

# INSIGHT Data Use Application: Guidance and Application Form

## Guidance to applicants:

1. The information that you provide in this Data Use Application form will be used to inform our application review process, including due diligence and ultimately any future Data Licensing. We therefore strongly recommend that you contact INSIGHT to discuss your potential data requirements beforehand.

We can be contacted by email at

[enquiries@insight.hdrhub.org](mailto:enquiries@insight.hdrhub.org)

Not all enquiries progress to the application stage, since it may be that your application can be met through existing published data, or you may need to discuss your proposal with the INSIGHT team.

2. If you are ready to make a formal application, we require you to complete all sections of this form and provide all relevant supporting information and enclosures where applicable.

### Some important elements of an application that often require further discussion include:

- technical details
- whether INSIGHT holds the data requested
- the purpose for accessing data
- the benefits accessing the data could yield for health and social care, patients and the public, in the UK and globally

The application will be referred back to you if more information is required in any of these areas.

## Information Governance

We require assurance of appropriate Information Governance awareness and training in advance of considering a Data Use Application. Currently we require you to confirm you have read:

**The National Data Guardian (NDG) Standards for Healthcare Data**

**The Data Security Standard Overall Guide – (DSP) Toolkit**

**National Data Guardian for Healthcare (CDG) Review: Review of Data Security, Consent and Opt-Out assets**

And undertaken appropriate training, such as:

- **Data Security and Awareness Level 1 Training** (organised by elearning for healthcare)
- **MRC Regulatory Support Centre: Research, GDPR & Confidentiality** (we will accept a certificate of completion)
- **Safe Researcher Training** (organised by the UK Data Service)

Information Governance awareness and training must be current. Training must be updated every three years, and where your training certificate expires within the time period of your study, you will be required to renew your training.

**Please note**, this is not an exhaustive list, so if you have undertaken other relevant Information Governance training that covers comparable areas, please provide details in your application.

**INSIGHT's datasets can be reviewed via our [website](#) and the [HDR UK Innovation Gateway](#)**

## Access to data:

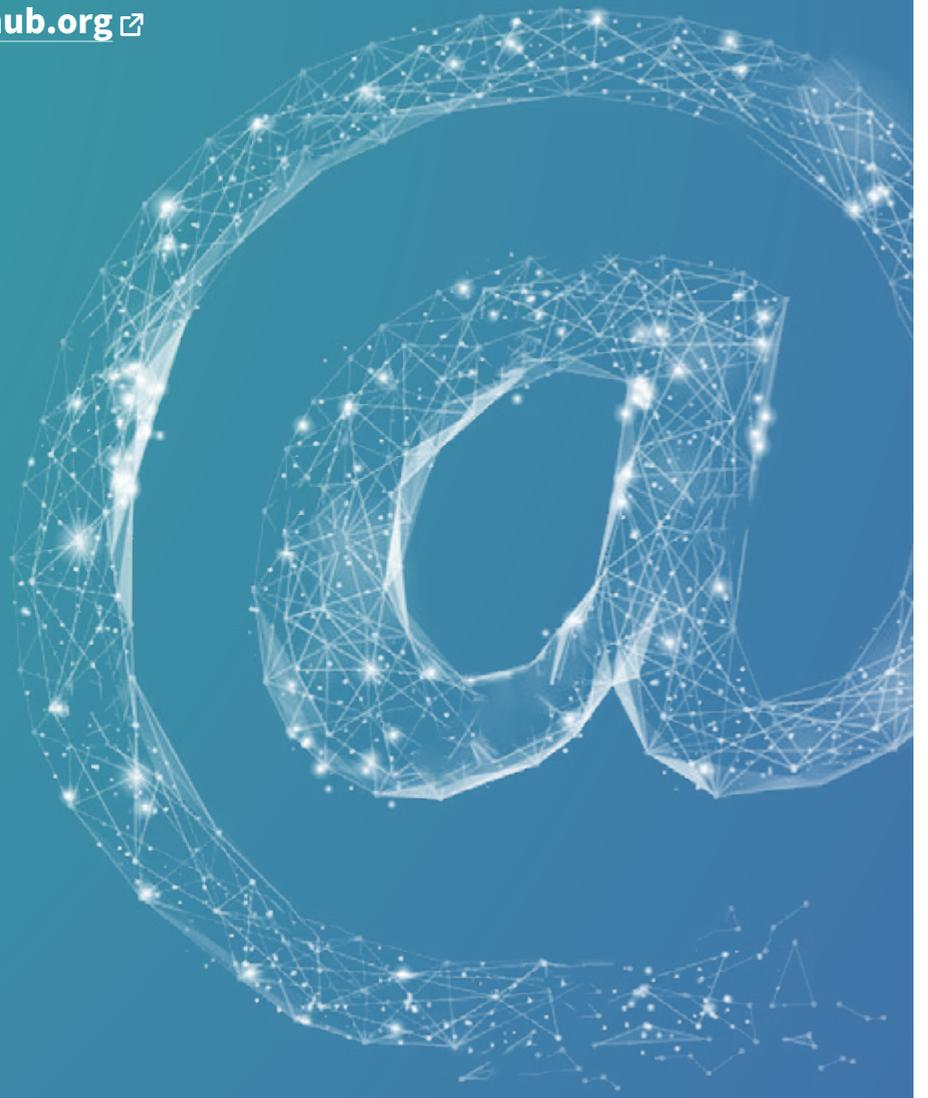
Subject to a successful Data Use Application, the representative Data Controller under INSIGHT would engage with you and your organisation to establish a Data Licensing Agreement.

