How to complete a Data Activity Risk-assessment Tool (DART) entry: RESEARCH BASED ACTIVITY

Overview

DART registrations are a living document that is used to both document the ways in which we process and hold all types of data and to risk assess how we manage it within the university.

You can return to it to update or correct after submission and even after sign-off should changes occur. If, after sign off, something about your project changes, please update the registration and let us know. Mandatory fields are marked within the platform with a * though we encourage you to complete as many fields as you can.

We generally and strongly recommend you differentiate research projects by ethics application. So, one ethics application should mean one project registration. There may be exceptions, such as where there are substantially different sub-studies or when there are separate ethics applications for the same project (such as when international projects require ethics applications in different countries) but this should be the general rule.

This walkthrough goes screen-by-screen and question-by-question to explain what we expect to see in each field for a DART research entry and assist you in its completion.

DART Support / Contacts

Primary contacts for DART within the university are as follows;

Faculty / Division	Name and e-mail
Faculty of Engineering	Inez McGregor / foedataprotection@imperial.ac.uk
Faculty of Medicine	Faculty of Medicine GDPR Coordination Team / fom.gdpr@imperial.ac.uk
Faculty of Natural Sciences	Debbie Oram / fonsdataprotection@imperial.ac.uk
Imperial Business School	Central Data Protection Team / data-protection@imperial.ac.uk
Support Services	Central Data Protection Team / data-protection@imperial.ac.uk

Further to the above, Imperial has Data Protection Coordinators who assist with their areas / divisions / teams and faculties when completing DART entries;

- Contact us / https://www.imperial.ac.uk/admin-services/governance/policies-and-guidance/contact-us/

Registering a new Data Activity

General guidance

- To navigate DART you can either;
  - Use the home screen button (top left hand corner), which will take you back to the beginning of DART.
  - Use the left / right arrows presented on the screens (left will take you back a screen, right will take you forward a screen). If DART does not allow you to move forward, please check all questions have been answered.
- This walkthrough is aimed at teams filling out a DART registration for research and so the expectation would be that you would choose research in this menu. Separate walkthroughs will be provided for non-research projects.
- Any questions / sections which include an asterisk (*) are mandatory meaning progression cannot occur to the next screen until answered.
- To save an incomplete section / screen, use the 'Save' option which can be found in the bottom right hand corner. Alternatively, your entry will be saved periodically when you move from screens and you will be notified when this occurs.
- You can request a PDF of your Activity Registration and associated Dataset(s) which will be sent to your inbox directly. It can be requested by clicking on the 'PDF' logo found on Screen 2 of your data activity risk-assessment screen.

Screen 1 – DART launch screen

1. Upon opening DART you will be presented 2 options which are directly linked to your profile;

   - Option 1 is listed as '**Register a new or view and edit existing data**'.

   - Option 2 is listed as '**Declare a new or view existing & centrally available datasets for your data activity risk-assessment**'.

2. For the purposes of registering a new activity, click on Option 1

Screen 2 – Your data activity risk-assessments

3   To register a new activity, click '**Create a new data activity risk-assessment'**.
4   Upon commencement of a new entry, a new entry will be add to this screen which will allow you to access any / all of the DART entries (which include you or have been created by you) next time you access DART.
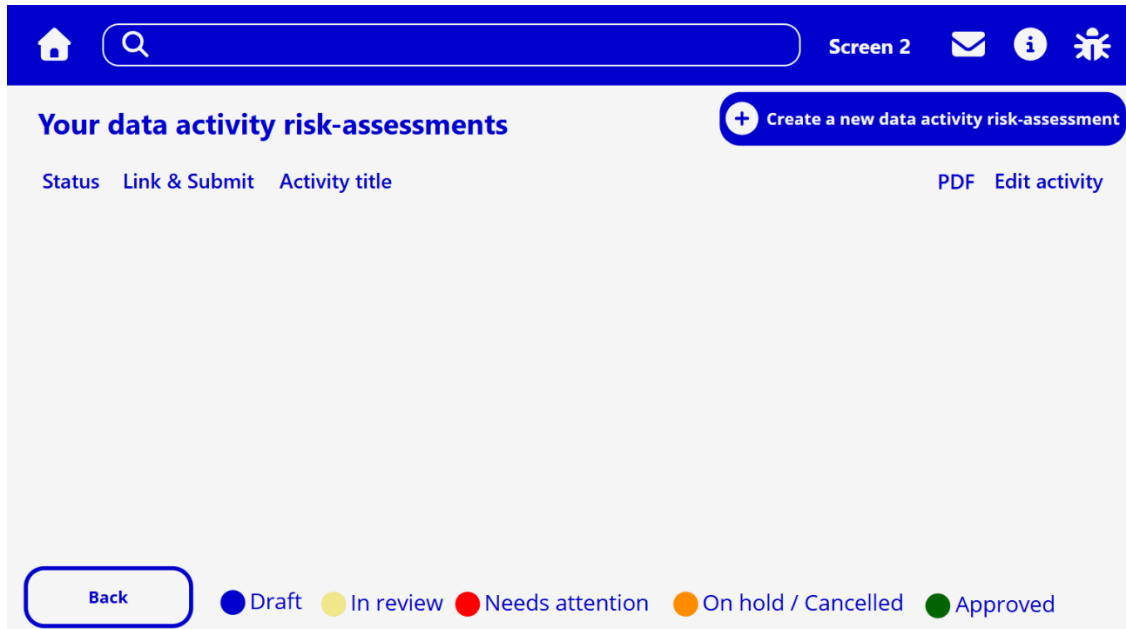
Creating a new Assessment

