HDRUK
Health Data Research UK

HDRUK
Health Data Research UK

HDRUK
Health Data Research UK

# Contents


Innovation Gateway 1

Introduction 6

- Fast tracking innovation in health data 6
- The Gateway - Innovation Gateway self-directed tour 6
- The Gateway facilitates discovery and understanding of available datasets 7
- The information held in the Gateway 8

Making data discoverable and accessible via the Gateway 13

- Creating an account on the gateway 14
- Creating and managing teams on the Gateway 16
- Adding a New Dataset to the Gateway (Onboarding metadata) 17
- Managing datasets on the Gateway 20
- Versioning and Archiving Datasets 23

Grouping your datasets for easy discovery with collections - Video tutorial: Collections 25

- Why create a collection? 25
- Collection categories 25
- Editing a collection 29
- Archiving a collection 30
- Deleting a collection 30
- Future plans 30

Why and how researchers and innovators make data access requests 31

- How applicants request access - Video tutorial: How to submit a data access request 31
- Data Access Request (DAR) Forms 31
  - Short form Data Access Requests 32
  - Long form Data Access Requests 32

5 safes DAR form authorisation workflow 35

- How Data Custodians manage and process data access requests 35
  - Creating a team and adding members – Video tutorial: Setting up a team, creating workflows and reviewing an application 35

Customise your “How to request access” information 35

Customise your Data Access Request form 36

Creating a Data Access Request workflow - – Video tutorial: Setting up a team, creating workflows and reviewing an application 36

Reviewing and managing a 5 safes Data Access Request application – Video tutorial: Setting up a team, creating workflows and reviewing an application 37

HDRUK
Health Data Research UK

# Purpose

The purpose of this document is to provide new data custodians (or a new member of a data custodian team) with an overview of the Innovation Gateway ("Gateway"). It walks them through the functions and processes that they will typically utilise as they upload and manage their metadata on the Gateway as well as providing access to that data and the interactions they will have with the users requesting access.

The scope of the document includes the entire set of data custodian processes including: onboarding of metadata and technical metadata to the Gateway; supporting data utility; creating collections; and responding to Data Access Requests (DARs). The ultimate goal is to enable the reader to understand the full user journey from a data custodian perspective, plan adequate resources and become as self-sufficient as possible.

This Playbook should be used alongside the other guidance resources available through the Gateway Community site including *How-to videos* as well as the Community Forum.

# Introduction

## Fast tracking innovation in health data

At HDR UK we passionately believe that every health and care interaction and research endeavour will be enhanced by access to large scale data and advanced analytics. Challenges to human health and health system sustainability are increasing globally. By making health data available to researchers and innovators we can better understand diseases and discover ways to prevent, treat and cure them.
The UK has some of the richest health data and unique research capabilities in the world, making it possible to make national-scale improvements to health and care. However, it is not always easy for researchers and innovators to discover the data that is available to support their research/work and to then apply for that data. This not only restricts innovation, but also results in a lot of irrelevant inquiries, and therefore wasted time, to already stretched data custodians.

We built the Innovation Gateway (the 'Gateway') to address these issues, making it easier for people to find the data to support their work and to simplify and streamline the process of requesting access requests to data for both custodians and researchers. All with the aim of accelerating research.

## The Gateway - Innovation Gateway self-directed tour

The Gateway is a web portal providing a common entry point to discover and request access to UK health datasets for research and innovation. It holds detailed information about the datasets held by members of the UK Health Data Research Alliance, such as a description, size of the population, and the legal basis for access. It does not, however, hold or store any datasets, patient or health data.

Innovators can search for research projects, publications, training courses and health data tools, such as those related to COVID-19. Researchers can create and collaborate on projects, interact with data custodians, request access to datasets, and connect to other researchers via the community forum.

The information provided about the datasets is supplied by the relevant data controller and is continually updated.

## The vision

The Vision for the Gateway is to be the primary go-to resource for:

- Discovering data, tools, best practice and collective knowledge and experience related to accessing resources to further health research.
- Posting research papers, datasets, tools and sharing learning and other resources knowing that the Gateway community will enable it to reach the broadest set of interested parties.

It contains metadata for over 500 datasets (and growing) that are discoverable using the Gateway's easy-to-use interface and as its popularity grows so will the content.

HDRUK
Health Data Research UK

## The Gateway facilitates discovery and understanding of available datasets

As a Data Custodian, making your data discoverable through the Gateway provides the single simplest action you can take to:

1. help your data reach a significant and growing community of researchers and innovators from across academia and industry;
2. list your data alongside other resources (datasets, papers and tools) enabling innovators and researchers to identify novel linkages between datasets to support their research;
3. help you manage access requests to use your data through a single electronic interface and streamlined process.

This third point is significant, especially if you currently operate a paper-based Data Access Request (DAR) process which are typically onerous to operate and problematic to maintain. The Gateway provides a standardised, electronic Data Access Request (DAR) process (based on the 5 safes model from the Office for National Statistics) for innovators and researchers to request access to your data from you - the data custodian - and for you to review and approve/reject the request. By removing the need for paper processes, the Gateway provides supportive functionality to both data requesters and custodians by enabling:

- secure communications between requestor and data custodian in advance of submitting a DAR;
- customisable DAR form for the Data Custodian's organisation based on a bank of over 100 pre-defined questions.
- researchers to work together to complete a DAR through the use of embedded collaboration tools;
- the ability for both custodians and requesters to manage and monitor DARs.

# The information held in the Gateway

The Gateway does not hold a copy of Data Custodian data. Instead, it stores summary information used to describe each of the datasets that the Data Custodian holds. This is commonly known as Metadata and contains information such as: where the dataset has come from, a description of the dataset, the time-period, and the geographical areas the dataset covers.

This metadata is helpful for a researcher to understand whether a dataset will be of use to them without them having to see the data itself. Obviously, the more metadata that is provided about each dataset and the more accurate it is, the easier it is for the researcher or innovator to decide whether it is helpful for their work. This metadata also allows them to check they are eligible for access before making a request from the Data Custodian. This helps to reduce the number of invalid requests that custodians have to handle.

Each dataset entry on the Gateway has its available metadata described under the following headings - (Video tutorial: Viewing dataset details):

## Summary and metadata quality score

The top of each entry provides a summary banner containing the name of the dataset, its publisher/custodian, how many times the entry has been viewed and a metadata quality score.

As described earlier, higher quality metadata makes it easier for users to decide if a particular dataset is of interest to them. To help Data Custodians and users to easily see the quality of a dataset's metadata, a metadata richness score is calculated for each dataset record on the gateway and displayed on the gateway as Gold, Silver, Bronze or Not rated.

Metadata quality is calculated across four categories:

- Completeness Percent
- Weighted Completeness Percent
- Error Percent
- Weighted Error Percent

Together they produce a Weighted Quality Score and from that a quality rating is applied based on the table below.

| Weighted)/Quality Score | Rating |
|---|---|
| < 66 | Not rated |
| >66 & <= 76 | Bronze |
| - >76 & <=86 | Silver |
| >86 | Gold |
| | Platinum |

More information on how the scores are calculated is provided at the following link https://github.com/HDRUK/datasets/tree/master/reports#hdr-uk-data-documentation-scores

Below this summary banner are a number of tabs providing additional information about the dataset. These are described below.

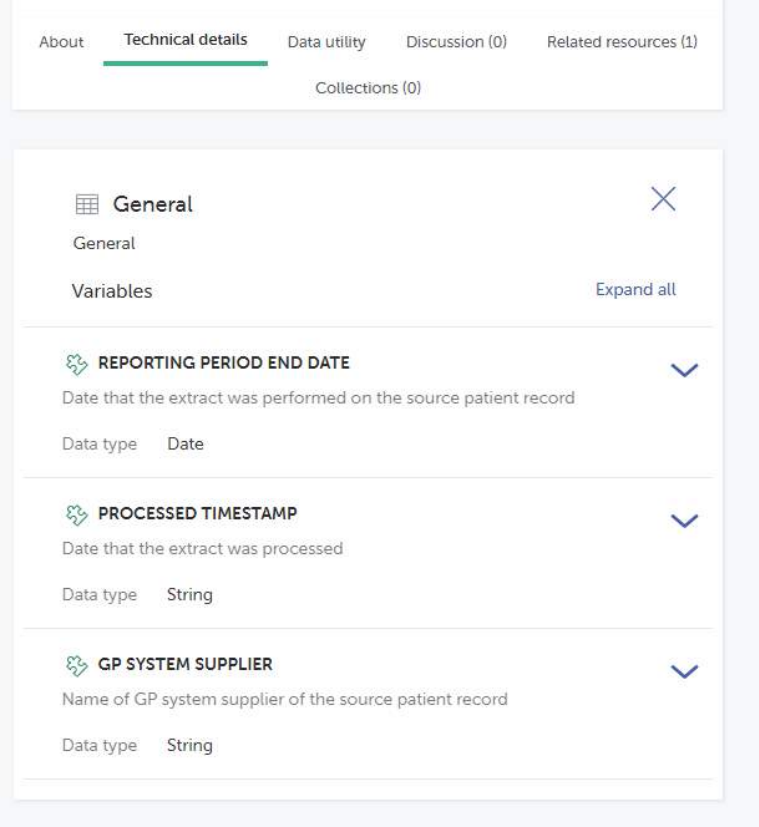
## About

This tab provides the high level overview of the available dataset broken down as follows:

- **Abstract**
- **Description**
- **Details**: Latest release date; Publishing frequency; and Resource creator
- **Coverage**: Time period covered; Any time lag in the dataset; Geographical Coverage; Typical age range of participants in the dataset; Availability of physical samples in the dataset; Any follow-up period for patients in the dataset; Any pathway represented in the dataset.
- **Formats and standards:** A list of semantic annotations used within the dataset; Any standardised data models used; Language of the dataset; Format of the dataset e.g. csv
- **Provenance:** Purpose for which the dataset was collected; Whether derived datasets are available and if so the type or derivation used.
- **Data access request:** Jurisdiction; Data controller; and Data Processor, if relevant
- **Related resources:** Any keystone paper associated with the dataset; any active projects using the dataset; and any tools or models that have been created for the dataset
- **Linked datasets:** Any datasets linked to this dataset

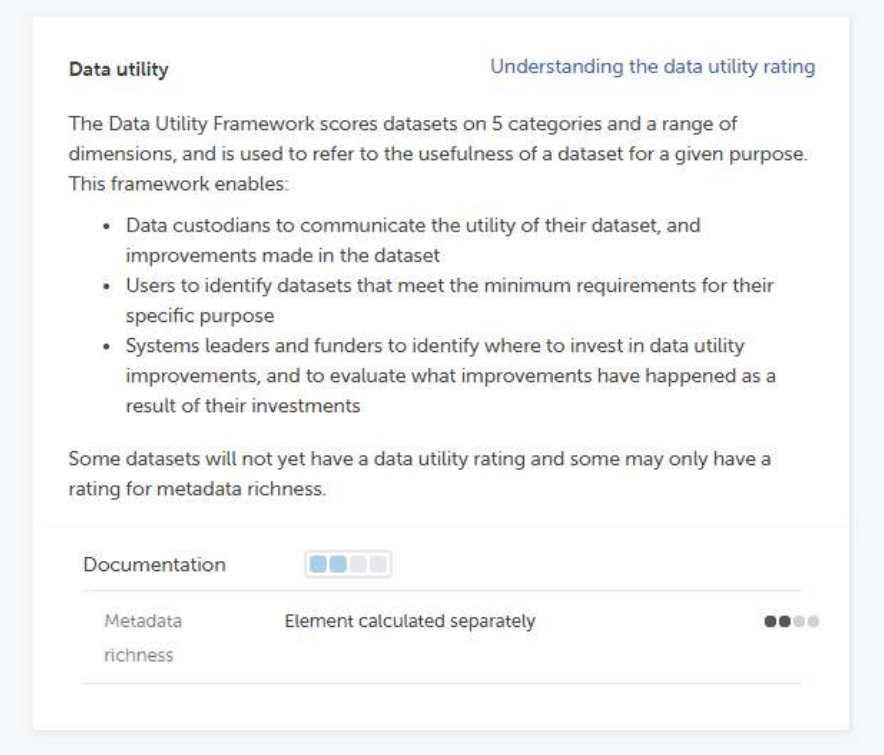
## Technical details

Technical metadata provides information on the technical properties of the dataset file associated with the data model including for example, field name, data types. An example is provided in the image below.



