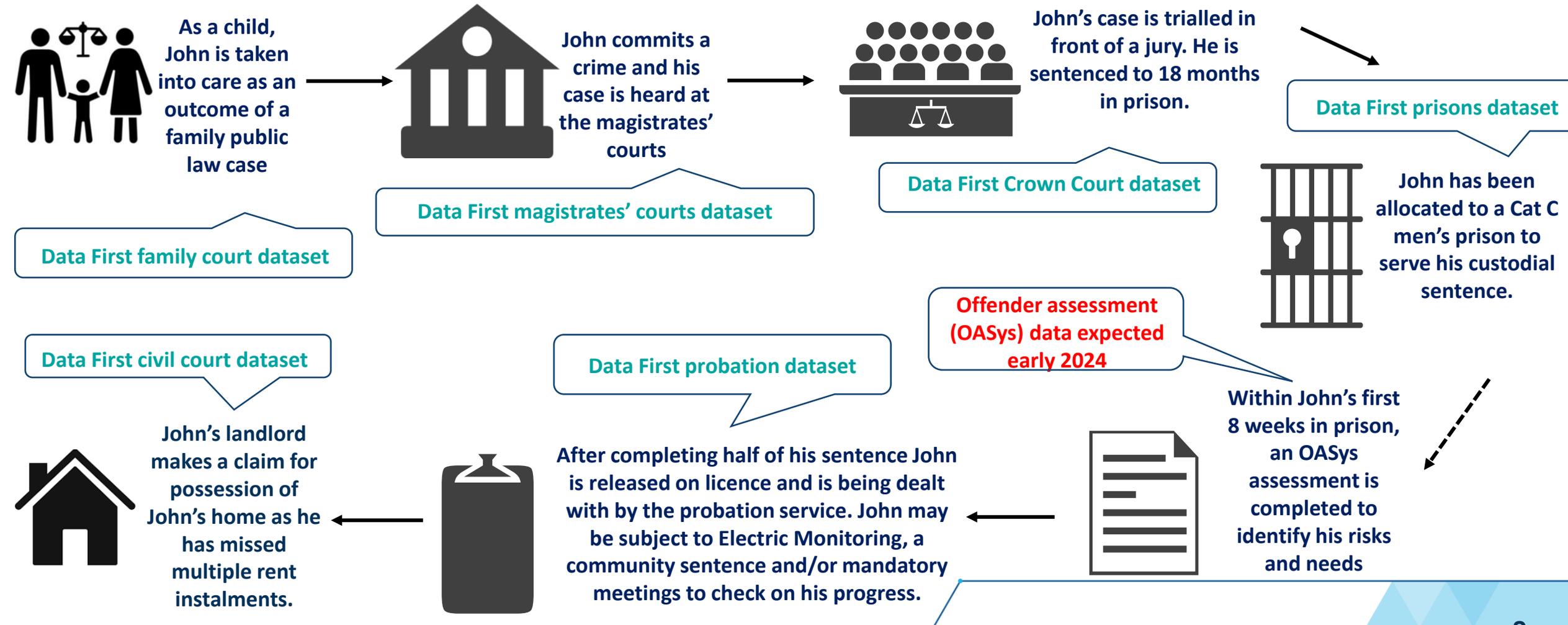
Coverage Periods



Information contained in the Ministry of Justice Data First: Cross-Justice System - England and Wales linked dataset



Data First cross-justice system data - user journey



Datasets overview – MoJ-DfE data share



The MoJ DfE data-share

What are the links between educational attainment and prolific vs. time-limited offending?

What is the influence of early education or children's social care on later prison outcomes including length of prison sentences and re-offending?

What are the short- and long-term effects of exclusion on educational attainment and offending?

What can it tell us?

- Increases our understanding between **childhood characteristics, education outcomes and (re)-offending.**
- Includes data on young people's **educational attainment** and characteristics such as free school meal eligibility, Special Educational Needs, absence and exclusions linked to the Police National Computer data on their **criminal histories** and interactions with the criminal justice system.