## Costs of data access:

INSIGHT operates a Full Economic Cost recovery model and, subject to the specific data requirements, we may need to request payment from you to cover the cost of administering and processing your Data Use Application.

The costs associated with a Data License Agreement will be discussed, subject to a successful Data Use Application, during contractual negotiation with the representative Data Controller.

**If you have any queries, please contact**  
**[enquiries@insight.hdrhub.org](mailto:enquiries@insight.hdrhub.org)** 



# Data Use Application Form

## Dataset(s):

Please provide name of dataset(s) as listed on the [HDR Innovation Gateway](#), or state 'bespoke'

Title of your research project for this Data Use Application

This application is part of a [National Core Studies project](#)

---

## Safe people

### Who is going to be accessing the data?

Safe people should have the right motivations for accessing research data and understand the legal and ethical considerations when using data that may be sensitive or confidential. Safe people should also have sufficient skills, knowledge and experience to work with the data effectively. Researchers may need to undergo specific training or accreditation before accessing certain data or research environments, and demonstrate that they are part of a bona fide research organisation.

### The purpose of this section is to ensure that:

- details of people who will be accessing the data and the people who are responsible for completing the application are identified
- any individual or organisation that intends to access the data requested is identified
- all identified individuals have the necessary accreditation and/or expertise to work with the data effectively

Please identify the Lead Applicant, Co-Applicants and other significant project team members as a minimum, including those who will access the data.

This includes those from the applicant organisation and any third party organisation(s) who will access the data for the purposes of this request.

### Applicant 1

**Full name**

**Job title**

**Organisation** (with which employment contract is held)

**Email address**

**ORCID iD**

**Role** (on project)

- Principal Investigator
- Collaborator
- Team member

Other (please state):

**Will this person access the data requested?**

Yes      No

**Has this person undertaken professional training or education on the topic of Information Governance?**

Yes      No

(see [Information Governance](#) ↑)

Please provide details of most recent relevant training attended or planned.

### Other individual

**Full name**

**Job title**

**Organisation** (with which employment contract is held)

**Email address**

**ORCID iD**

**Role** (on project)

- Principal Investigator
- Collaborator
- Team member

Other (please state):

**Will this person access the data requested?**

Yes      No

**Has this person undertaken professional training or education on the topic of Information Governance?**

Yes      No

(see [Information Governance](#) ↑)

Please provide details of most recent relevant training attended or planned.

Please provide evidence of team’s expertise and experience relevant to delivering the project

Please attach CVs for the applicants identified

CVs should include as a minimum:

- details of most relevant publications
- relevant outputs
- relevant training (to include Information Governance training)

**If you wish to add further individuals, please use [Annex 1: Additional individuals](#) ↴**

# Safe project

## What is the purpose of accessing the data?

Safe projects are those that have a valid research purpose with a defined public benefit. For access to data to be granted, researchers need to demonstrate that their proposal is an appropriate and ethical use of the data and that it is intended to deliver clear public benefits. The purpose of this section is to ensure that:

- the project rationale is explained in lay terms
- the research purpose has a defined public benefit; this can be new knowledge, new treatments, improved pathways of care, new techniques of training staff, etc.
- how the data requested will be used to achieve the project objectives is articulated

## About this application

### This application is...

A new application

A renewal of an existing approval

An amendment to an existing application

Related to a previous application (approved or not)

An extension of an existing approval

---

## Project details

### Title of project (300-character limit)

The title should identify the main area of your research so that another researcher could understand if it might be relevant to their area of study