Screen 3 – Project / activity oversight screen

5   Upon starting a new activity entry, a unique reference number will be listed in the bottom left hand corner of the screen (example DA-000000)
6   Please provide;
 - **'A title for your project / activity'**
   - It's most helpful if the project name matches what you've used in the ethics application where appropriate.
   - Max. 300 characters
 - **'A precis / explanation / overview of your project / activity'**
   - Max. 10000 characters
 - **'Cite the Information Asset Owner (IAO)'**
   - The IAO is normally the PI or CI.
   - Upon clicking on the 'find people' button, if you start to write either their first name or surname a list of possible personnel will be shown. Once identified, click on the name so it becomes visible and highlighted in the box.
   - To change the listed name, click on the box again and utilise the (x) listed next to the name and replace.

Screen 4 – Type of activity

7   You will be presented 2 options;

 - Option 1 is listed as **'I am working on a research project'**

 - Option 2 is listed as **'My work is not related to a research project'**
8   For the purposes of logging a research activity, click Option 1, then press the right pointing arrow.

Screen 5 – identifying possible data subjects

9   Click the option that best represents this activity;
 - **This data could be used to identify:**
   - Living person/s
   - Dead person/s
   - Both living and dead person/s
   - None of the above
Whilst data protection legislation does not apply to the dead, we may still have reputational concerns or contractual responsibilities to adhere too. Furthemore, in the case of things like genomic data or family data, data collected from the dead can still tell us about the living (for example in the case of hereditary illnesses), which would make it personal data of others.

Screen 6 – identifying data types

10 On this screen we ask which specific data types / fields are included;
 - **'Please select all, if any, personal identifiers included in your data'**
11 Click all options that relate to the data involved in this activity
12 This screen lists a number of data types commonly collected as part of research projects, please click and highlight all that apply for this activity.
13 For any data types that are not listed, click on the '**Other**' option and then list them in the free text box which will appear.

Screen 7 – Risks

14 On this screen we ask what, if any, potential risks or limitations the data to be used within the activity represents to the university;
  - **'This data poses contractual or regulatory risk or could be used for malicious intent at risk to the university'**
15 Click all options that relate to the data involved in this activity. You should consider any limitations of use, whether the content poses any reputational, financial, physical risks, where the data has come from and / or whether any controls are in place or been agreed as part of the data collection.

