Ministry of Justice

Data First:

Privacy and data protection

Harnessing the potential of linked administrative data for the justice system.

Version 2.0 April 2021

ADR UK

Contents

About Data First

Data First is an ambitious data-linking programme led by the Ministry of Justice (MoJ) and funded by ADR UK (Administrative Data Research UK).

Data First aims to unlock the potential of the wealth of data already created by MoJ by linking administrative datasets from across the justice system and beyond. Internal data linking work will match civil, family and criminal justice administrative records held by MoJ. External data linking will seek to bring justice data together with data held by other government departments (OGDs).

As part of Data First, MoJ are committed to sharing deidentified and research-ready datasets with accredited researchers, across government and academia, in a secure, ethical and responsible way. Working in partnership with academics to facilitate and promote research in the justice space, will create a sustainable body of knowledge on justice system users and their interactions across the family, civil and criminal courts, and their needs, pathways and outcomes across a range of public services. This will provide robust evidence to inform the development of government policy and drive progress in tackling social and justice problems.

Data First forms a crucial part of MoJ's wider ambitions to maximise the use of evidence and external expertise to shape policymaking and improve justice outcomes. The linked administrative data made accessible via the project will enable some of the critical evidence gaps outlined in the departments Areas of Research Interest to be explored in collaboration with our partners. In doing so, the aim is to strengthen the strategic research capabilities across government and academia and reinforce the impact of evidence at all stages of policy development and evaluation.

How we use data

This document outlines how Data First uses and shares data for research purposes and how this data is kept safe and secure.

Adding value by linking and sharing

Data First harnesses the potential of existing administrative data by linking internal justice datasets (and administrative data from OGDs), for accredited researcher access via secure platforms. It does not involve the collection of any new data. The project reflects a drive across government to make better use of the information it routinely collects when users access public services such as the courts.

In line with legislation

MoJ is permitted to process data supplied by the police, the Crown Prosecution Service (CPS), courts and prisons by virtue of its common law powers for the administration of justice.

Data is only shared with OGDs where a suitable legal gateway exists, and is processed in compliance with all applicable data protection legislation including the General Data Protection Regulation (GDPR) and Data Protection Act 2018.

Data First has been developed within the framework established under the Digital Economy Act (DEA) (2017) which enables government to prepare administrative data for the purposes of research, and to provide de-identified versions of those data to researchers and projects accredited by the UK Statistics Authority (UKSA). This is the legal basis for the project and is the framework for ensuring that data is used responsibly, ethically and will not identify individuals.

In order for research projects to be approved they must comply with the Research Code of Practice and Accreditation Criteria which was approved by the UK Parliament in July 2018.

For permitted research purposes

Provisions for using data collected in the course of the department's operations for research and statistical purposes are set out in the department's Personal Information Charter and that of Her Majesty's Courts and Tribunals Service (HMCTS).

Data Sharing Agreements (DSAs) between government departments specify the permitted uses of the data as well as the safeguards in place. These stipulate that datasets shared under Data First are accessible for approved research projects only. They do not permit datasets to be re-used for operational purposes, nor to identify individual justice system users or their outcomes. No analysis of the performance of the courts, or of individual judicial decisions will be permitted.

Taking wider views into account

Data First is supported by an external Academic Advisory Group comprising experts with wide-ranging technical expertise and justice knowledge, including on methodological data linkage, justice research, and the ethical use of administrative data for public policy. They provide advice and constructive challenge to the Data First team and help to ensure that the project follows best practice.

The MoJ ethics advisory group and external ethicists will be consulted as the programme develops to ensure that it, and the research projects that it enables, are ethically sound. This will include the consideration of the impact on data subjects and their data protection.

To reflect the commitment of Data First to represent and protect the needs and interests of justice system users, a User Representation Panel has also been established. The panel is made up of representatives from across the third sector, who work closely with justice system users across the criminal, civil and family justice jurisdictions, to represent their views and needs throughout the project.

