

Probation & criminal justice system linking dataset

Key messages from the Data First User Representation Panel

March 2022



Overview

This report records a meeting between [Data First](#) project partners and the [Data First User Representation Panel](#) to discuss the potential of a new 'probation & criminal justice system linking dataset'. It captures the main findings of the discussion with the intention to inform the future use of the linking dataset, including research priorities and scope for the growth of the dataset, with a recommendation to account for lived experience when analysing crime and justice data.

Particular areas of interest for stakeholders included a need for greater understanding of how different communities interact with the criminal justice system, rates of recall into the prison and the impact of pre-sentence reports.

This meeting was held on Monday 31 January 2022; a full list of attendees and areas for future research can be found in Appendices 1 and 2, respectively.

1. Introduction

Engagement with stakeholders, such as policymakers, public service providers and public advocacy groups, is essential to ensure that research using public sector administrative data is truly in the interests of those it hopes to benefit. Each of these stakeholders has an interest in knowing whether the focus and methods of research using public sector data are ethical, robust and useful, and that any potential negative consequences have been considered and mitigated.

On 31 January 2022, stakeholders came together to discuss the ‘probation & criminal justice system linking dataset’. This new de-identified linking dataset from the Data First programme (see Box 1) is now available for researchers to apply to use for research in the public interest. The event provided an opportunity for stakeholders to ask questions about and share their views on the linking dataset, and to suggest priority areas of research that would benefit the public. The event consisted of an introduction of the probation and criminal justice system linking dataset from a member of the Ministry of Justice (MoJ) and a facilitated discussion between the Data First project partners and the Data First User Representation Panel (hereafter referred to as the ‘User Representation Panel’). This panel consists of members who have the knowledge and expertise to represent the interests of justice system users in discussion around their data and its research uses. A full list of attendees is given in Appendix 1. This report covers the key messages arising from discussions at the event.

Box 1: Overview of the probation and criminal justice system linking dataset

This dataset enables a more comprehensive understanding of how people interact with the criminal justice system in England and Wales. The new linking dataset includes four datasets already made available via Data First:

- Probation dataset – which includes 1.9 million records from across England and Wales from January 2014 to December 2020
- Prisoner custodial journey level dataset – which includes 1.3 million records from across England and Wales from January 2011 to September 2021
- Crown Court defendant case level dataset – which includes 1 million records from across England and Wales from January 2013 to December 2020
- Magistrates’ court defendant case level dataset – which includes 13.1 million records from across England and Wales from January 2011 to December 2020

This de-identified linking dataset is available to external researchers via the Office for National Statistics (ONS) [Secure Research Service](#). This research will aim to improve our understanding of the experiences of the criminal justice system. Researchers will need to be [accredited](#) and submit a successful application to access the data.

2. The research value of this data

Commencing in 2019, Data First is an ambitious data linkage initiative that aims to harness the potential of the wealth of data already created by MoJ. The primary motivation for Data First is to link data from across the justice system and beyond to create a sustainable body of knowledge on justice system users and their interactions with government and across the family, civil and criminal courts. This will generate evidence to underpin the development of government policies and drive real progress in tackling social and justice problems.

Data First has been releasing de-identified, research-ready data into the Office for National Statistics (ONS) Secure Research Service for accredited researchers to apply to use since May 2020. Of this data, the following are included in the new probation and criminal justice system linking dataset:

- [Probation dataset](#)
- [Prisoner custodial journey level dataset](#)
- [Crown Court level defendant case level dataset](#)
- [Magistrate’s court defendant case level dataset](#)