**Provide a lay summary of the project (300-word limit)**

The lay summary will be published on the INSIGHT website and should be in plain English, stating the:

- project aims
- scientific rationale
- project duration
- public health impact

---

**What is the anticipated start date of the project? (DD/MM/YYYY)**

---

**What is the anticipated end date of the project? (DD/MM/YYYY)**

**What are the project aims, objectives and rationale? (300-word limit)**

Please describe the specific aims of the project including:

- any hypotheses that you hope to test
- how the data requested are required to help address these aims
- details of any peer review
- referencing of any supporting evidence

**How will the data requested be used to achieve the project objectives?**

Please provide full details of your research methodology. This must include justification of sample size, analyses proposed, statistical methods, additional data sources such as linked data, and any plans for collaborative work. This information will be key to assessing whether your proposal will be feasible, deliver clear public good, and be an appropriate use of data.

Ensure you:

- specify the method(s) of analysis you plan to use (such as regression)
- as far as possible, try to articulate the outcome or dependent variable(s)
- indicate the starting-point for modelling process – acknowledging that the model may evolve
- explain (where relevant) how any potential selection/casual bias will be addressed (e.g. by including a control group with information on how this control group will be created)
- provide methodology references, if a non-standard methodology is proposed
- include information about any contribution to the field of research methodology that you believe may result from your research
- include an explanation of how your methodological approach will answer the research question(s) set out in the project when employing methods not covered by any of the above (e.g. correlation or basic descriptive analysis only will be used, noting that such analysis might be more applicable for exploratory research)

---

**How will the project benefit the public and what is the anticipated impact?**

Please describe the specific aims of the project including:

- any hypotheses that you hope to test
- how the data requested are required to help address these aims
- details of any peer review
- referencing of any supporting evidence

---

**Can you provide an outline of the patient and public involvement and engagement (PPIE) strategies of the study or a brief explanation of why they are not planned? (300-word limit)**

Alongside established practices of public engagement, including sharing the results and disseminating the findings in an accessible way to the general public, please also describe how you have included public and patients' views when designing your research questions and this project.

**For clarity, we expect:**

- inclusion of public and patients' views in the design of the proposal (this could be direct interaction or based on published surveys and research work, e.g. by Understanding Patient Data and UseMyData).
- accessible ways to communicate the objectives and outputs of your project to the general public.

---

**Do you anticipate any risk to individual privacy, and if so, what steps have you made in your proposal to mitigate these?**

Please describe any steps taken to mitigate potential risks to individual privacy.

---

**In considering your proposal, do you believe it could disadvantage any group or individual? Explain what steps you have taken to avoid this.**

Please explain if you believe you could disadvantage any group or individual in considering your proposal, and explain what steps you have taken to avoid this.

---

**Please provide up to 6 key words which best summarise your proposed research project**

### Will the research project enhance future benefits from our research database?

Examples might include:

- derived analyses of existing data, e.g. calculated or corrected values from raw measurements
- classification or collapsing of terms
- new labels to accompany image sets, e.g. assigning features and/or diagnoses to new images
- new analyses of image sets, e.g. providing segmentation or other manipulations or analyses of images that may be of value to other researchers

---

### Funder information

A funder is the organisation or body providing the financial resource to make the project possible, and may be different to the organisation detailed in the Safe people section.

**Does your project have a funder?**                      Yes                      No

**If Yes, please provide name:**

**If No,** please provide details of how you intend to fund the study

**If No,** please attach evidence of independent peer review.

## Sponsor information

The sponsor is usually, but does not have to be, the main funder of the research. The sponsor takes primary responsibility for ensuring that the design of the project meets appropriate standards and that arrangements are in place to ensure appropriate conduct and reporting.

**Associated applicant** Please confirm the name of the applicant linked to this sponsor organisation. This information will be used to form the legal documents.

**Organisation name** Please provide legal name to appear on legal documents

**Registered address** (line 1)

**Registered address** (line 2)

**City**

**Postcode**

**Country**

**Sector**

**Size** (small, medium, large)

**Additional details** Please provide any additional information

### Contracts person

Please provide a contact email address for the sponsor organisation or the Lead person to liaise with regarding agreements.

**Full name**

**Contact email address**

**Contact telephone**

If you wish to add a further sponsor, please use [Annex 2: Additional sponsor](#) ↴

## Safe data

Safe data ensure that researchers have a clear legal basis for accessing the data and do not inadvertently learn something about the data subjects during the course of their analysis, minimising the risks of re-identification. The minimisation of this risk could be achieved by removing direct identifiers, aggregating values, banding variables, or other statistical techniques that may make re-identification more difficult. Sensitive or confidential data could not be considered to be completely safe because of the residual risk to a data subject's confidentiality, hence other limitations on access will need to be applied.