## Data utility framework

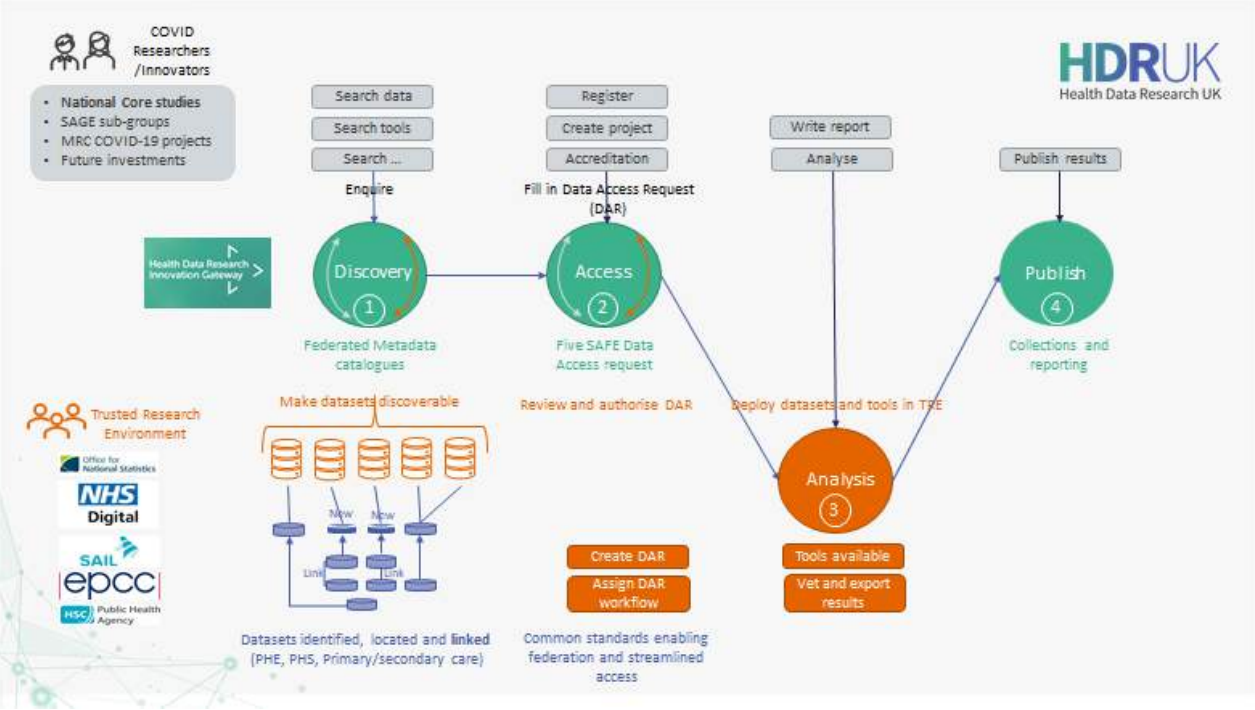
The Data Utility Framework scores datasets on five categories and a range of dimensions and is used to refer to the usefulness of a dataset for a given purpose.  The image below explains data utility in more detail.




More detail on Data Utility Evaluation is provided at the following link: https://www.hdruk.ac.uk/help-with-your-data/ways-to-improve-data-quality/data-utility-evaluation/

# Data custodian journey and user journey

Now that we have explained the metadata that is stored on the Gateway, it is helpful to set out the high-level user journeys undertaken by researchers and custodians as part of the Innovation Gateway.



# Making data discoverable and accessible via the Gateway

The purpose of this section of the document is to give an overview of how a data custodian or processor can make their datasets discoverable on the Health Data Research Innovation Gateway ('the Gateway').

The Gateway does not hold any custodian data; instead, it stores summary information used to describe each of the datasets[1] that the Data Custodian holds. This is commonly known as *metadata*, and contains information such as: where the dataset has come from, a description of the dataset, the time-period, and the geographical areas the dataset covers.

This metadata allows a researcher to understand whether a dataset will be of use to them without having to see the data itself. The more metadata that is provided about each dataset and the more accurate it is, the easier it is for a researcher or innovator to decide whether it is helpful for their work. This metadata also allows them to check if they are eligible for access before making a request from the Data Custodian; this helps to reduce the number of invalid requests that custodians must handle.

For a data custodian to make their datasets discoverable on the Gateway, they will need to load metadata about their datasets onto the Gateway - a process called *onboarding*. This document will outline the steps required for a custodian to onboard new datasets or to manage existing datasets on the Gateway:

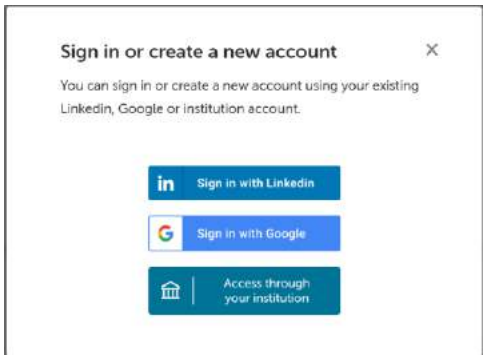
- Creating an account on the Gateway.
- Creating and managing teams.
- Managing and versioning datasets.
- Adding a new dataset to the gateway.

For more in-depth information around the metadata that is stored on the Gateway, as well as the specifications required for each metadata field can be found on the HDR UK Github page.

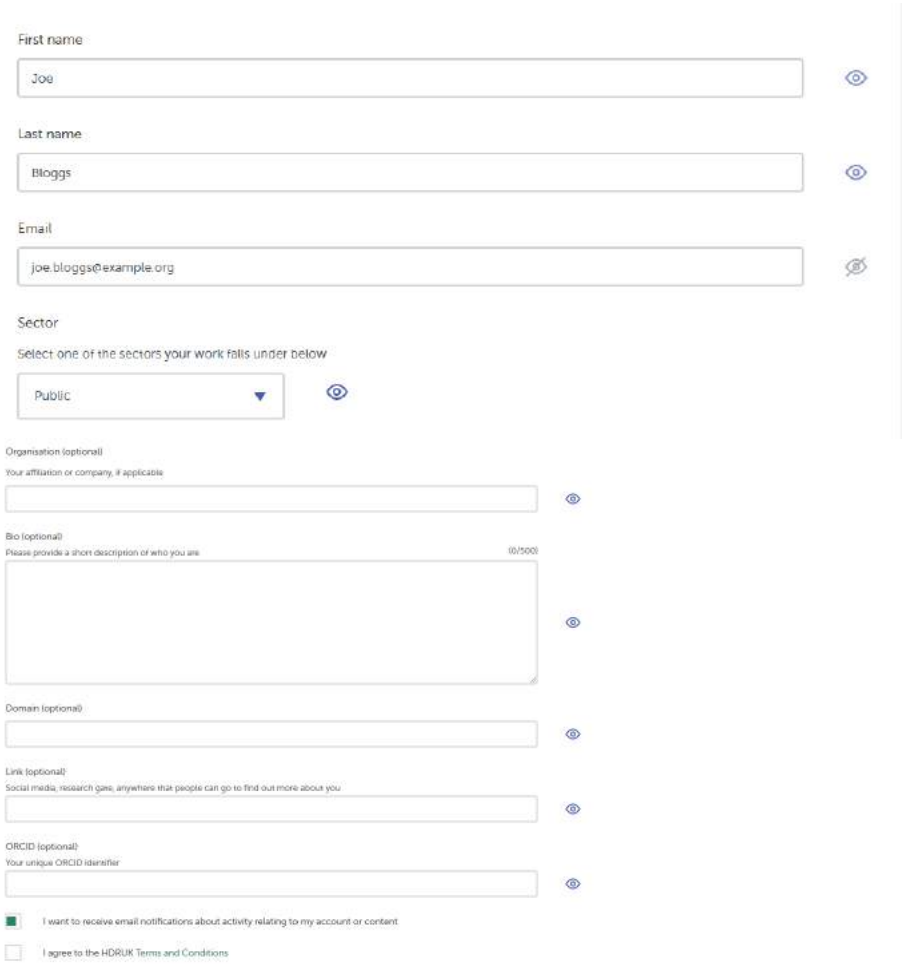
[1] A dataset is a collection of related individual pieces of data but in the case of health data, identifiable information (e.g. name or NHS number) is removed and data is de-identified where possible

# Creating an account on the gateway

To sign up for an account on the Gateway, first select the 'Sign In | Sign Up' option in the top-right corner of the Gateway homepage. Here, you can create an account on the Gateway either using a LinkedIn or Google account. Alternatively, you can also sign in through your institutional login.



Next, you will need to provide some additional details to create your profile: your name, email address, and sector. There are also several optional fields that you can choose to include in your profile such as your organisation, a bio, a link to a research or social media profile, or an ORCID identifier.



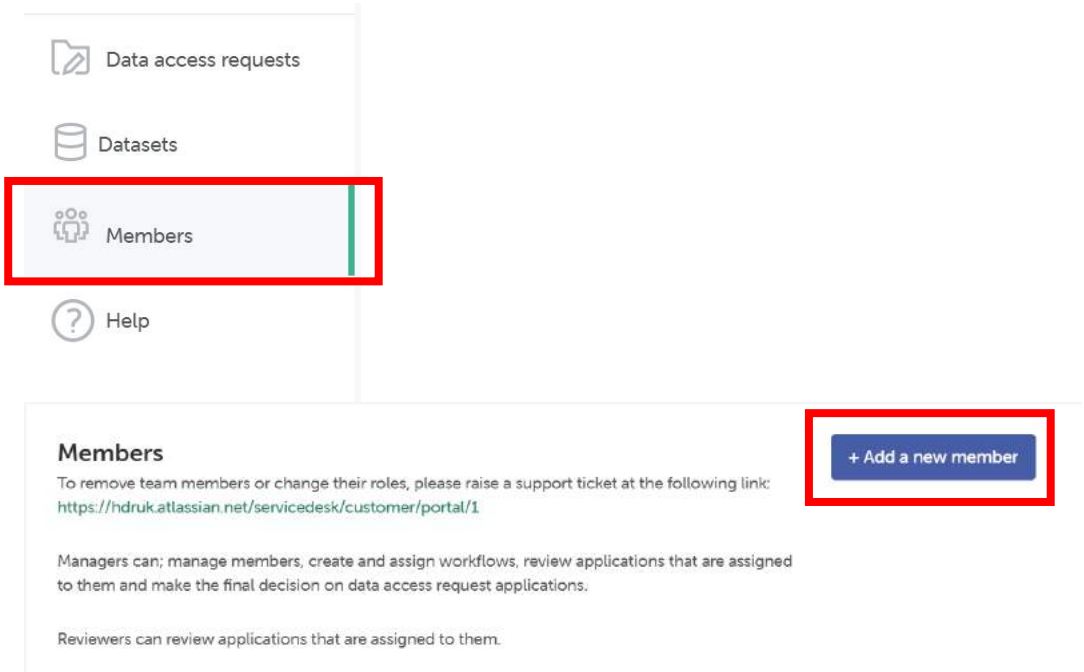
# Creating and managing teams on the Gateway

In order to onboard and manage metadata on the Gateway, you must first be a member of a team. Currently, to create a team, email support@healthdatagateway.org, or raise a support ticket at https://hdruk.atlassian.net/servicedesk/customer/portal/1.

Once you have created your team, you can manage your team members in your account section. To do this, start by navigating to your account page from the top right corner of the gateway homepage.

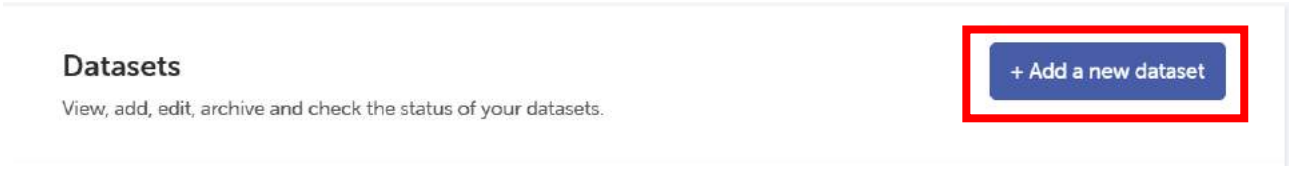
Here you can select your team from the dropdown menu beside your name on the left hand side of the page. You are now acting on behalf of your team.

In the members tab of this page, you can add new members to your team, assuming that they also have an account on the Gateway. To remove team members or change their roles, please raise a support ticket at the following link: https://hdruk.atlassian.net/servicedesk/customer/portal/1.