Is there an association between particular interactions with the education system and offending, and if so, is one of these factors typically the driver?

Are specific interventions more successful on offenders depending on their education attainment?

What are the protective and risk factors for the likelihood of committing crimes and how do these differ across crime types?

An example MoJ and DfE user journey



Achieved level 4 in English and Maths at KS2



Identified as having Special Educational Needs



Permanently excluded from school



Child is subject to a care order

58% of those cautioned or sentenced for a serious violence offence achieved a level 4 or above in Maths, and 51% achieved the equivalent in English.

44% received their first permanent exclusion over a year before their first serious violence offence

New for 2024: Include all episodes of social care, this will allow us to explore the sequencing of care experience and offending.

New for 2024: Prison education data



Released from prison



Completes a construction course in prison



Sentenced to prison



Commits a serious violence offence

Evidence gaps and research priorities



Ministry of Justice

Areas of Research Interest (2020)

IMPROVING JUSTICE OUTCOMES



through innovative research, data, and analysis

Applying for access to Data First datasets



Applying for access to Data First datasets

Three steps for accessing data:

- 1) Become an [accredited researcher](#).
- 2) Apply for access to the specific data required for a research project from relevant data owners using the MoJ and HMCTS '[Secure access to data](#)' combined form for access to data held in the ONS SRS or complete the [scoping form and Information Governance Review Panel \(IGRP\) form](#) for data held by SAIL Databank.
- 3) Complete the [Research Project Application](#) form and receive final approval from UKSA RAP (not required for MoJ-DfE data share).

Researcher support



What you can expect from us

- Advice and guidance on producing a successful application
- Provision of detailed data catalogues
- Availability to provide insights into dataset queries throughout the project
- Timely review and clearance of outputs
- Informal academic peer review
- Opportunities to present and share research

Q&A Part 1



Examples of Past Research



Past Research using Data First Datasets

The British Journal of Criminology, 2024, XX, 1–22
<https://doi.org/10.1093/bjc/zaae005>
Original Manuscript

OXFORD

Ethnic Inequalities in Sentencing: Evidence from the Crown Court in England and Wales

Kitty Lympelopoulou 

*K. Lympelopoulou, School of Society and Culture, University of Plymouth, Drake Circus, Plymouth PL4 8AA; email: Kitty.Lympelopoulou@plymouth.ac.uk

In recent years, there has been considerable policy and academic interest in the existence of ethnic inequalities in the Criminal Justice System. A large body of sentencing research has been dedicated to exploring whether ethnic minority defendants are treated more harshly than similarly situated white defendants. This paper extends this research utilizing Ministry of Justice linked criminal justice datasets and multilevel models to assess the effect of ethnicity and other defendant case and contextual factors on sentencing outcomes in the Crown Court. The analysis shows that legal characteristics such as plea, pre-trial detention, offence type and severity are important factors determining sentencing outcomes although they do not fully explain disparities in these outcomes between ethnic groups. Ethnic disparities in imprisonment persist and, in some cases, become more pronounced after controlling for defendant case and court factors. In contrast, ethnic disparities in sentence length are largely explained by legal factors, and after adjusting for other predictors of sentencing outcomes, observed differences between most (but not all) ethnic minority groups and the white British disappear.

Past Research using Data First Datasets



The image shows the cover of a report. At the top left is the ADR UK logo, which consists of two overlapping circles (one white, one dark) and the text 'ADR UK'. At the top right is the text 'Data Insight'. The main title 'School funding, pupil performance and crime' is centered in a large, bold, black font. Below the title, the author 'Author: Dr Will Cook' and the date 'Date: October 2023' are listed in a smaller font. The background is a teal color with a pattern of faint, light-colored circles.

This Data Insight summarises findings from working with the Ministry of Justice & Department for Education linked dataset - England. Using data from the National Pupil Database and the Police National Computer, the research analyses the effect of school funding programmes on crime. The data was made available through the Ministry of Justice Data First initiative funded by ADR UK.