### The purpose of this section is to ensure that:

- there is a clear legal basis for accessing the requested data
- the data requested is proportionate to the requirement of the project
- all data requested is necessary in order to achieve the public benefit declared
- data subjects cannot be identified by your team by cross-referencing data sets from anywhere else

These are the information assets which your proposal seeks to access and use. You should consider this definition to be wide in scope and include any source of information which you propose to access and use.

### Data fields

**Please indicate the data necessary to conduct the study, the data fields required and the justifications for each field.**

Please list and make clear the details on the information from each dataset that you would like to have access to. Data requested can only be provided for processing data for a specific purpose. Please justify why this dataset is required to answer the research question within this application.

### Inclusion and exclusion criteria (including data parameters)

**A description of the precise criteria which define the patients to be included and to be excluded from the data extract you are requesting should be provided.**

This should include precise date parameters for the start and end of the range requested (dd/mm/yy) and explain which dated project field will be used to define the requested cohort (e.g. date of admission or date of operation).

---

### Analysis

**Do you wish to commission the Data Controller to conduct the analysis for you, minimising your exposure to the data?** Yes      No

---

**Do you intend for the datasets requested to be linked with any additional datasets, other than the datasets listed in this application?** Yes      No

INSIGHT provides anonymised data, therefore for data to be linked, INSIGHT would need to undertake this linkage.

---

### **Please summarise the risks / mitigations considered.**

Linking datasets can increase the potential risk of identification. Please provide a summary of the risk(s) you have considered and how you intend to mitigate these.

## Statistical analysis

**What forms of statistical analysis are planned?**

---

**How is it intended that this will be presented in the final output?**

---

**What is the smallest cell value that is likely to be generated by this analysis and how will this be managed to avoid disclosure? (300-word limit)**

For example, an age range of 90-100 in a rare disease may result in a cell value of 1.

These questions are to help us understand the different uses to which INSIGHT data is being put to.

**Will the INSIGHT data be used for algorithm generation and training?**      Yes      No

---

**Will the INSIGHT data be used for internal validation?**      Yes      No

---

**Will the INSIGHT data be used for external validation?**      Yes      No      Other (please state):

---

### Ethics approval

**Do you seek for your project to be approved under the generic favourable ethical opinion of the INSIGHT Research Database (Ref: 20/WS/0087)?**      Yes      No

If No, please provide details of REC approval

**REC Committee name**

**REC Reference**

**Additional information**

If No, please attach evidence of REC Approval.

# Safe settings

Safe settings are analytics environments where researchers can access and analyse the requested datasets in a safe and ethical way. Safe settings encompass the physical environment and procedural arrangements such as the supervision and auditing regimes. For safe settings, the likelihood of both deliberate and accidental disclosure needs to be explicitly considered.

## The purpose of this section is to ensure that:

- researchers access requested data in a secure and controlled setting such as a Trusted Research Environment (TRE) that limits the unauthorised use of the data
- practical controls and appropriate restrictions are in place if researchers access data through non-TRE environment. There may be requirements that data is held on restricted access servers, encrypted and only decrypted at the point of use.

This section details in what way the proposal aims to store and use data, and controls in place to minimise risks associated with this storage and use. If you have indicated that your proposal seeks to store and use data exclusively through a recognised trusted research environment, then you do not need to complete this section.

In relation to personal data, means any operation or set of operations which is performed on personal data or on sets of personal data (whether or not by automated means, such as collection, recording, organisation, structuring, storage, alteration, retrieval, consultation, use, disclosure, dissemination, restriction, erasure or destruction).

All Locations where processing will be undertaken, for the avoidance of doubt storage is considered processing. For each separate organisation processing data which is not fully anonymous a separate partner organisation form must also be completed.

Processing, in relation to information or data means obtaining, recording or holding the information or data or carrying out any operation or set of operations on the information or data, including:

- a) organisation, adaptation or alteration of the information or data,
- b) retrieval, consultation or use of the information or data,
- c) disclosure of the information or data by transmission, dissemination or otherwise making available, or
- d) alignment, combination, blocking, erasure or destruction of the information or data

## Storage and processing

<b>Will the data be accessed within a trusted research environment?</b>	Yes	No
---	-----	----

### If Yes, please identify which one

Secure e-Research Platform (SeRP)

NI Honest Broker Service (NI HBS)

Scottish National Safe Haven (SNSH)

NHS Digital

SAIL Databank

ONS Secure Research Service

Other (please state):

**If No, please complete the following section and provide a data flow diagram:**

As a minimum this should define the flow of data from INSIGHT to the recipients. This should clearly detail which data or subsets of the data are shared with which recipient organisations and will be used as the basis for contracting.