# Adding a New Dataset to the Gateway (Onboarding metadata)

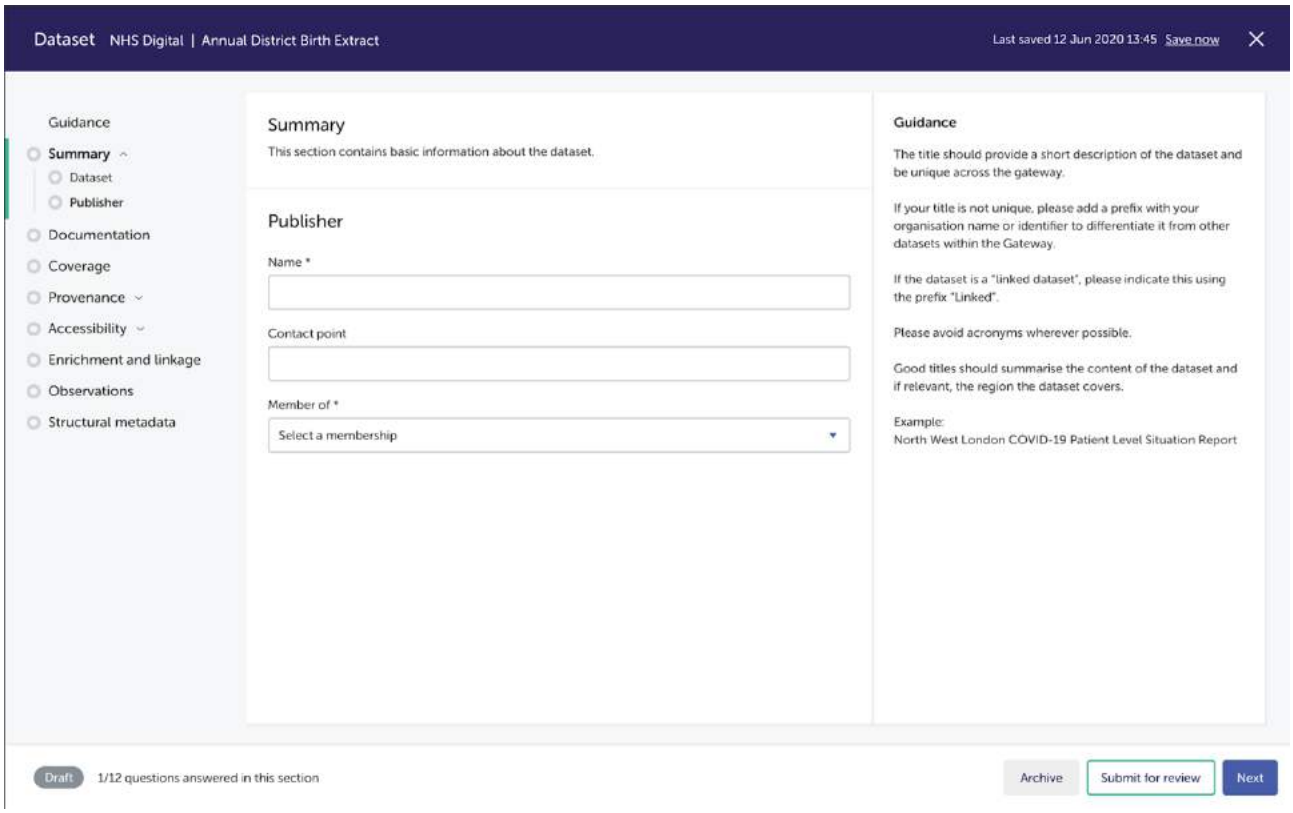
Once you are set up on your team, select 'Datasets' from the menu on your account page. Here, you will also see the option to 'Add a new dataset', which will bring you to the metadata onboarding platform.



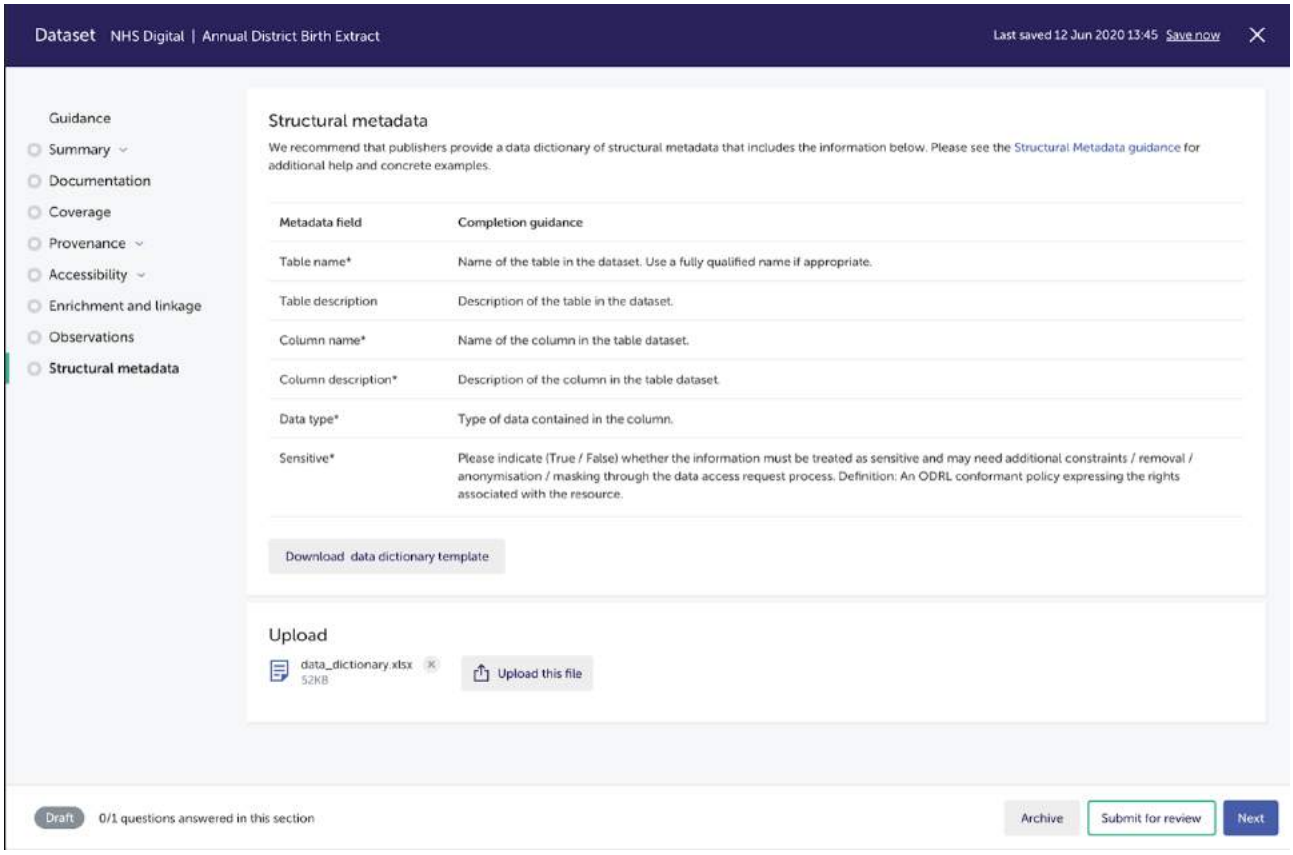

You will be guided through the process of filling out the metadata fields according to the specifications for each of the 8 sections, which can be seen on the left hand side of the screen. The action bar at the bottom shows the status of the dataset, your progress through the onboarding process, as well as options to archive the dataset, submit it for review, or to move on to the next section.

Before you begin, you will be presented with information on the approval process, how to ask HDR UK further questions, best practice and further guidance, uploading multiple datasets, and understanding the metadata score.

In each section, you can see the title, and explanation of what the section is for. On the rest of the page you can see the input fields and guidance for completing them. For further guidance on completing each of the fields select the '?' button beside the input. The guidance will then appear on the right hand side. An asterisk beside the input field means that these inputs are mandatory. You must complete all mandatory fields before you are allowed to submit your application for review.

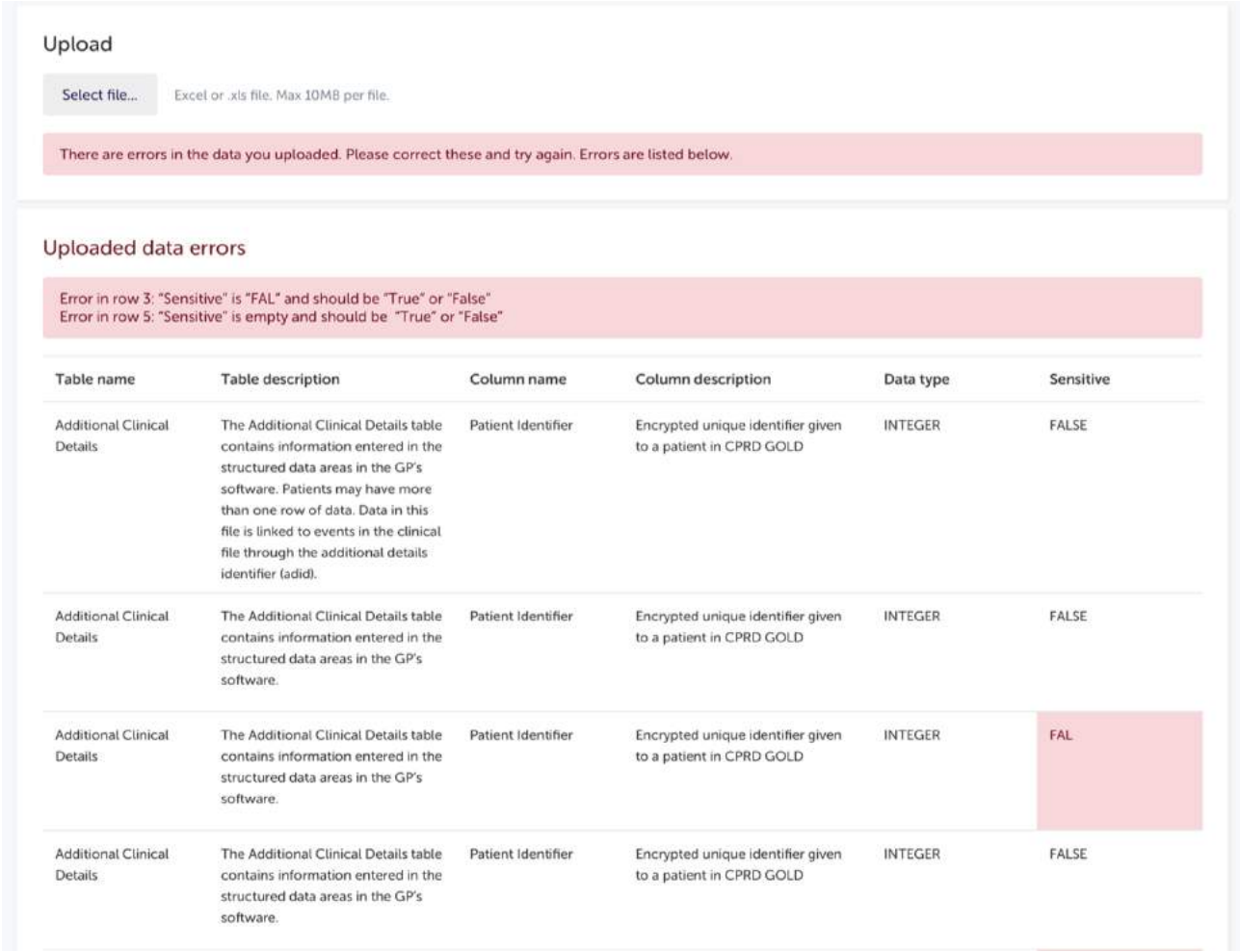


The last section of the onboarding process, structural metadata, is for adding a data dictionary. The data dictionary should be submitted using the data dictionary template provided in this form. To download it, click ‘download data dictionary template’. It will provide you with a template that lists out the following fields. Guidance is available here to help you fill in the template.




Once you have completed the data dictionary spreadsheet, you can attach it using the ‘select file’ button under Upload. Only Excel or .xls files are accepted, and there is a file size restriction of 10Mb. When you are happy with your file selection, click ‘upload this file.’

Once the file is uploaded you will either see a successful upload, or the system will identify any errors. If there are errors you will receive a message informing you that errors have been found and that you need to upload the file again with these errors corrected. The entire spreadsheet will be available for you to view, and at the top it will indicate in exactly which row and column the error is. The errors are highlighted in red.



Upload

Select file... Excel or .xls file. Max 10MB per file.

There are errors in the data you uploaded. Please correct these and try again. Errors are listed below.