How we protect personal information

Linking data

Data relating to the same person is linked by using identifying information, such as names, date of birth and addresses, that are held in the source data. This information will not be shared with researchers and will be replaced in the data linking process by a meaningless identifier (one that has been generated for these datasets and is not used in any existing operational systems). Other identifiers used within the justice system, such as case IDs, will also be replaced.

See the Data First user guide for more information on data linking and identifiers.

Considering risks and sensitivities

MoJ recognises the risks associated with sharing data. As well as only allowing researchers to access data after identifiers have been removed, there are rigorous safeguards in place to ensure data cannot be accessed by any unauthorised persons, or for any reason other than approved research projects.

Information classed as ‘special category’ data under GDPR are sensitive, as is information about criminal histories. MoJ recognises the risk of this information being disclosed and requires additional justifications for processing and sharing this data, in line with GDPR.

Personal descriptors such as sex or gender, ethnicity or age, might be shared if it is insufficient to identify someone. Some fields or combinations of data do not identify a person directly but increase the risk of identification. This could occur, for example, if very few people match a set of characteristics or activities recorded in the data. To protect individuals, the application process will consider the sensitivity of both the individual data items and the full combination of data being requested. The ethical issues around the use of data are factored into the decision-making process. All research undertaken under Data First must be approved on the basis that it is in the public interest.

Limiting access to what is needed

MoJ will take steps to minimise the amount of data accessed for research. Researchers will need to justify why they require the use of each specified field and the coverage of the dataset within their data access application.

For more sensitive data, greater justification is needed. For example, where two fields give different levels of details (such as offence categorisations) researchers would need to justify that their research question requires use of the more extensive version.

Restricting and monitoring use and outputs

Users will access data in a secure setting without internet access and with limited external resources. This limits the risk of identification as researchers will not be able to combine the data with information outside of the dataset and use this to infer information about individuals.

The current primary processor for accessing the data is the Office for National Statistics (ONS) Secure Research Service (SRS) (see below: How we share data). Users' activity is closely monitored in the ONS SRS and access will be revoked if they go beyond the permitted uses of the data. Attempting to identify individuals is strictly against the terms and conditions which researchers must agree.

All outputs are checked to ensure they remain within the scope of the project that has been approved and that personal information is protected as far as is reasonable, for example, through minimum cell counts and statistical disclosure control.

Data is only kept as long as necessary and is then securely deleted.

How we share data

Sharing with other government departments and agencies

Data shared between the MoJ and OGDs is governed by a Data Sharing Agreement (DSA). This establishes a framework for appropriate processing, including the contents and duration of the share, permitted uses of the data, the legal basis and justification for processing, and the protections in place throughout its lifecycle.

A Data Protection Impact Assessment (DPIA) will be completed prior to data sharing, to identify and assess risks to individuals' privacy and ensure appropriate protections are in place to minimise them. As the share evolves, so too will the details in the DPIA and DSA report.

For data linking to take place, personal identifiers (such as name, address and date of birth) are transferred between departments. Access to this information is strictly limited to those who need it in order to carry out the linking process. These identifiers are transferred separately from all other information about individuals and deleted immediately following linking, to minimise the risk of identification. New, meaningless IDs assigned to individuals throughout linkage are the only IDs provided for use in analysis. Only the minimum of information (both fields and coverage) justified for sharing is supplied.

All data are transferred using an agreed mechanism and stored in a setting that is compliant with the HMG Security Policy Framework.

Sharing with accredited processors and trusted third parties

To facilitate sharing with researchers outside of government, MoJ will share deidentified, research-ready data with accredited processors under the Digital Economy Act (2017) to make access securely available for research purposes.

ONS Secure Research Service

Currently, data is shared with one accredited processor: the ONS SRS. The SRS is designed to deliver full compliance with the statutory Conditions and Code of Practice of the DEA (2017) and has substantial data expertise, especially in data management, metadata, and the checking of outputs before they leave the centre.