Screen 8

16 When completing a research based activity, Screen 8 is skipped.

Screen 9 – Basic information

17 On this screen you will be asked to provide basic details for the activity. Please provide;

  - **'Lead Faculty'**
    - o This will help identify a suitable assessor
  - **'Department'**
    - o The department should normally be the department with which the researchers are affiliated.
  - **'Cost Centre'**
    - o This will normally be your section.
  - **'Information Asset Administrator (IAA)'**
    - o The Information Asset Administrator (IAA) is probably the person completing this form, but should in any case be the individual/s who is/are the first points of contact about the project or the data associated with the project. You can include multiple entries in this field if there are multiple contacts. Note that any names entered into this field will get access to the entry in DART.
  - **'Additional details for IAOs or IAAs'**
    - o This field will be most relevant when one of the IAOs is not affiliated with Imperial or does not have Imperial contact details.
  - **'What type of activity is this?'**
    - o For research this will already be populated, to be more nuance, please use the drop down which lists more granular options
  - **'What is the project / activity status?'**
    - o Active - Active projects are going ahead and are either collecting or processing data.
    - o Closed - Closed projects have been completed but not yet archived.
    - o To Be Transferred - Projects which are to be transferred are projects which will be transferred into or out of Imperial.
    - o Transferred - Transferred projects are ones which have been transferred to us from other institutions, such as when a PI or researcher changes affiliations.
    - o Not sure/Planning/In setup/On hold This is for projects which are still being planned, temporarily on hold, or might not even go forward.
  - **'Does this project/activity have defined dates?'**
    - o Although not all projects do, we generally expect research projects to have some sense as to when they'll begin and end.

Screen 10 – Systems and data outsourcing

18 On this screen you will be asked to provide details regarding the Systems to be used, Sharing and Outsourcing and Risk Mitigations;

Systems
- **'Does or will the activity involve the use of any systems new to Imperial? (such as survey platforms or cloud storage)'**
  - o If you are introducing a new service (free or paid for) to process data that isn't already in use within Imperial, choose yes.
- **'Is an unregistered system being developed for this project / activity by external developers?'**
  - o If you are hiring a team external to Imperial to develop an app or data collection system or to develop software for processing the data, please click yes.

Sharing and Outsourcing
- **Do you outsource or will you outsource data processing (e.g. use of cloud based apps)?**
  - o If you're paying or otherwise engaging the services of a third party to process any data (including, collection, storage, deletion, transfer, analysis, hosting, etc.) then choose Yes and select the options for whatever service the third party is providing. We will need to know if you're using any university-contracted cloud based tools like Qualtrics, Microsoft 365 or SharePoint, etc. but we won't require any additional contracts, etc. for that.
- **Does this project / activity have collaboration with other organisation(s), e.g. Universities, industry, partners etc., that you share or will share data with?**
  - o We need to know if you're sharing data with other organisations as part of the project. If you're thinking about as yet unknown future sharing for secondary research, although we will need to make sure that's mentioned in the PIS, we don't need to know about that here.
- **Is data transferred or going to be transferred outside of the UK?**
  - o Depending on what country you're transferring data to, we may need additional clauses added to data sharing or data processing agreements.

Risk mitigation
- **Will you be using new/emerging technologies for processing personal data (e.g. AI, data mining algorithms etc.?**
- **Please describe current mitigation(s) in place to protect personal data when using these new/emerging technologies**
  - o We're especially concerned about technologies (like AI, machine learning, novel kinds of genomic or -omic processing, etc.) that may increase the risk of unexpected identification, invasive processing and/or decisions being made about people through automated means. Novel technologies may also involve unknown risks associated with them that require consideration and mitigation.
- **If using AI, please give details**
  - o As part of implementation of AI, please ensure you have reviewed the university AI Code of Practice and any additional guidance set regarding how AI can / cannot be used / https://www.imperial.ac.uk/media/imperial-college/administration-and-support-services/legal-services-office/public/data-protection/DPA-CoP-08---Use-of-Artificial-Intelligence-(AI).pdf

Screen 11 – Transparency information

19 On this screen you will be asked to provide details to cover transparency requirements and details;

- **Will you be attempting to anonymise line level data in this activity?**
  - o We will need you to explain your approaches to anonymisation by filling out an anonymisation form /

https://imperiallondon.sharepoint.com/:w:/r/sites/cf/DPIA/_layouts/15/Doc.aspx?sourcedoc=%7B0CEF1415-CBD8-4A60-BEA6-EB91D7021841%7D&file=Template%20-%20Anonymisation%20Form.docx&action=default&mobileredirect=true&CID=2D202C31-FB8A-4DBF-9D3E-D8F1912FC85C if you believe this to be the case. If you have a pseudonymised dataset in mind, choose no. Please note, Pseudonymisation is not anonymisation and 'pseudo-anonymised' is not a recognised term under legislation. We are also not asking about anonymisation of data for publication such as aggregate tables or charts. We expect researchers to take standard statistical disclosure controls for those.