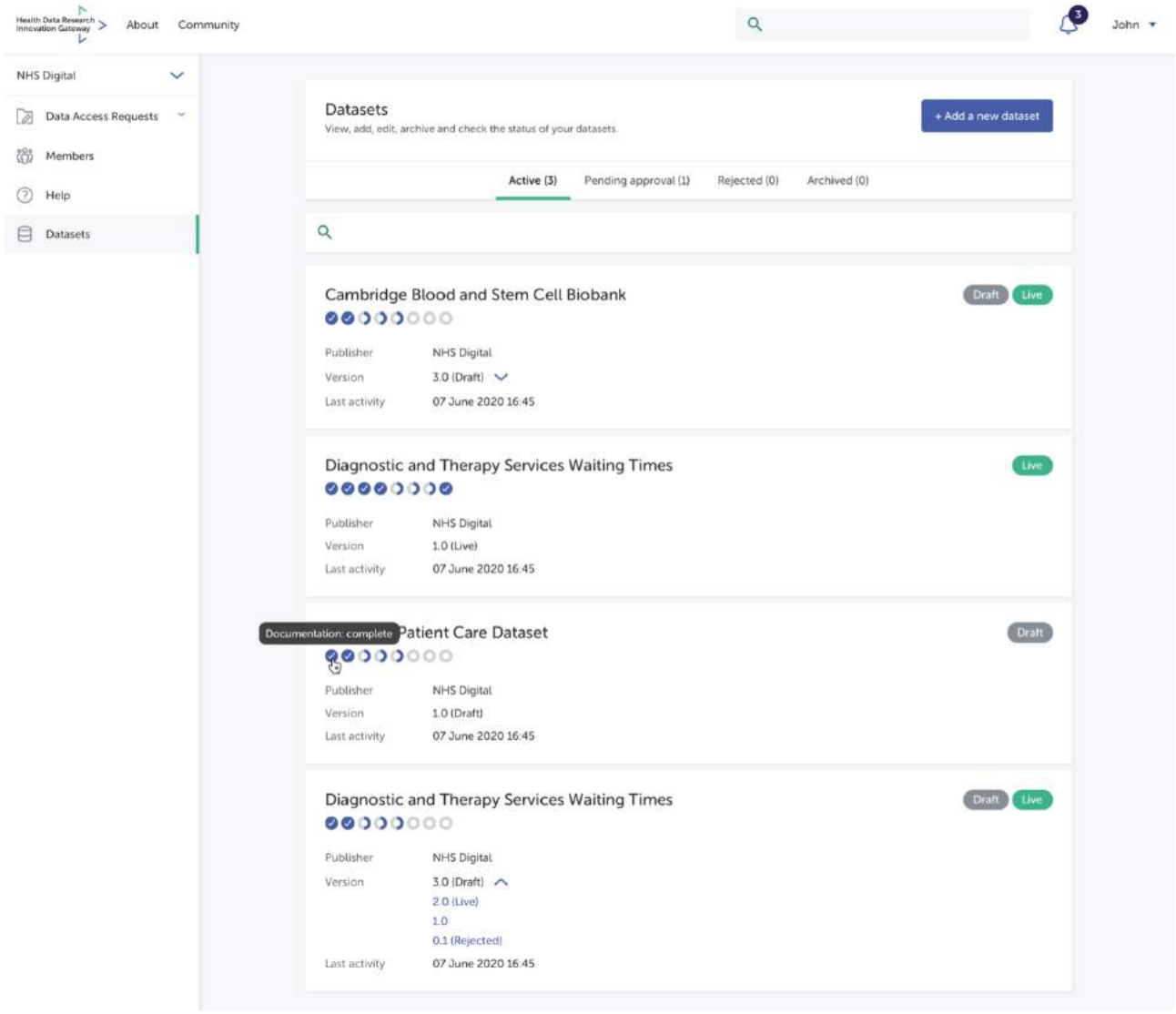
Uploaded data errors

Error in row 3: "Sensitive" is "FAL" and should be "True" or "False"
Error in row 5: "Sensitive" is empty and should be "True" or "False"

| Table name | Table description | Column name | Column description | Data type | Sensitive |
|---|---|---|---|---|---|
| Additional Clinical Details | The Additional Clinical Details table contains information entered in the structured data areas in the GP's software. Patients may have more than one row of data. Data in this file is linked to events in the clinical file through the additional details identifier (adid). | Patient Identifier | Encrypted unique identifier given to a patient in CPRD GOLD | INTEGER | FALSE |
| Additional Clinical Details | The Additional Clinical Details table contains information entered in the structured data areas in the GP's software. | Patient Identifier | Encrypted unique identifier given to a patient in CPRD GOLD | INTEGER | FALSE |
| Additional Clinical Details | The Additional Clinical Details table contains information entered in the structured data areas in the GP's software. | Patient Identifier | Encrypted unique identifier given to a patient in CPRD GOLD | INTEGER | FAL |
| Additional Clinical Details | The Additional Clinical Details table contains information entered in the structured data areas in the GP's software. | Patient Identifier | Encrypted unique identifier given to a patient in CPRD GOLD | INTEGER | FALSE |

# Managing datasets on the Gateway

In the datasets tab of your account page you can view, edit, archive, and check the status of your datasets. You will be able to view which datasets are active, pending approval, rejected, or archived. This page also gives you an overview of the progress that has been made through the onboarding metadata process for each of these datasets thus far.



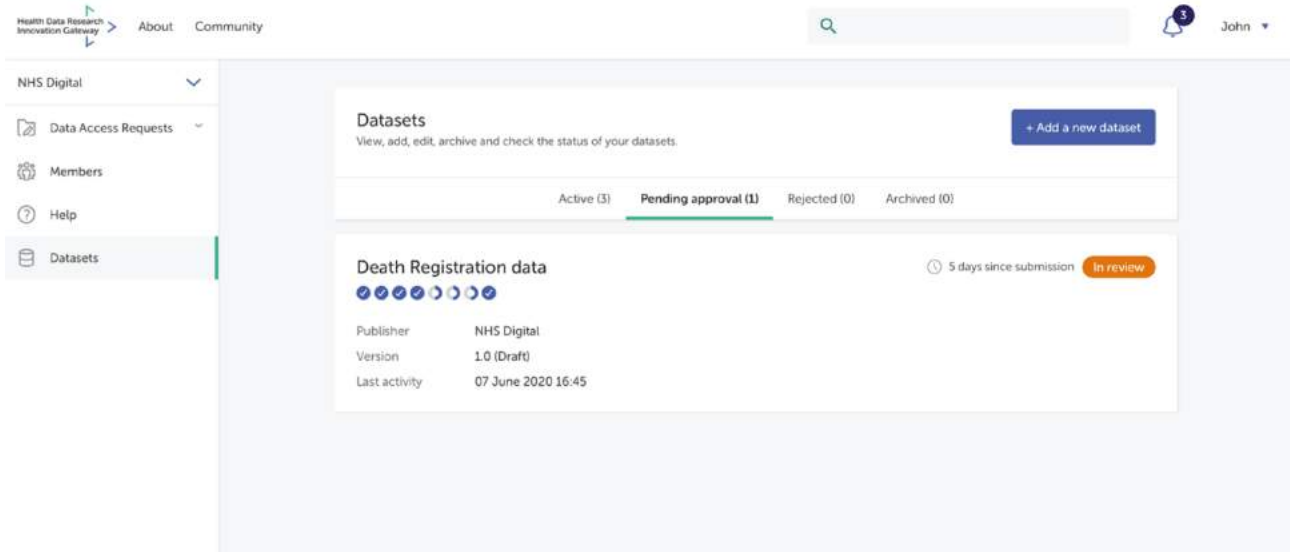

The status tags down the right-hand-side show whether a dataset is in draft, live, or both:

- Live: Dataset is live on the Gateway and is searchable and available to request access.
- Draft: The dataset is not yet live and is yet to be published.
- Both live and draft: There is a live version of the dataset, however a new version is being edited.

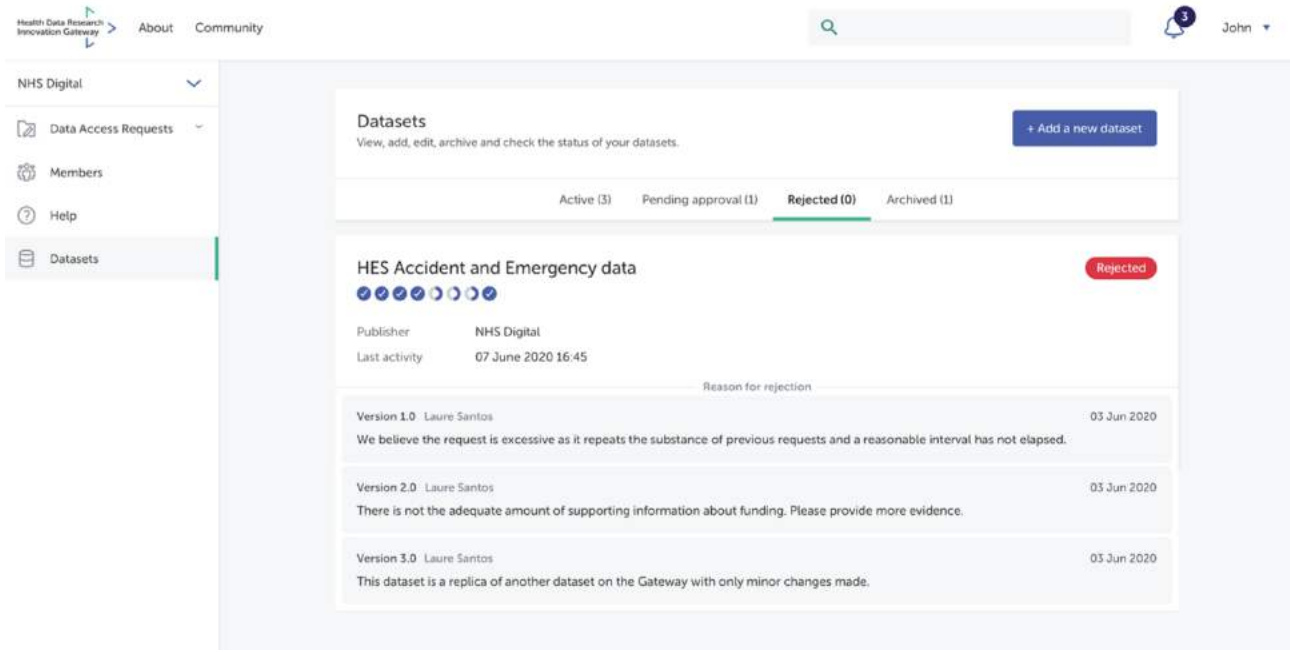
The circles below the dataset titles represent sections of the metadata onboarding process. When a circle is filled in with a tick, it means that it has been fully completed, and a half circle represents a section that is partially completed. An empty circle means you haven't started the section.

You can also see the publisher, the version of the dataset, and the latest activity on the dataset. If there are several versions, select the arrow to see all past versions, you have the ability to click and view these previous versions, including any rejected versions. You can quickly see the status of each version next to the number, for example draft or live.

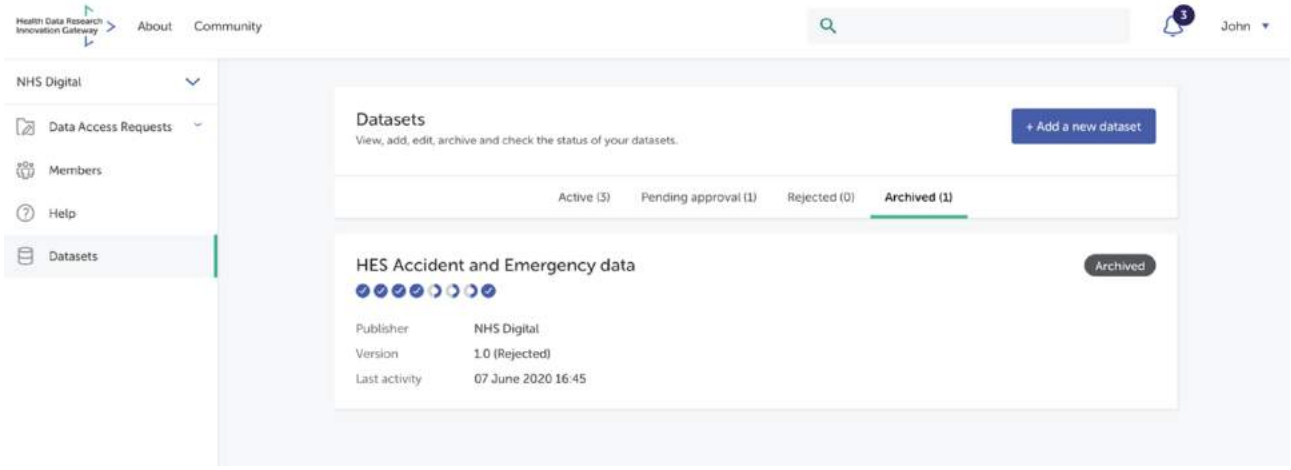
On the pending approval tab you are able to see which datasets you have submitted for approval but have not yet been approved or rejected by the HDR UK admin team. You can see that it's in review with the orange status tag, together with the amount of time that has passed since you submitted it for approval.