Past Research using Data First Datasets



Data First: Criminal Courts Linked Data

**An exploratory analysis of returning
defendants and the potential of linked
criminal courts data from 2011 to 2019
in England and Wales**

**Tom Jackson, Caris Greyson, Dr Ian Rickard and
Professor Andromachi Tseloni**
Ministry of Justice

Ministry of Justice Analytical Series
2022



Ministry
of Justice

Data First Evaluation Fellowship

Using linked administrative data to evaluate policy interventions in the justice system

Dr Georgina Mathlin

Postdoctoral Evaluation Fellow, Centre for Psychiatry and Mental Health, Wolfson Institute of Population Health, QMUL

g.mathlin@qmul.ac.uk / g.mathlin1@justice.gov.uk

Background to the Evaluation Fellowships

- A collaboration between the Ministry of Justice, the Economic and Social Research Council (ESRC) and the Cabinet Office Evaluation Accelerator Fund (EAF) to explore the feasibility of using **administrative data to evaluate policy and practice interventions in the justice system**.
- Two academic Fellows, **Ian Brunton Smith** (University of Surrey) and **Georgina Mathlin** (Queen Mary University of London) appointed in February 2023.
- To use **quasi-experimental methods on linked justice datasets** to establish whether specific policies and practices are effective in delivering intended departmental outcomes of **reducing reoffending, protecting the public and delivering swift access to justice**.
- Ian and Georgina are embedded within MoJ Data & Analysis as part of the **Evidence & Partnership Hub** commitment to **maximise the use of academic expertise** to address evidence priorities as set out in the MoJ **Areas of Research Interest (ARI)**.
- 18-month project, co-funded by the ESRC and EAF, which began with a **three-month scoping phase** to identify key policies and interventions and **priority evaluation questions** to shape the evaluation programme. The **12-month implementation phase** began in June 2023.

Reflections on using administrative data to evaluate justice services

Home > News & publications

Reflections on using administrative data to evaluate justice services

Categories: [Blogs](#), [ADR UK Research Fellows](#), [ADR UK Partnership](#), [Crime & justice](#)

15 March 2024

Dr Georgina Mathlin is undertaking a [Cabinet Office-funded Evaluation Fellowship](#) at the Ministry of Justice (MoJ), a jointly led opportunity by the Economic and Social Research Council and MoJ. This blog reflects on her experiences of working with administrative datasets made available through the MoJ Data First programme (funded by ADR UK).

There is a vast amount of administrative data collected across the justice system, which has been made available through the

Scoping Phase

1. Initial call for key interventions/policies that may benefit from being evaluated using administrative data
2. Initial meetings set up to discuss proposed project
3. Initial onboarding with the Analytical Platform and accessing MoJ linked datasets as jumping off point for prioritisation phase
4. Data deep-dives to explore feasibility of proposed projects
5. Follow up meetings and project mapping

Evaluation programme

1. Evaluating the use of combined punitive (UPW) and rehabilitative order requirements (RAR)

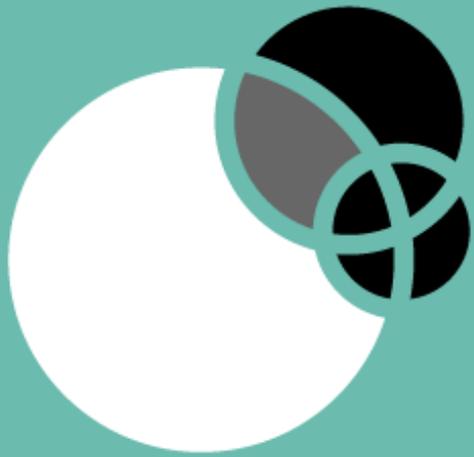
- Who is given a combined order requirement compared to either requirement alone?
- Are there any differences in requirement outcomes for those given a combined order requirement compared to an individual?