- **Please select Legal Basis (under data protection legislation) for processing of data**
    - o For research projects in the Faculty of Medicine, your legal bases will almost without exception be those in the menu beginning with 6(1)(e) and 9(2)(j) and none of the others. Please select only those two. If you want more information as to why this is the case, please consult the FoM Code of Practice / LINK, for health research data in the first instance
- **Data on roughly how many individuals will be processed as a result of these activities?**
    - o Please let us know how many data subjects you're processing or collecting data about.
- **Do you have a privacy notice for the project / activity?**
    - o If you've completed a SETREC/ICREC PIS template, then you should have completed a transparency / privacy notice already, and you should choose "Yes – local but based on university templates." If you've taken a different route through ethics or if your project does not require ethics, you will still most likely need to create a privacy notice. If that is the case, please contact us to confirm and so that we can provide an updated template.
- **Who authorised the privacy notice?**
    - o If you're using the SETREC/ICREC template, please put 'RGIT' (Research governance and integrity) / LINK.

Will this activity involve any of the following?
- **Evaluating, scoring, profiling or predicting from data**
    - o This is likely to apply to research projects. Examples of common mitigations include ethics assessment, transparency, pseudonymisation…
- **Will the data be used in automated decision making about individuals with legal or similar effects?**
    - o This is unlikely to apply to research projects, but let us know if you think it might.
- **Will the data be used to contact individuals in ways which they may find intrusive?**
    - o If you are cold-contacting participants or contacting participants from a list (say, one from a participant database or a list of participants associated with a different project) please check "Yes". In the "Please describe current mitigations in place to protect personal data", please describe where the relevant contact details came from and why you believe we are permitted to recruit from that list. Please also describe any mitigations in place to make sure that we're using a minimum necessary data to contact individuals.
- **Will the data be used to make decisions or take actions against individuals in ways which can have a significant impact on them?**
    - o Impact in this case means anything that the individual might find harmful. Most research will not need to fill in details here but an example of a study that has the potential to impact an individual would be an interventional clinical trial.
- **Will the data be used for systematic monitoring or observing of individuals where they may not be aware of it?**
    - o This would be relevant where a project is using CCTV or mass collecting personal data without consent.
- **Does this project / activity require processing of NHS patient data**

- o Projects which involve processing NHS patient sometimes require additional technical protections and may require processing in a DSPT-or-equivalent certified environment. Any patient data classified as medical data would also require a mechanism to satisfy the duty of confidentiality.
- **Does this project / activity require processing Children and/or Vulnerable persons data?**
  - o We need to take particular care when processing data from children or vulnerable persons since the requirement for transparency may be more difficult to satisfy.
- **Special Category Data (this includes health data)**
  - o This includes health data very broadly construed, ethnicity data, any kind of genetic data, biometric data (including things like voice recordings or images of the face), and information about sex lives or sexual orientation.
- **Do you hold consent forms for the activity?**
  - o Consent forms contain people's names so are personal data and consideration needs to be given to how to protect them.
- **If relying on consent, how will you ensure that consent is correctly informed and with no power imbalance?**
  - o Please describe how and when data subjects are given consent forms and any actions taken to ensure people are freely giving their consent and understand what they are signing up too.
- **Do you use individuals' contact details to tell them about communication, including research, public engagement, outreach, etc.**
  This may qualify as marketing
  - o Communications conducted for the purpose of the research project they have consented to normally doesn't qualify, though communications informing potential subjects about other or future studies they might like to join may be deemed marketing and require specific and separate consent.
- **Is the project run as a clinical study?**
  - o This is relevant because clinical studies sometimes have different requirements for data storage or collection than non-clinical studies. Some systems can't be used for storage of clinical data, etc.
- **Does your project use human tissue or organs?**
  - o It's helpful for us to know this so that we can understand what data is likely to be generated from any samples being used.
- **Do you have approval from CAG under section 251?**
  - o If you are processing patient data without their consent to access their medical data, you are likely to need approval from the Confidentiality Approval Group. If you are not consenting patients and don't believe CAG is required, we will likely need to know that you've at least confirmed that CAG isn't required.
- **Do you have HRA approval for the project?**
  - o HRA Approval is a comprehensive process that ensures and supports ethical, legal, and governance compliance for research involving NHS patients, staff, or data within England.