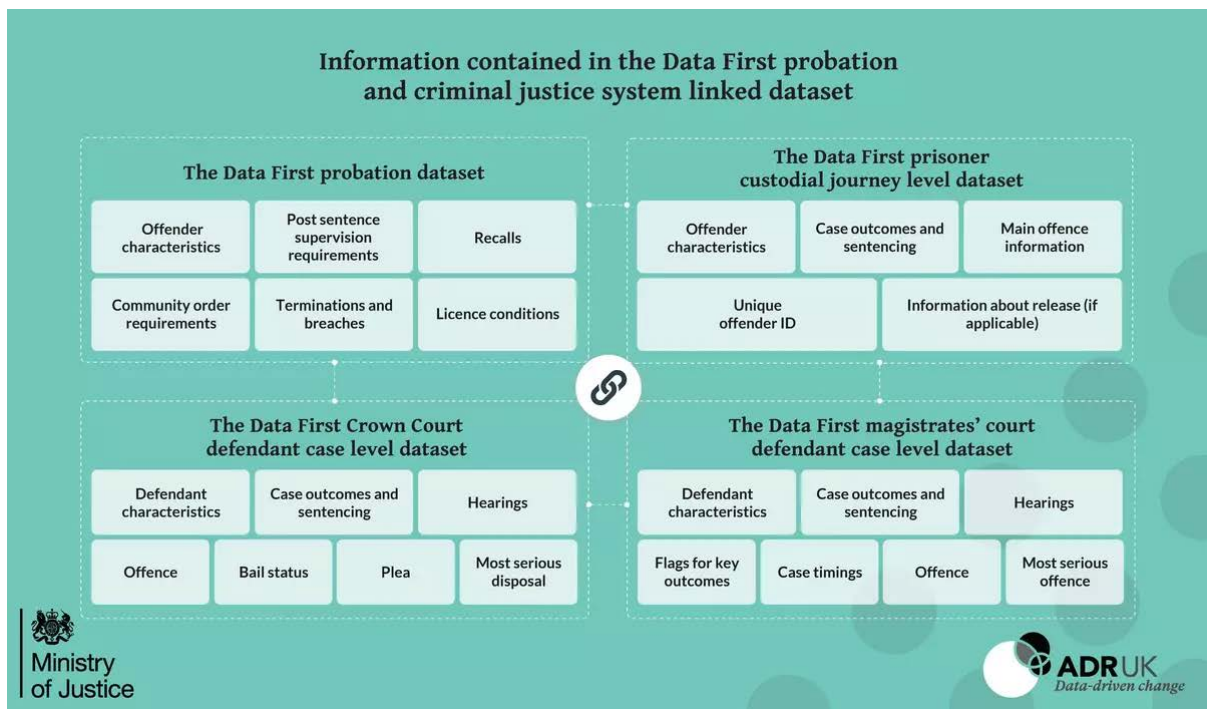


Figure 1: The probation and criminal justice system linking dataset

2.1. The probation dataset

In January 2022, probation data was released into the ONS Secure Research Service. The Probation dataset holds 1.9 million de-duplicated records across England and Wales from January 2014 to December 2020. It brings together data on offenders and events, sentence components and pre-sentence reports in a research-ready format.

The probation dataset is summarised below:

Data First probation dataset

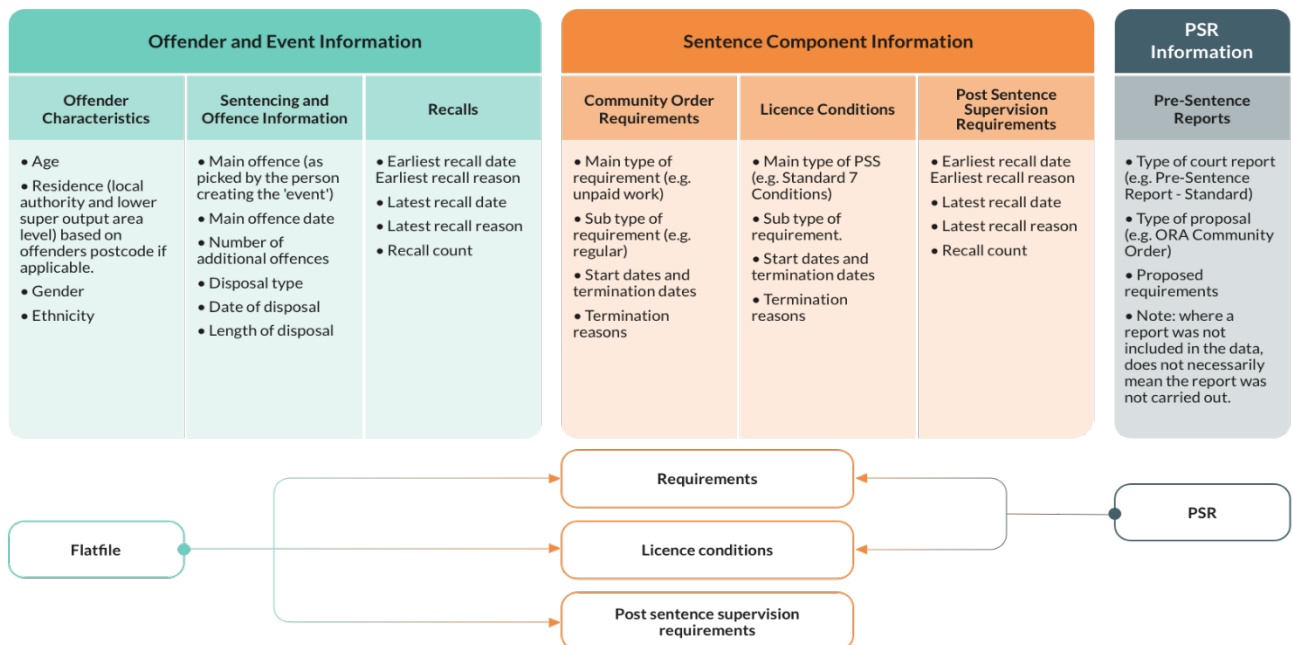


Figure 2: The probation dataset

As with all ADR UK datasets, the probation dataset has been de-identified and there are [rigorous safeguards](#) in place to protect it from re-identification. When accessible to researchers, data does not include any personal identifiers, with any elements that could be traced directly back to individuals – such as names, contact details or any identifying numbers – having been removed. What is left is a set of attributes of individuals along with recorded interactions with public services, allowing for relationships between these to be analysed.

2.2. The probation and criminal justice system linking dataset

The probation and criminal justice system linking dataset was transferred to the ONS Secure Research Service in February 2022, following a suite of datasets released from the Data First programme. The linking dataset acts as a lookup to identify where records in the various datasets refer to the same people. Each of the individual datasets can generate unique and granular insight on aspects of the criminal justice system, however, this dataset enables the 'linking' of this data at a person level, enabling research that aims to explore the end-to-end user journeys across the criminal justice system including how offenders interact between key justice services.

3. Research priorities for the linking dataset

Following an introduction to the dataset, the User Representation Panel discussed their research priorities for the linking dataset. Below are the main findings from the discussion.

3.1. How different communities interact with the criminal justice system

Key to the research potential of the probation and criminal justice system linking dataset is the ability to link between the four Data First criminal justice system datasets which are population-level data. This will enable public service providers the opportunity to cross-reference whether the demographics of those who access their services reflects the demographics of justice system users as recorded within the Data First criminal justice system datasets. This information can enable, for instance, a more accurate evaluation between those who access support services and those who interact with the justice system and if there are any underrepresented groups that aren't being reached. Demographic information concerning different actors in the criminal justice system – from victims and witnesses to justice system users - can also shed light on patterns of repeat court use and ways to direct support services that can mitigate these outcomes.

It is also possible to use demographic information to identify perceived and actual biases within the criminal justice system, as found in independent reviews on this topic. A representative from Revolving Doors drew attention to the Lammy review, which indicates that some individuals face disparities in the justice system as a result of their identities, including gender, race and neurodiversity. It was flagged that capturing the nuance and intersectionality between different characteristics can help us better understand these disparities and how services could be improved to mitigate and reverse these.

A representative from Cafcass highlighted how more robust demographic information on family background may also improve understanding of pathways into the court system. This was echoed by a representative from the Prison Advice and Care Trust, stating a greater understanding of the family background may also improve understanding of reoffending. While the linking dataset doesn't currently contain data on family background, it was suggested a linkage of this kind could yield greater understanding into the demographics of those who enter the court system and how services can be better targeted to support those most vulnerable to crime.

3.2. Rates of recall into custody

Central to understanding the relationship between prisons, probation and the criminal courts is more information on the rate of recall back into custody. There are many dimensions to understanding the impacts of rates of recall, including how the incident was documented, the length of recall (for instance, a fixed or undetermined length of time before returning into custody), and type of sentence an individual initially received. Greater knowledge of the wider circumstances concerning recall rates can yield insights into whether other factors, such as demographic information, location, support received during time in prison and on probation, for instance, have an impact on an individual reentering the criminal justice system.