2. Evaluating the impact of ROTL 2019 policy on offender characteristics, ROTL policy implementation and re-entry to the criminal justice system

- What were the 2019 ROTL policy changes intended outcomes?
 - Did the policy change have an impact on offender characteristics?
 - Did the policy change have an impact on the policy implementation?
 - Did the policy change impact criminal justice outcomes (time to re-entry into the CJS)?

Reflections

- 😊 Ability to link datasets together to understand offender journeys through the CJS
 - 😞 Can be challenging – often multiple offences so finding the right information across each dataset to link together can be difficult
- 😊 Large amount of available data
 - 😞 Administrative data takes time to pre-process into a ‘research-ready’ state
- 😊 Minimal missing data
 - 😞 Important to identify and explore patterns of missingness where they do exist
- 😊 Improved data and analysis skills
- 😊 Learning from MoJ policy-facing analysts
- 😊 Due to scoping phase the projects have clear policy implications which can directly reach policymakers



ADR UK
Data-driven change

A national data linkage study to assess the extent, nature and outcomes of ‘serious and organised crime’ prosecuted before the criminal courts in England and Wales.

Dr Tim McSweeney

ADR UK Data First Research Fellow



About the project

Secondary analysis of 12.6 million de-identified, individual-level, linked records from the criminal courts and prison system in England and Wales (E&W) over an eight-year period (2013-2020).

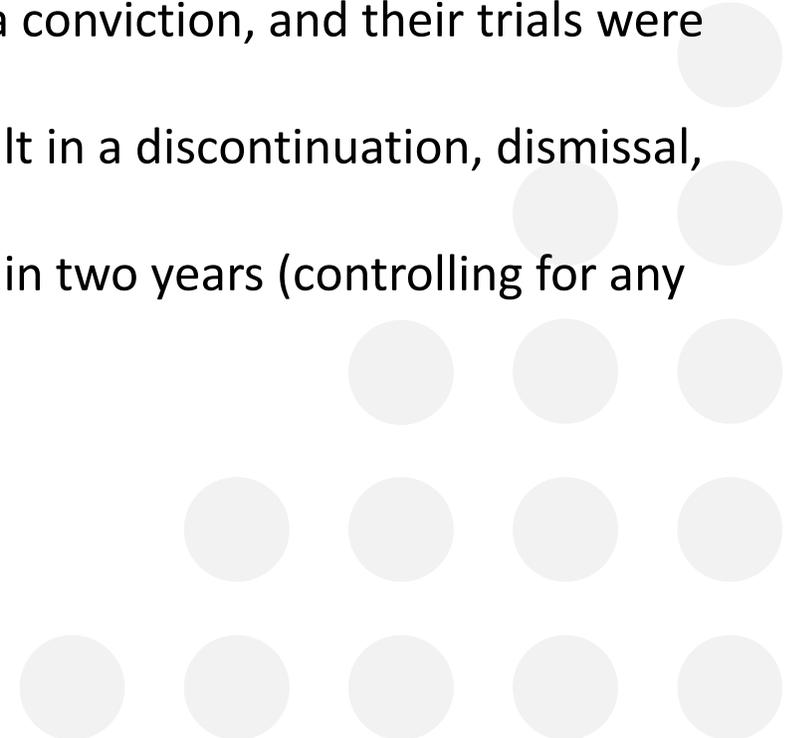
It sought to describe the **extent, nature** and **outcomes** of appearances and cases heard before the Crown Court in E&W considered to be linked to 'serious and organised crime' (SOC).

Drawing on a definition developed and used by previous research (Home Office, 2013; Ashby 2015), **SOC was defined** in the current project as any prosecutions for offences considered to:

- involve a degree of ***planning, control and co-ordination***,
- which would attract a ***minimum custodial sentence of three years*** upon conviction, and
- where ***co-defendants were involved***.

Findings indicate that...