Please provide details of the organisation doing the processing and/or storage:

**ICO registered name of organisation**

**ICO registered number**

**What type of security does this organisation have in place?**

Data security and Protection Toolkit (DSP Toolkit)

SLSP

ISO 27001

Other (please state):

If you wish to add any further organisations, please use [Annex 3: Additional organisations](#) ↴

## Dataflow

### Please provide detailed information on data flows.

As a minimum, this should define the flow of data from INSIGHT to the recipients. This should clearly detail which data or subsets of the data are shared with which recipient organisations and will be used as the basis for contracting.

**Please include a data flow diagram for the requested data and any additional datasets intended to be linked.**

## Safe outputs

Safe outputs ensure that all research outputs cannot be used to identify data subjects. They typically include 'descriptive statistics' that have been sufficiently aggregated such that identification is near enough impossible, and modelled outputs, which are inherently non-confidential. The purpose of this section is to ensure that:

- Controls are in place to minimise risks associated with planned outputs and publications
- The researchers aim to openly publish their results to enable use, scrutiny and further research.

Please include any plans for dissemination and publication of the data and results arising from your proposal. Please also specify any controls in place to minimise risks associated with publication. Dissemination can take place in a variety of ways and through many mechanisms, including through electronic media, print media or word of mouth.

### Outputs dissemination plans

#### How will proposal findings be disseminated, to what audience and in what format?

Please describe how you plan to disseminate the results from your proposal. As the public might not read scientific literature or attend conferences, please consider how the results or findings will be disseminated to the wider public and how this fits with the public benefit of the proposal. Please indicate if you plan to publish your findings in an open access journal.

#### Please include any milestones for outputs dissemination.

Provide an outline of your plan, including timeline for output dissemination.

#### What steps will be taken to ensure that individuals cannot be identified? Please describe what disclosure control policy will be applied. (300-word limit)

Please describe the steps you will take to ensure the confidentiality of the data when disseminating or publishing your findings. This may include the application of disclosure control procedures, aggregation of data or other approaches.

## Archiving

Please state the date until which you will retain the data

Please indicate the reason for this date

What method of destruction will be used when this period has expired?

What evidence will be provided that destruction has occurred and when?

---

## Enclosure checklist

- Training
- Please confirm enclosure of CVs for all team members listed
- Evidence of independent peer review (if required)
- Evidence of REC approval (if required)
- Dataflow diagram (if required)

---

Once you have submitted your application, we will email you to acknowledge receipt. The INSIGHT team will then review your application and respond, normally within five working days.

# Annex 1: Additional individuals

Full name

Job title

Organisation (with which employment contract is held)

Email address

ORCID iD

## Role

A role is a function that the applicant plays on the project. Examples might include:

**Principle Investigator** - the primary individual responsible for the project (this would usually be the Primary applicant, but please identify any co-Leads here if appropriate) - Collaborator - co-applicant, project lead e.g. statistics/ technical lead etc

**Team member** - Project Manager/ Coordinator/ Administrative support

Principal Investigator

Collaborator

Team member

Other (please state):

Will this person access the data requested?

Yes

No

Has this person undertaken professional training or education on the topic of Information Governance?

Yes

No

(see [Information Governance](#) ↗)

Please provide details of most recent relevant training attended or planned.

[Return to 'Safe people' section](#) ↗

## Annex 2: Additional sponsor

**Associated applicant** Please confirm the name of the applicant linked to this sponsor organisation.  
This information will be used to form the legal documents.

**Organisation name** Please provide legal name to appear on legal documents

**Registered address** (line 1)

**Registered address** (line 2)

**City**

**Postcode**

**Country**

**Sector**      NHS      Academia      Charity      Other, please state

**Additional details** Please provide any additional information

### Contracts person

Please provide a contact email address for the sponsor organisation or the Lead person to liaise with regarding agreements.

**Full name**

**Contact email address**

**Contact telephone**

[Return to 'Sponsor information' section](#) ↑

## Annex 3: Additional organisations

ICO registered name of organisation

ICO registered number

What type of security does this organisation have in place?

Data security and Protection Toolkit (DSP Toolkit)

SLSP

ISO 27001

Other (please state):

[Return to 'Safe settings' section ↗](#)