On the rejected tab you are able to see any datasets you submitted for approval that have then been rejected by the HDRUK admin team, alongside the reason for the rejection. You can re-submit a rejected dataset by creating a new version and making the appropriate changes.

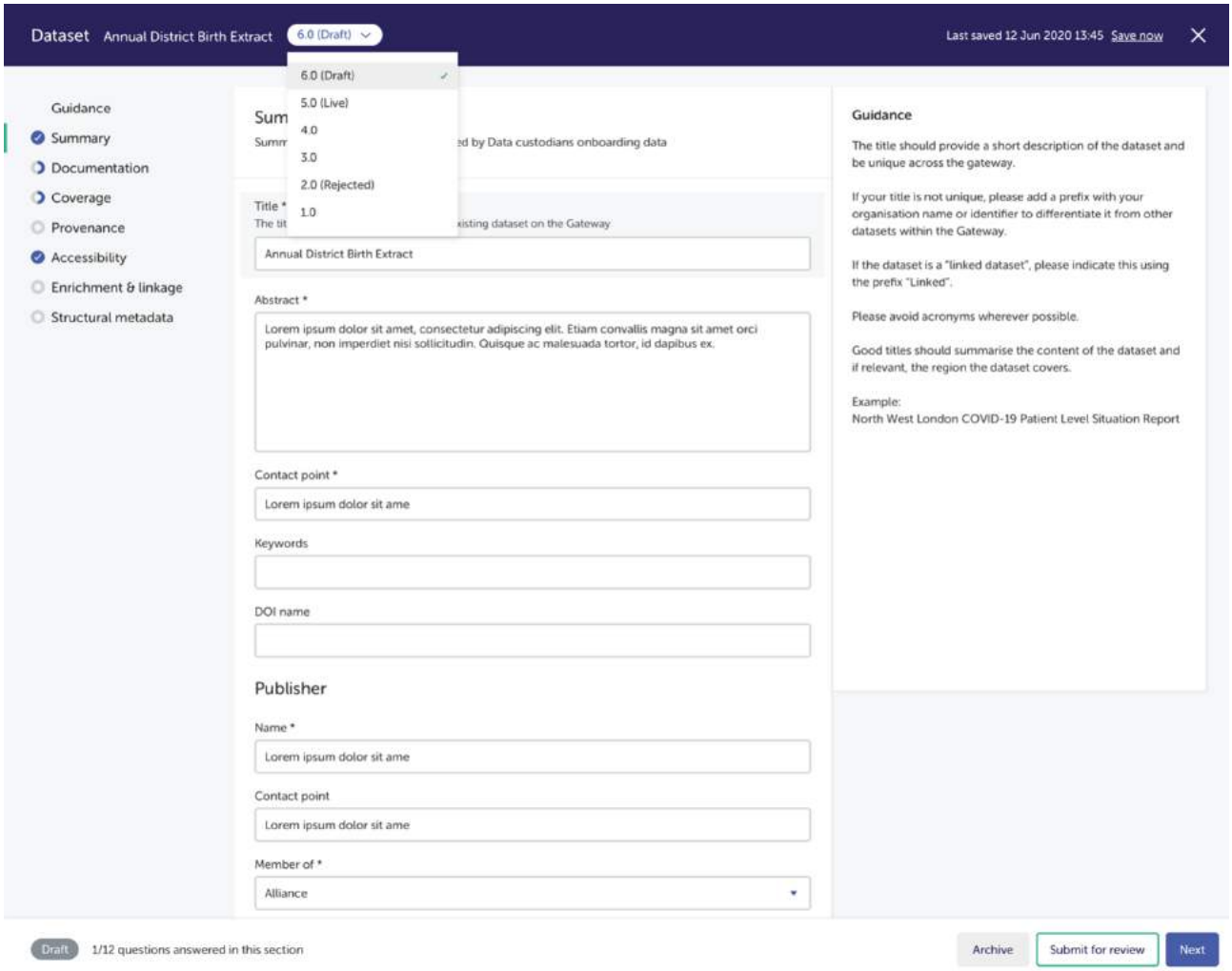


Finally, the archived tab is where you will find datasets that you have archived. Archiving your dataset will remove it from any search results on the Innovation Gateway. However, anyone with a direct link will still be able to access the dataset page, but they will see a message letting them know that it has been archived.



# Versioning and Archiving Datasets

You can create new versions of live or rejected datasets. To do this, select the dataset, and click ‘create a new version’ from the action bar at the bottom of the screen. There will now be a new version of your dataset - as indicated by the chip next to the title in the top blue bar. You can also navigate between versions using this drop down chip.




New versions of datasets will go through the same approval process as a new dataset, and you will receive an email once a decision has been made. The original version of your dataset will remain live and searchable until your new version is approved. If a user has a direct link to the old version of a dataset, they will still be able to access it, but will be informed that it is an older version and they will be prompted to visit the newest version.

If a dataset you have submitted is rejected, you may create a new version from this rejected dataset. Simply navigate to the rejected tab and select the dataset you would like to re-submit. Select the ‘create a new version’ button and make all necessary changes. Please pay close attention to the feedback that the admin team provided you with when they rejected this dataset. You can then submit this updated dataset for review.

Archiving a dataset will remove your dataset from the search results. Users with a direct link will still be able to navigate to this dataset, and they will see a message explaining that it has been archived.

To archive, click on a live dataset and select the archive button in the action bar. You will receive a warning letting you know what happens when you archive a dataset. Select ‘archive this version.’ You can see from the status tag that it is archived. It will also appear in the archived tab of your dashboard.

You can un-archive a dataset by selecting your chosen dataset from the archived tab. Select ‘un-archived’ from the action bar. This will make the dataset live and searchable again.

# Grouping your datasets for easy discovery with collections - Video tutorial: Collections

## Why create a collection?

Collections are used to categorise related resources in one single space. Anyone on the Gateway can create collections and share them with other researchers. Collections can be populated with tools, projects, papers, people, courses and datasets and can be created by any user on the Gateway. They can also revolve around a certain disease - such as COVID-19, a data research hub - such as INSIGHT or any other grouping you require.

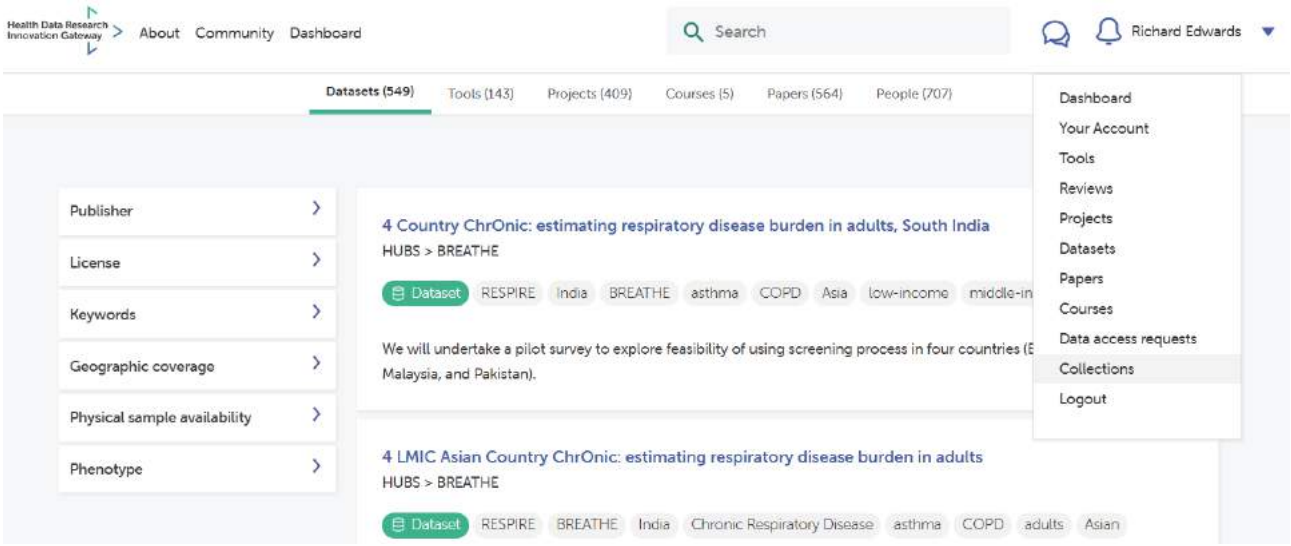
On the homepage of the Gateway you will find the featured collections. These have been selected by or suggested to HDRUK to be featured, in order to increase their visibility. You may submit your own collections to be featured via a feature request ticket. But by default, collections are private and only you can see collections you created. If you would like to share one of your collections with a collaborator, you can simply send them a link to the collection. In addition, if you think that your collection would benefit others by being published on the Gateway home page then contact a member of the Gateway team with your reasons and they will consider the request.

## Collection categories

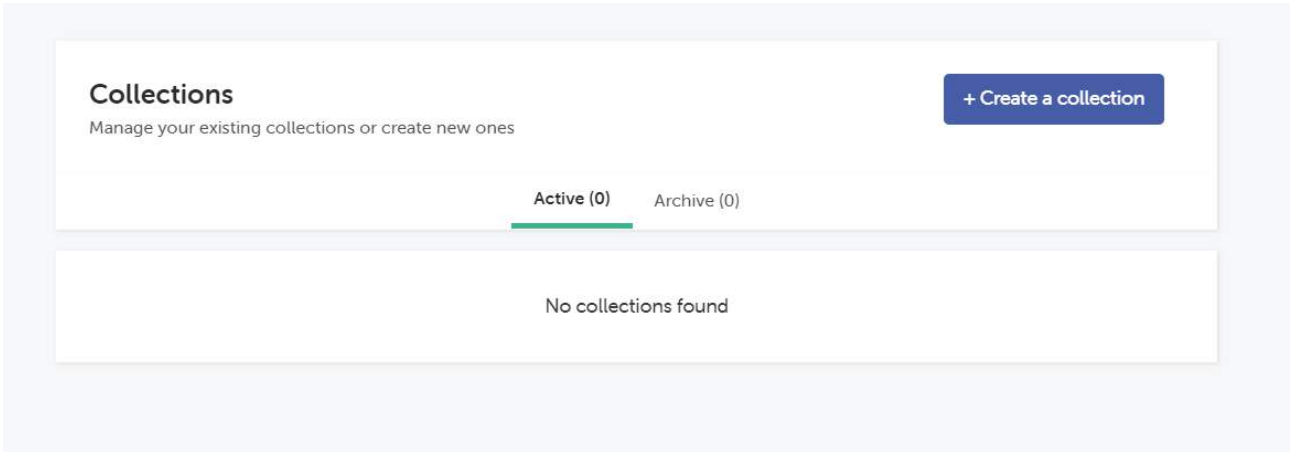
Collection categories are another feature of the Innovation Gateway allowing you to group collections by theme. For instance, in the National Core Studies category you can find 6 collections that relate to one of the National Core Studies. You can view all featured collections and categories with the 'Explore all' button, or through the 'view all featured collections' button via the Collections tab at the top of the homepage.

# How to create a collection

1. To create your own collection, you first need to register onto the Gateway and create an account.
2. Sign in and navigate to your account via the top right hand corner. You will see the collections tab in the left hand navigation on your account. Here you will be able to manage your existing collections or create a new one. You can see which collections are active and which have been archived.