- SOC defendants are predominantly men, White, and aged in their early 30s.
- On average, SOC appearances generated higher levels of crime-related 'harm'.
- SOC appearances were concentrated in certain locations.
- Compared to other Crown Court appearances, SOC was typically linked to more affluent areas.
- SOC appearances were more likely to involve a guilty plea, result in a conviction, and their trials were longer.
- Those SOC appearances heard before a jury were more likely to result in a discontinuation, dismissal, or acquittal.
- Fewer SOC defendants reappeared in court for further offences within two years (controlling for any time spent in custody).



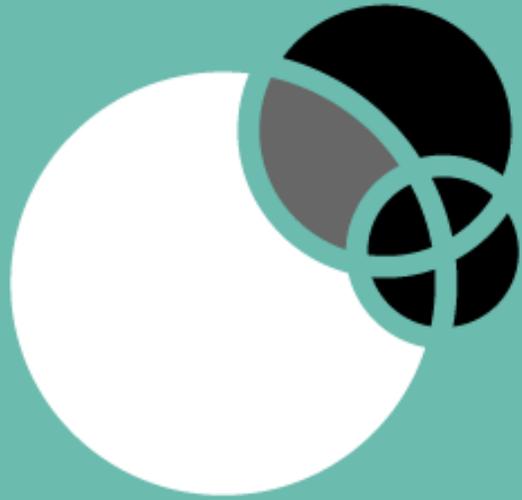
Finally, some limitations and implications

The main limitations encountered by the project were largely due to the absence of:

- a **dedicated SOC flag or marker** within the datasets used;
- information on the **full range of offences** being prosecuted;
- any data on **complainants and victims** (and areas being impacted); and
- details of **aggravating and mitigating factors** (which may have been relevant to sentencing decisions).

Some implications for research and policy in this area include developing a better understanding of:

- Why are SOC-related defendants **more likely to enter a guilty plea** at Crown Court?
- What are the drivers behind SOC-related appearances involving **female** defendants being **more likely to result in a discontinuation, dismissal or acquittal**?
- The **custodial journeys** of those imprisoned for SOC offences and their impact on prison safety and security.
- **‘What works’ in the effective management of SOC-related cases** (e.g., in terms of the most effective forms of post-sentence supervision, licence conditions and post-release requirements to reduce the risk of breach, recall and/or reconviction)?



ADR UK
Data-driven change

**Administrative data is an invaluable resource
for public good. *We're using it.***

Dr Tim McSweeney, Anglia Ruskin University
(email: timothy.mcsweeney@aru.ac.uk)



ADR UK 2024 Research Fellowship



ADR UK 2024 Research Fellowship

- ADR UK are inviting funding applications for research fellowships to conduct research and analysis of ADR England datasets including the three Data First datasets:
 - **Data First: Cross-Justice System – England and Wales**
 - **Data First: Family Court linked to Cafcass and Census 2021 – England and Wales**
 - **Ministry of Justice & Department for Education linked dataset – England**
- The Research Fellowships can be up to 18 months in duration, and up to a maximum of £200,000 at 100% full economic cost.
- MoJ are hosting drop-in sessions for researchers interested in using Data First datasets
- This funding opportunity has a **mandatory register of interest stage, which closes at 16:00 on 28 March 2024**
- If you have any questions please contact us at datafirst@justice.gov.uk or you can contact ADR UK at ADR Fellowships@esrc.ukri.org.

Q&A Part 2



Get in touch:

datafirst@justice.gov.uk

(for signing up to the Data First mailing list and any Data First related queries)

datalinkingteam@justice.gov.uk

(for queries regarding the external data linking shares e.g. MoJ-DfE share)



ADR Participant Survey

Find out more:

- Existing datasets, user guide & privacy statement: <https://www.gov.uk/guidance/ministry-of-justice-data-first>
- Application form and approved projects:
<https://www.gov.uk/government/publications/moj-data-first-application-form-for-secure-access-to-data>
- Splink data linking package: <https://github.com/moj-analytical-services/splink>