A representative from Revolving Doors highlighted how positive, trusting and consistent relationships are widely recognised by both practitioner practitioners and people under probation supervision as core to good probation practice. Greater understanding of the quality of this communication (taking into account communication method, frequency of communication, and the number of practitioners the person has communicated with) could therefore support better understanding of the efficacy of Probation. Factors to consider include:

- communication (including when, frequency and type, i.e. written or verbal), and
- if an individual communicated with multiple or the same probation officer.

The granular information within the Data First criminal justice system datasets, particularly the probation dataset, will enable a greater understanding of the wider context of rates of recall, which in turn will enable a more efficient evaluation of successful rehabilitation.

The User Representation Panel members noted that in their experience those on shorter sentences tend to experience poorer communication with probation services whilst they are in prison custody than those given lengthier custodial sentences. This has been observed to negatively impact an individual's sense of preparedness and decision making as they re-enter the community. Pre-sentence reports, which provide information on the circumstances of an individual's case, can shed light on the root causes of crime and people's support needs. A large component of the probation dataset includes details on pre-sentencing reports, information which, when linked to the other criminal justice datasets, may therefore close some of these evidence gaps.

In their experience, members of the User Representational Panel found the type of support, such as emotional or educational, received by an individual on probation can also affect the likelihood of recall. A representative from Khulisa shared the need to better understand what works in 'Through the Gate' services aimed at mitigating offending, and whether prioritising social and emotional skills provision can help individuals on probation better engage with society. In addition different types of sentences may also shed light onto recall rates, for instance the linked data could be used to explore:

- whether those given community rehabilitation services fare better than those sentenced to prison for similar offences, and
- if those with community rehabilitation services receive timely access to support.

This information can help inform whether there is a better alternative to short prison sentences for low-level offences, including if alternative sentences have an impact on reoffending rates. A better understanding of efficacy of short sentences and alternative forms of sentences can inform approaches to help reduce crime and support rehabilitation.

It is evident that to evaluate rates of recall, researchers need a holistic understanding of the journey of justice system users. The probation and criminal justice system linking dataset creates the opportunity to enable greater understanding of what works and for whom.

3.3. The impact of pre-sentence reports

Crucial to filling evidence gaps is a greater understanding of the impact of pre-sentence reports. Pre-sentence reports can be drafted at any time before a sentence and the level of detail is dependent on a variety of factors (i.e. if they are drafted on the day of sentence, they can be less detailed). Experiential evidence from the User Representation Panel suggests the quality, number and timing of pre-sentence reports may have an impact on the outcomes of justice system users and their sense of judicial trust. A representative from Revolving Doors highlighted an existing knowledge gap on whether there is more judicial trust in certain areas, regions or within certain courts around both pre-sentence reports and the recommendations around sentencing they include.

When creating the probation dataset, it became clear from the data that on occasion, requests from the court do not match what is delivered in the sentence. For instance, courts may request a full pre-sentence report but may receive a fast pre-sentence report, due to various reasons. Understanding what, if any impact this mismatch of requests has, can help inform evidence gaps pertaining to levels of trust in the quality of the pre-sentence reports and whether the content of the reports address what is needed for someone on probation.

A representative from Revolving Doors expanded on this by stating that pre-sentence are reports in critical in helping sentencers to build a fuller picture of defendants' circumstances, including mitigating factors such as being the victim of domestic violence or coercion. A fast pre-sentence report, with limited time to build the trust required to support defendants to talk openly about these mitigating circumstances, may therefore fail to communicate factors that are key to sentencers in determining appropriate sentencing.

3.4. Accounting for lived experience

The User Representation Panel raised the importance of lived experience across many actors within the criminal justice space. It was noted that understanding the perspectives of those on probation, in addition to that of probation practitioners, is crucial to evaluating and improving services and gaining insight into why some groups have better outcomes than others. Meanwhile, understanding the role of family and personal and social connections can provide insight into reasons for offending and reoffending, and how to build connections to mitigate experiences with the criminal justice system.