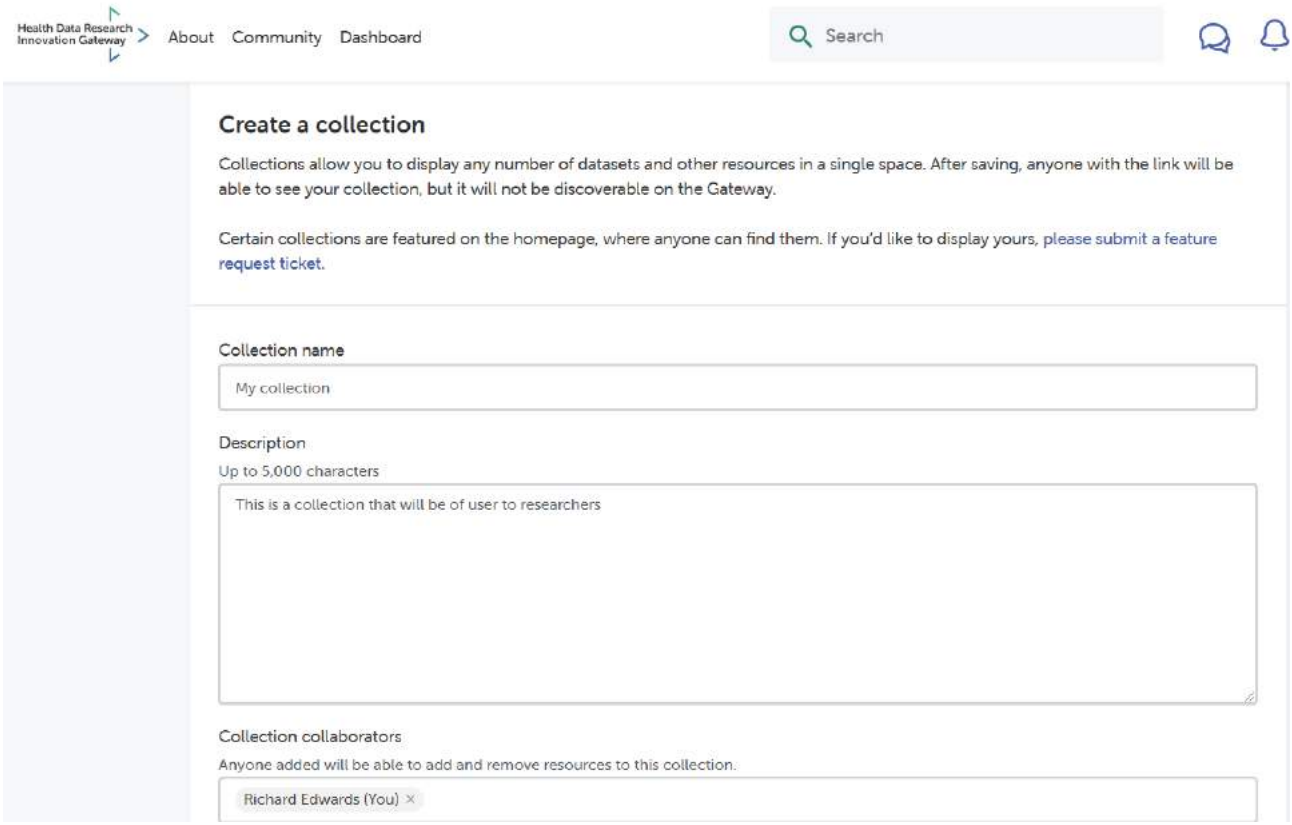



1. To create a new collection click the 'create a collection' button.

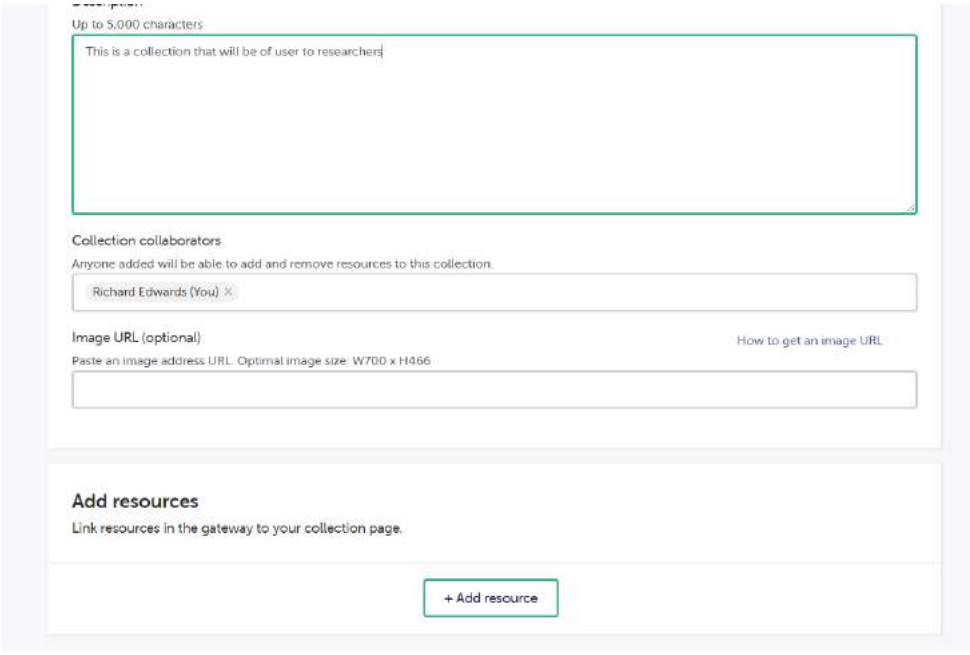



You will be presented with the Collection creation form.

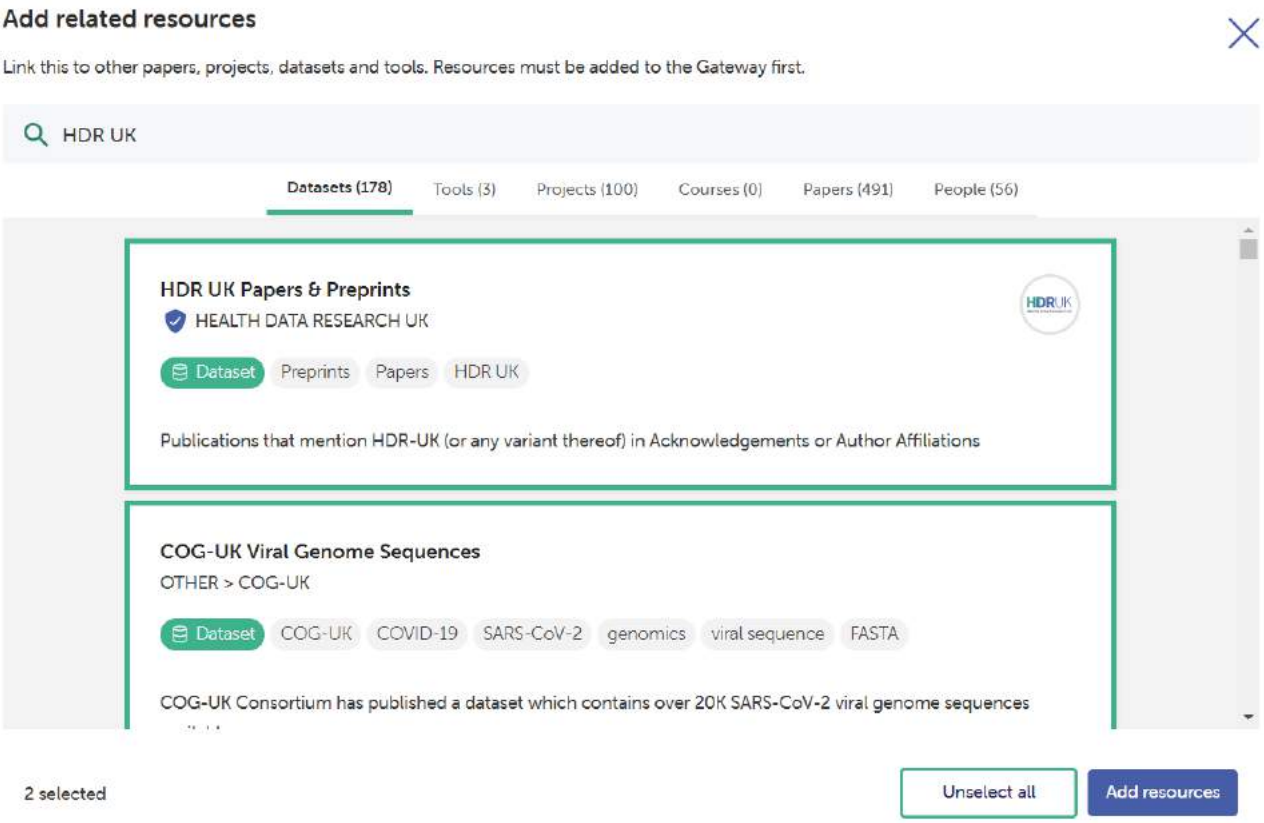
2. Enter the name for your collection, provide a description and add collaborators. You can also add an image using an image URL. Note that collaborators need to also have a Gateway account and will be able to add and remove resources to the collection.




3. You now need to add resources to your collection. First click the + Add resource button to open a search modal.

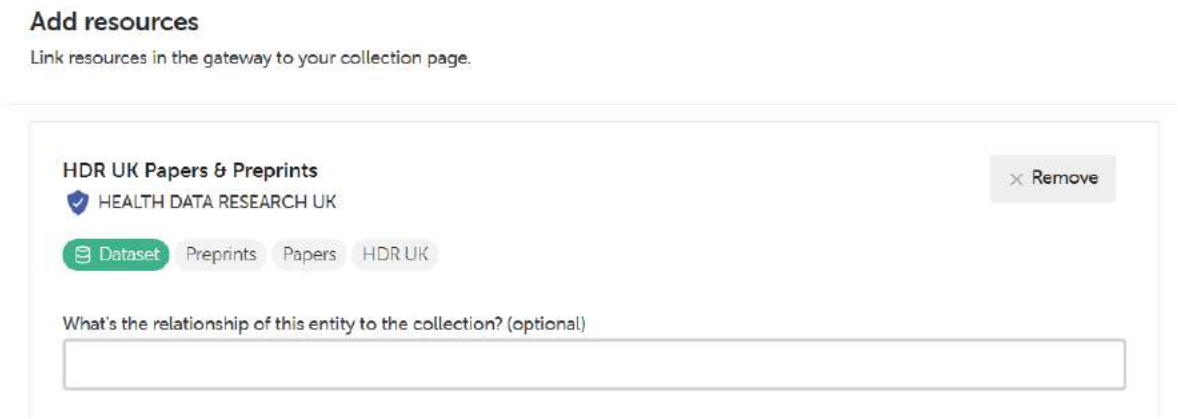



Now search for the resources in the same way that you do in the main search page. Results are divided into categories like the main search page - datasets, tools, projects, courses, papers and people.

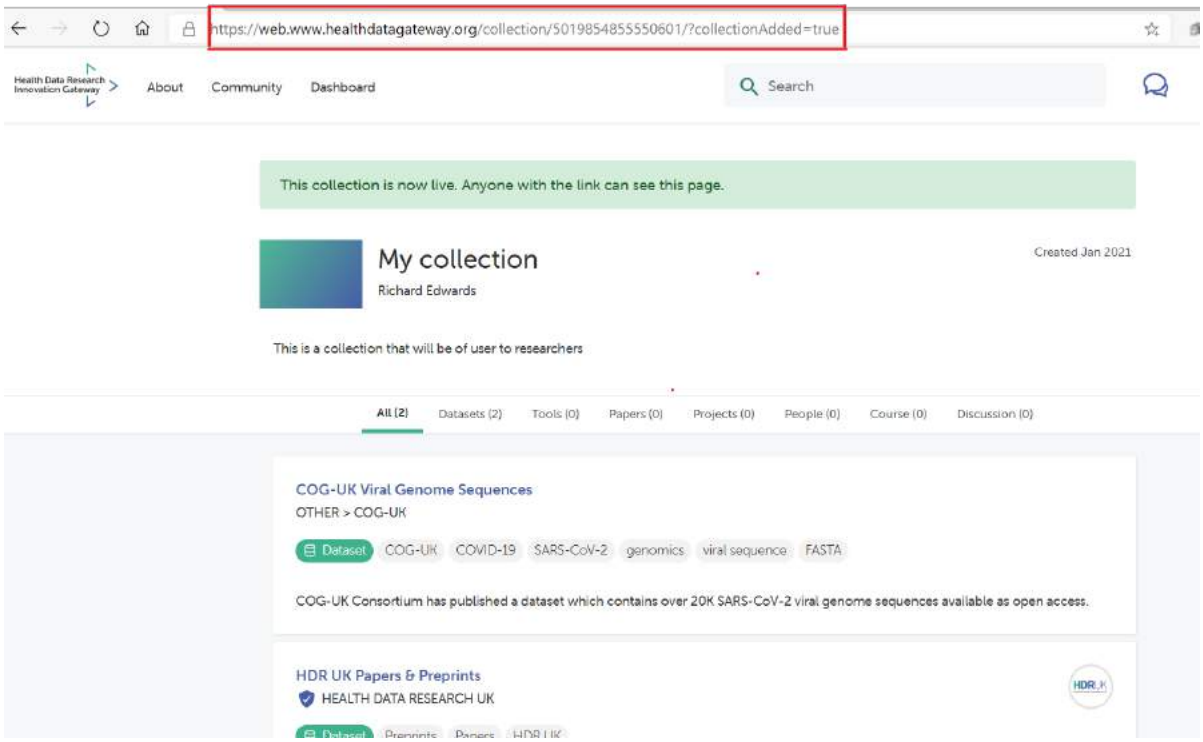



Select as many as you like from all the different categories. The number of resources you've selected is shown on the bottom left. You can unselect resources individually by clicking on them again. Alternatively, if you want to unselect all of the resources you can use the Unselect all button. Once you are happy with the list of resources from your search click Add resources. You will be returned to the collection screen with your selected resources added.