Each dataset shared between the MoJ and ONS will be subject to a DSA and DPIA.

The de-identified extracts of the data will be transferred using an agreed mechanism and stored in a setting that is compliant with HMG Security Policy Framework.

Sharing with external researchers

Researchers will be given access to the deidentified data according to the needs of their accredited research project only. Access will be governed by an application process whereby users specify what data they require for analysis.

Access to the data is limited to a fixed time period agreed by MoJ. Research outputs will consist of aggregate data only. Users are monitored and outputs are checked to ensure analyses remain within scope of the approved project.

Data First uses the ONS 'Five Safes' framework to ensure information is kept safe and secure:

Safe People – Trained and accredited researchers are trusted to use data appropriately.

Data First enables access to data extracts to accredited researchers, who have met the conditions set out in the UKSA Research Code of Practice and Accreditation criteria.

Assurances will be sought in the application process that researchers have appropriate skills in working with data and have undertaken relevant data protection training.

Safe Projects – Data are only used for valuable, ethical research that delivers clear public benefits.

All research applications to access Data First datasets must be approved by data owners via data access panels at MoJ, HMCTS and/or partner departments (where applicable). They consider whether the research proposal is ethical; whether access is necessary to address the research questions; whether there are any data protection concerns; the research proposal methodology; and benefit to the public good.

Applications that have secured data owner approval will then need to secure project accreditation by the UKSA Research Accreditation Panel (RAP). Projects must meet the accreditation criteria published by the UKSA in the Research Code of Practice and Accreditation criteria.

Safe Settings – Access to data is only possible using secure technology systems.

Access to the SRS is provided in safe settings on ONS sites or other certified sites, or through an organisation which has a current and certified Assured Organisational Connectivity Agreement with ONS.

Safe Outputs – All research outputs are checked to ensure they cannot identify data subjects.

Research outputs will consist of aggregate data only and will be checked thoroughly to ensure that they are in line with the stated purposes of the project and that personal information is protected.

In line with the Digital Economy Act (2017; Bar on further disclosure of personal information), all research outputs from the SRS will undergo thorough Statistical Disclosure Control (SDC) prior to publication to protect the identity of individual persons and/or businesses in those releases.

Safe Data – Researchers can only use data that have been de-identified.

All data provided through Data First will be de-identified and risks and sensitivities around sharing the combination of data requested will be considered before access is granted (see how we protect personal information).

Approving access to data

Applications by external researchers will be subject to internal governance processes based on their individual merit.

The application form and guidance can be found on gov.uk.

- MoJ governance processes require consideration by the Data Access Group (DAG) and approval from the Data Access Governance Board (DAGB) chaired by MoJ's Chief Statistician.
- Data First datasets containing prisons or probation data collected by HMPPS also follow MoJ processes. This has been agreed with the National Research Committee (NRC) who handle requests to access other HMPPS data.
- If access to data on courts and tribunals is requested HMCTS processes apply. These require consideration by their Data Access Panel (DAP). If deemed necessary, applications may be escalated to the Data Governance Authority (DGA) or Shadow Data Governance Panel (SDGP).
- If access to data owned by OGDs is requested, they will require consideration by the separate relevant processes of that department.

A brief summary of successful applications to access the data will be published on gov.uk. There may be exceptions where we do not publish information, for example, to protect the privacy of researchers working on particularly sensitive topics.

Accessing your data

Information on how to request access to copies of your personal information can be found in the MoJ's personal information charter or HMCTS personal information charter.

Reviewing the information

The data will be reviewed and refreshed approximately annually to include new data, provided the administrative system has continued being used to record equivalent information. This privacy statement and our data protection processes will be updated to reflect any relevant changes.

Contact us

If you would like more information or have any questions, please contact us at: datafirst@justice.gov.uk

Any enquiries regarding this publication should be sent to us at datafirst@justice.gov.uk