While further exploration into nuances that arise from lived experience is more suitable for qualitative research, the linkage of probation practitioner data with data on individuals on probation can provide insights on how, for instance, caseload and rates of recall can affect outcomes of individuals on probation. Likewise, a better understanding of referrals made by probation practitioners, including the types of referrals and whether they are taken up by the courts, can help better understand if input from probation practitioners impacts outcomes for individuals on probation seeking support services, such as housing. This information can inform rates of reoffending as well as whether there are barriers to service access and whether that impacts outcomes.

Conclusion

This meeting with the User Representation Panel highlighted the multifaceted dimensions of the criminal justice system, and the need for greater understanding of the end-to-end user journey to shed light on what works and areas which need improving. It is evident that a multitude of factors, from timing to types of interventions, sentences, pre-sentence reports, support, and so on need to be considered when evaluating the efficacy of rehabilitation methods and mitigations of crime, reoffending and recall.

The User Representation Panel expressed interest in the findings of all MoJ data to be shared in the form of short reports, highlighting areas in which there are gaps in service provision that could be filled by third sector organisations. They also encouraged researchers working on this theme to propose opportunities for collaboration and for futures of the data to mitigate ethical considerations concerning the use of crime and justice data by engaging with third sector groups working on or on the behalf of justice system users.

Acknowledgements

The Data First project is led by the [Ministry of Justice \(MoJ\)](#). It is funded by [ADR UK \(Administrative Data Research UK\)](#). ADR UK is funded by the [Economic and Social Research Council](#) (part of [UK Research and Innovation](#)).

Author

Shayda Kashef, ADR UK Strategic Hub, shayda.kashef@ukri.org

Visit the [ADR UK website](#)



[@ADR_UK](#)



Appendix 1: List of attendees, ordered alphabetically by organisation¹

Chair: [Christine Boase](#), Administrative Data Research UK

| | |
|--|--|
| <i>Data First project partners</i> | Representatives |
| Administrative Data Research UK | Christine Boase, Strategic Lead for Communications & Campaigns |
| Administrative Data Research UK | Gregory Meredith, Senior Research & Impact Manager |
| Administrative Data Research UK | Bogusia Wojciechowska, Research & Impact Manager |
| Administrative Data Research UK | Shayda Kashef, Public Engagement Manager |
| Administrative Data Research UK | Grainne Wrigley, Communications & Engagement Officer |
| Administrative Data Research UK | Hannah Ratcliffe, Grants Delivery Manager |
| Ministry of Justice | Thomas Jackson, Senior Social Researcher, Evidence & Partnerships Hub |
| Ministry of Justice | Georgina Eaton, Statistician, Evidence & Partnerships Hub |
| Ministry of Justice | Kylie Hill, Statistician, Evidence & Partnerships Hub |
| Nottingham Trent University | Andromachi Tseloni, Professor of Quantitative Criminology and Academic Lead for Data First |
| Office for National Statistics | Alex Lock, Strategic Data Curation Officer |
| <i>Data First User Representation Panel</i> | Representatives |
| Cafcass | Saif Ullah, Senior Research & Evaluation Manager |
| Khulisa | Iman Haji, Senior Evidence & Impact Officer |
| The Prison Advice & Care Trust | Max Banfield, Data & Performance Information Manager |
| Prison Reform Trust | David Maguire, Project Director |
| Revolving Doors | Phillip Mullen, Research Manager |

¹ Invitations were also extended to other organisations within the Data First User Representation Panel who were unable to attend.

Appendix 2: Areas of interest for future

Below are suggest areas of research interest using linked criminal justice data as identified by the Data First User Representation Panel.

How different communities interact with the criminal justice system

- Is there a discrepancy between the demographics of those who access support services and those who interact with the criminal justice system?
- Are particular communities more at risk of entering the court system?
- Are there factors that contribute to different groups receiving different sentences?

Rates of recall into custody

- Do those on shorter sentences have a better or worse chance of rehabilitation?
- What impact does support received in prison have on the likelihood of an individual being recalled back into custody?
- What are the trajectories of people who received different types of support via 'Through the Gate' services?
- How do the outcomes of those given community rehabilitation services compare to those sentenced to prison?
- How do license period, conditions, and durations affect the potential for recalls?
- What are the enablers and barriers to effective sentences, including community-based, alternative or short custodial sentences?

The impact of pre-sentence reports

- What makes an effective pre-sentence report?
- Can the quality of a pre-sentence report affect the outcomes of an individual on probation?