4. You can repeat the process above to add more resources to your collection until you are happy you have included all the items you require.
5. You can remove any resources you no longer want to include in the collection by clicking the Remove button in the top right of each resource summary.

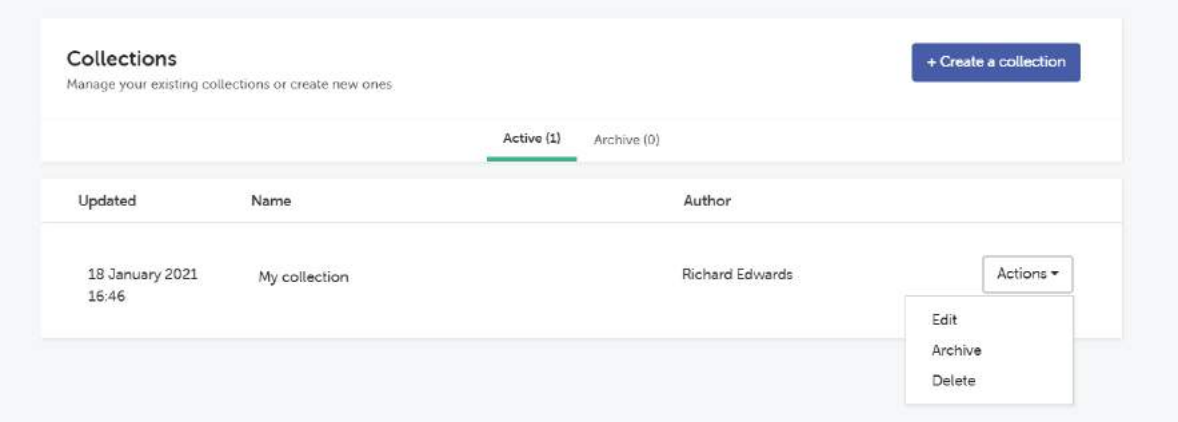


6. When you are finished, you can ‘save’ the collection using the button at the bottom of the page. As shown by the success banner on your collection page, you can share this collection with anyone by selecting the entire address URL, copying and pasting to the appropriate person. If you want this collection to be featured on the homepage you would do so via a feature request ticket.




## Editing a collection

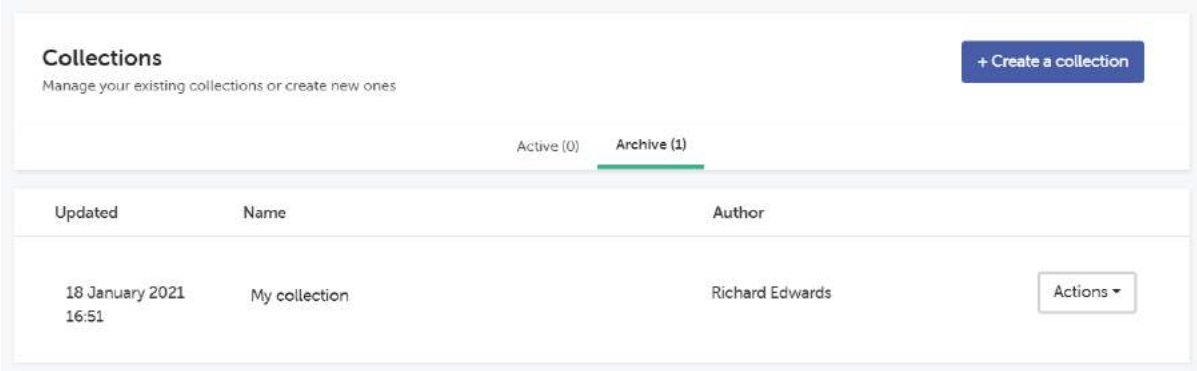
1. To edit a collection you first need to access the collections dashboard. You do this using the account drop-down at the top right of the screen. This is shown in the earlier images.
2. Active collections are shown in the Active tab. From here you can use the Actions menu to edit, archive or delete it. To edit, click edit and make any required changes the same way that you did when creating the collection. Once you have finished editing you save it the same way you did previously.



## Archiving a collection

Archiving a collection removes it from your active collection tab but doesn't delete it permanently. You can always unarchive a collection if you decide to in the future.

1. To archive a collection you first need to access the collections dashboard. You do this using the account drop-down at the top right of the screen. This is shown in the earlier images.
2. Active collections are shown in the active tab. From here you can use the Actions menu to edit, archive or delete it. This is shown in the earlier images. To Archive a collection, click Archive. You will be asked to confirm your decision. Click yes and your collection will be moved into the Archive tab.
3. You can unarchive a collection by navigating to the Archive tab and using the Actions menu to select Unarchive. You will be asked to confirm your decision. Click yes and your collection will be returned to the Active tab.




## Deleting a collection

Deleting a collection removes it completely from your account. Once deleted a collection cannot be recovered.

1. To delete a collection you first need to access the collections dashboard. You do this using the account drop-down at the top right of the screen. This is shown in the earlier images.
2. Active collections are shown in the Active tab, archived collections are shown in the Archive tab. You can delete a collection from either tab. From here you can use the Actions menu to edit, archive or delete it. To Delete a collection, click Delete. This is shown in the earlier images. You will be asked to confirm your decision. Click yes and your collection will be permanently deleted from the directory.

## Future plans

In the future - collections will be searchable on the Gateway. Because of this you will be able to make collections private or public depending on who you want to view your collections.

HDRUK
Health Data Research UK

# Why and how researchers and innovators make data access requests

The Gateway contains over 500 datasets (and growing) that are discoverable using the Gateway's easy-to-use interface. And a primary objective of the Gateway is to simplify, streamline and therefore reduce the time it takes users to submit, and data custodians to review and make a decision on user requests for access to data.

Now that your data is discoverable, innovators and researchers want to do more than just find it. Where it will support their work, they will want to request access to it. This is where the Data Access Request (DAR) functionality of the Gateway comes into play.

A DAR is the process by which a user can make contact with a data custodian to request access to their data. Their request is supported by additional information about the work they want to use the data to support.

## How applicants request access - Video tutorial: How to submit a data access request

Before applicants can request access, they first need to find a dataset that suits their research needs. They will also need a Gateway user account and must be logged in to the make an application.

- Applicants will see a Request access button if you are a custodian using the short enquiry form
- Applicants will see a *How to request access* button if you are a custodian using the new 5 safes application form.
- Applicants are encouraged to send a message to the custodian before starting the application process.

You can access a video showing how applicants can submit a data access request in the following link: https://discourse.healthdatagateway.org/t/how-to-request-access-to-a-dataset/456

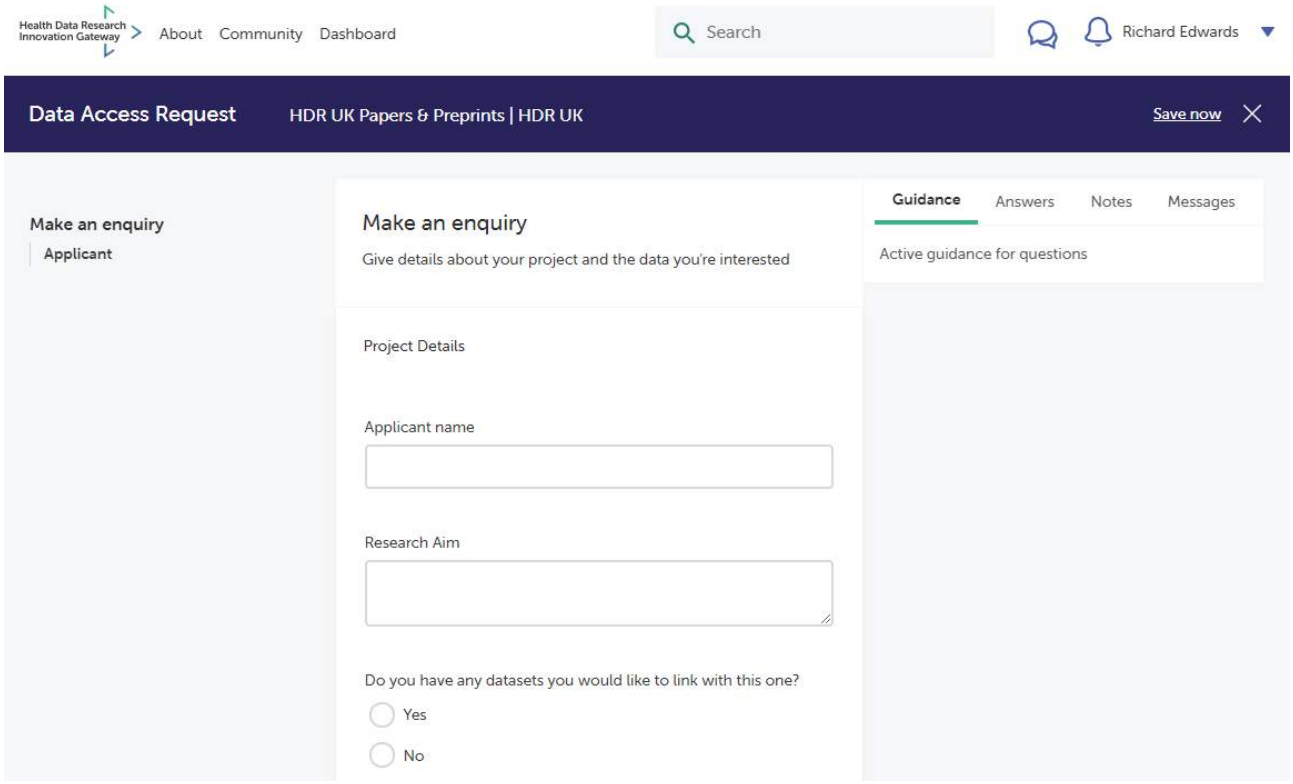
## Data Access Request (DAR) Forms

Depending on the custodian, users will either be able to send a data request enquiry to the custodian (Short Form Data Access Request) or, where a Data Custodian has implemented a 5 safes DAR Form on the Gateway (Long form Data Access Request), the entire DAR process can be conducted electronically through the Gateway itself. The 5 safes process offers other supportive functionality to both data requesters and custodians by enabling secure communications between requestor and data custodian in advance of submitting a DAR; powerful collaboration tools to allow collaborators to work together to complete the DAR, and the ability to monitor DAR progress.

More information on the 5 safes process is set out below and if you would like to put your Data Custodian forward for onboarding to 5 Safes or would just like more information, please make contact through Paola Quattroni at HDR UK.

## Short form Data Access Requests

Unless a Data Custodian has been onboarded to the 5 safes DAR process, all data requests will be handled through a simple enquiry form. An example of this is set out in the figure below.




The enquiry screen captures key information about the user's project and the data that they are interested in. It includes:

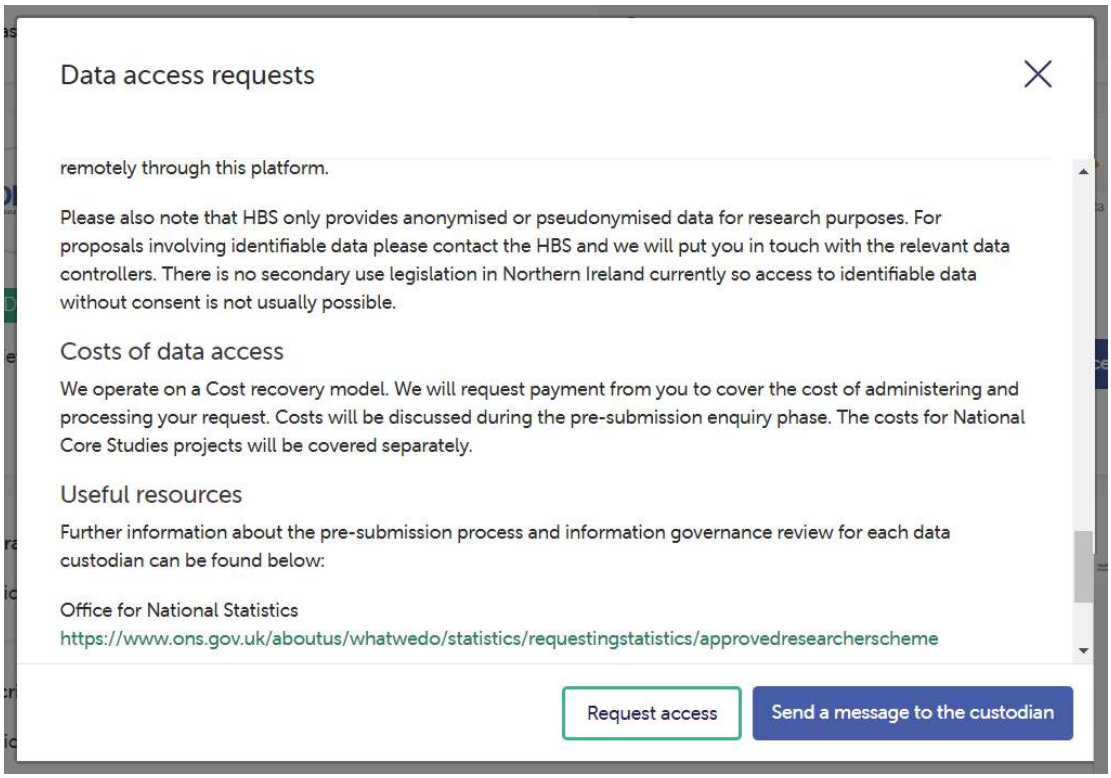
- Applicant name
- Research aim
- Whether there is a requirement to link datasets
- Which parts of the dataset the user is interested in
- Proposed project start date
- ICO number
- Research benefits (optional field)
- Ethical processing evidence (optional field)
- Contact telephone number (optional field)

Once the user has completed and submitted the enquiry form, the content of this will be emailed to the data custodian for consideration. For these types of enquiry there are no other steps that need to be undertaken on the Gateway.

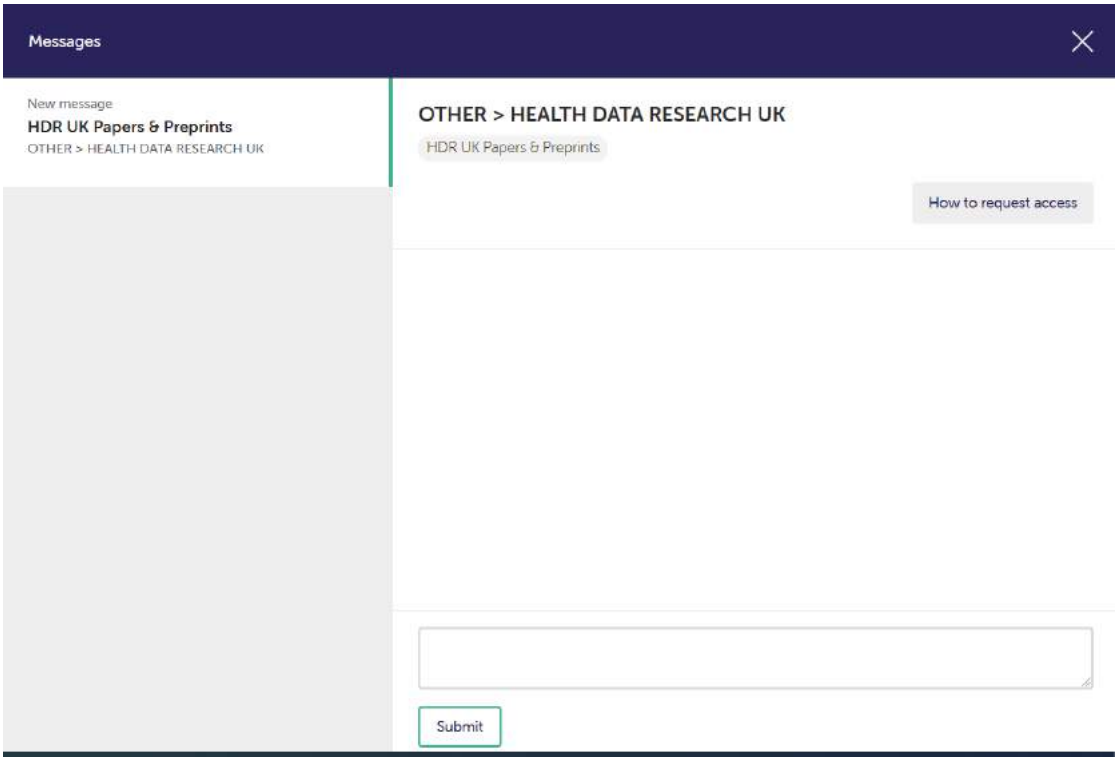
## Long form Data Access Requests

Where a Data Custodian has onboarded to the 5 safes DAR process, a much more advanced and fully featured set of DAR functionality is available to the custodian. In the majority of cases, this allows the

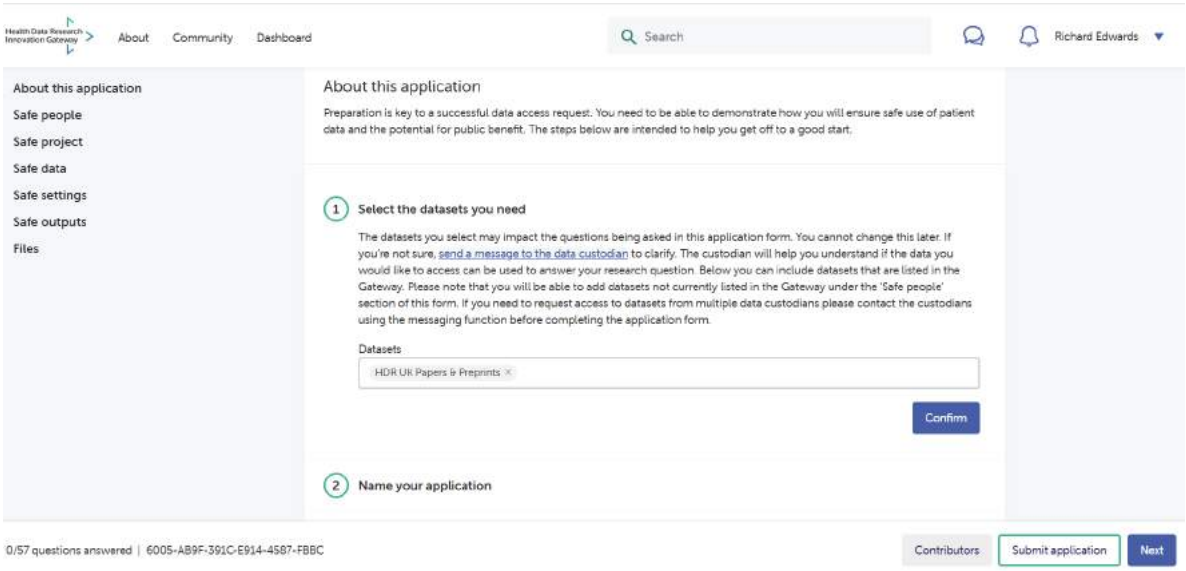
custodian to process the entire DAR electronically through the Gateway, helping avoid onerous and time consuming fragmented and/or paper based processes.




As illustrated in the images above and below, where a 5 safes DAR has been implemented, should they have any questions, users have the ability to message the custodian prior to submitting the application. This messaging function exists entirely within the Gateway application, is specific to each data request, and therefore is much simpler for custodians to administer compared to other tools such as email.



Once a user decides to submit a 5 safes DAR, they are required to select one or more datasets that they would like access to. Working with other contributors if relevant, the user(s) are then required to answer questions across each of the 5 domains of Safe people, Safe project, Safe data, Safe settings, and Safe outputs. This is illustrated in the image below.




Once the user(s) has completed and submitted the application the custodian is then able to review the application, query aspects of it and/or ask for further information and then decide on whether to approve the application (with or without conditions) or reject it. All of this can be done through the Gateway.

If you would like to put your Data Custodian forward for onboarding to 5 Safes or would just like more information, please make contact through Paola Quattroni.

# 5 safes DAR form authorisation workflow

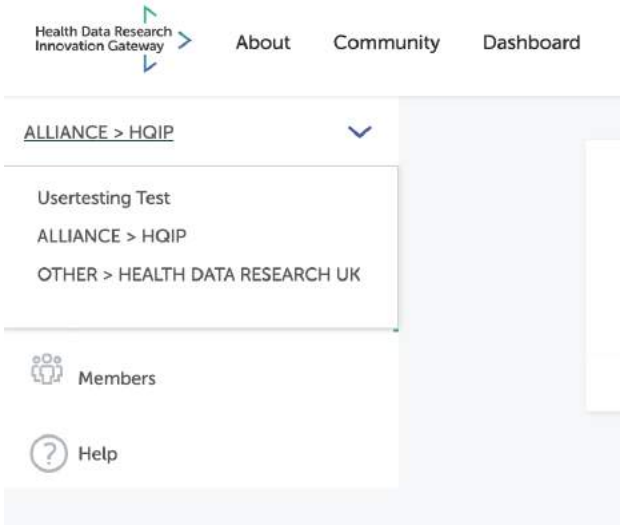
## How Data Custodians manage and process data access requests

### Creating a team and adding members – Video tutorial: Setting up a team, creating workflows and reviewing an application

A team on the Gateway is a group of associated members from a data custodian. Teams are used on the gateway by data custodians to manage datasets and data access request applications from users on the Gateway.

To create a new team, please get in touch with HDR UK via email. We are working on a new feature which will allow you to do this yourself in the future.

To add new members, go to your account, and switch between your name and your team's name on the top left hand corner.




Members can be added to your team, so long as they are members of the Gateway. There are two types of members who have different permissions.

- Managers can manage members, create and assign workflows, review applications that are assigned to them and make the final decision on data access request applications.
- Reviewers can only review applications that are assigned to them.

You can add team members via the Members section of your dashboard. To remove or change the role of a member, please contact HDR UK via email.

### Customise your "How to request access" information

This information will appear once the user clicks the How to Request Access button on one of your datasets. Each custodian can customise the text that appears for them. The same information will be shown for all datasets from the same custodian.

The modal will be displayed to all users at the beginning of their access journey. To ensure that they are prepared for the process, include all necessary information such as; what to do before they submit an application, when data can be released, the cost and other useful resources.

We are working on a new feature which will allow you to change this text yourself. For now, please get in touch with HDR UK via email to request any changes.

## Customise your Data Access Request form

You can customise the form that applicants are required to fill in when requesting access to your datasets.

- Select the questions you want on your form from our question bank. You can turn questions on/off, but not change the wording of a question.
- Add custom guidance to each question. Guidance can contain links to pages or documents.
- Add custom guidance on each page.

We are working on a new feature which will allow you to customise the form yourself. For now, please get in touch with HDR UK via email to request any changes.

If you need a question which is not on our question bank, get in touch. We are constantly reviewing the form and making improvements based on feedback form custodians and researchers.

## Creating a Data Access Request workflow - – Video tutorial: Setting up a team, creating workflows and reviewing an application

Workflows can help data custodians manage the review process after an application is received. You can assign people on your team to review certain parts of the form, and send automatic notifications when it's their turn to action.

Once a new application is submitted, managers will have the option to either make a decision (Approve, Reject) or assign a workflow. Workflows are optional, and won't do anything until they are assigned to an application.

To create a new workflow, go to Data Access Requests > Workflows > +Add a new workflow

- Create a phase and give it a name (i.e. Ethics). You can add as many phase as you like.
- Assign reviewers (members of your team) to the phase and sections of the form to be reviewed (safe people, safe project)
- Assign a deadline. Reviewers will receive a notification 3 days before the deadline.
- Phase 2 will start after phase 1 and so on.
- A phase will end when all reviewers have made a recommendation (issues found, no issues found).
- Managers of a team will always have the ability to manually skip the workflow review of an application to the next phase, in case something goes wrong or a reviewer isn't available.

You can have as many workflows as you like, and assign the appropriate workflow to an application after having reviewed its contents. For instance, you may have a Regular workflow and a Fast-track workflow, or a New application workflow vs a Extension workflow.

## Reviewing and managing a 5 safes Data Access Request application – Video tutorial: Setting up a team, creating workflows and reviewing an application

Managers will receive an email notification whenever a new application is submitted. Managers can then start the review process, and make a decision: approve, approve with conditions or reject. Managers can also assign a workflow, which will send the application to others in their team for review.