Screen 12 – Training details and Ethics

20 On this screen you will be asked to provide details to covering Training and Ethics;

- **Please provide details of any data protection training that you or your team may have undertaken outside the mandated Imperial Essentials course**
  - o This helps us understand if any additional or focused training has been undertaken that will assist in supporting our approaches to how data will be processed under the study.
- **Is this project primary or secondary research?**

- o If you are processing data that was originally collected for a different project or for another purpose, for the purpose of this form you're conducting secondary research. Otherwise, please choose primary or Both as appropriate.
- **Do you have ethics approval for the data processing activity**
  - o We will need to know the project is ethically approved or, if it hasn't been, why it hasn't been. This is because the legal justification we have to process people's data normally depends on the research we're doing being ethically approved.
- **Edge ID/Worktribe ethics number/ other ethics ID (i.e. IRAS, NHS REC number) if unknown/ not applicable**
  - o This is helpful if we need to link the project in other systems. Including this may shorten processing times because it may let us find needed information about the project without contacting you.
- **Activity ethics sponsor**
  - o We need to know which organisation is acting as the ethics sponsor. If it's not Imperial, another field will appear in which we ask who the sponsor is.
- **From which type of funding is the project funded?**
  - o This relates to the funding code letter.
- **Please provide the codes used for funding (e.g. P12345, G12345)**

Screen 13 – 20

21 When completing a research based activity, when you move on to the next screen, Screens 13 - 20 are initially skipped.

Screen 21 – Document upload

22 On this screen you will need to upload supporting documentation for your project / activity
- **Data flow diagram**
  - o This describes / shows the data lifecycle for the activity. A template is available
- **Document upload**

  Individual documents can be uploaded and each time this occurs a new document upload number will appear below it allowing multiple documents to be stored. We generally need the following documents and draft versions are fine whilst being created /finalised:
  - o Participant Information Sheet (PIS) – This is a document presented to participants that tells them about the project. You should have been provided a template for this. If it's a SETREC or ICREC template you should have a transparency notice in the second half which tells subjects how their data will be used. A transparency notice will be required with very few exceptions. If you're using personal data to conduct secondary research or otherwise have no interaction with data subjects, we will often require you to create a privacy notice and to publish it somewhere data subjects would reasonably expect to find it should they believe themselves to be among those whose data is processed as part of this project. DART review of this prior to submission to ethics will help reduce any requirement for ethics amendments.
  - o Consent form – You should have been provided a template for this. The DART assessment is usually focussed on what the consent form tells data subjects about how their data is being used.
  - o Protocol – For primary research you should have been provided a template for this. Where available, we will need to see this because it helps us better understand what your project involves, what data you're collecting and how.
  - o Ethics Application form – For us, this serves much the same purpose as the protocol.
  - o Agreements / contracts – If data is moving between us and other organisations as part of this project, we will usually require either a data processing agreement or

data sharing agreement. If you have those (including in draft form), please attach those.

- o Anonymisation form – If you're claiming that any of the datasets relating to people that you're collecting/ sharing / generating is anonymous, we will need to verify and document that. To accomplish that, we have an anonymisation form that you should complete and attach to the registration. This asks you to explain any anonymisation procedures undertaken on a dataset and / or your reasoning as to why you believe a dataset associated with your project is anonymous. Note, we only need these for datasets which you're claiming are anonymous – not pseudonymous.
- o Data dictionary – This is a list of fields being collected as part of a dataset. This is helpful to have for all projects and is especially useful if you're claiming a dataset is anonymous. Please include separate datasets on separate worksheets. We have a data dictionary template available, but do not require a specific format so long as you can provide us with a comprehensive list of the fields that are included as part of the dataset.

Should documents be replaced, amended, finalised and require removal from DART, this can be managed via the screen by identifying the relevant document and clicking the [X] listed next to the document title / name.

Screen 13 – Activity registered / data set creation

23 Once you have uploaded your supporting documentation, which can be amended / added to at any time, you will be taken to Screen 13 which confirms completion of your initial DART entry and from here you can either create a NEW dataset(s) or, if a dataset(s) has already been created, then you have an option to link and submit your completed DART entry

My Data activity Risk-assessments
- Click on this option to return to Screen 2 and review your activity list.

Declare or view existing datasets
- Click on this option to view all datasets available to you and/or create a new dataset

Link & Submit to this Risk-assessment
- Click on this option if the necessary dataset has already been created and you are therefore ready to link the data set and the activity registration for submission and review.
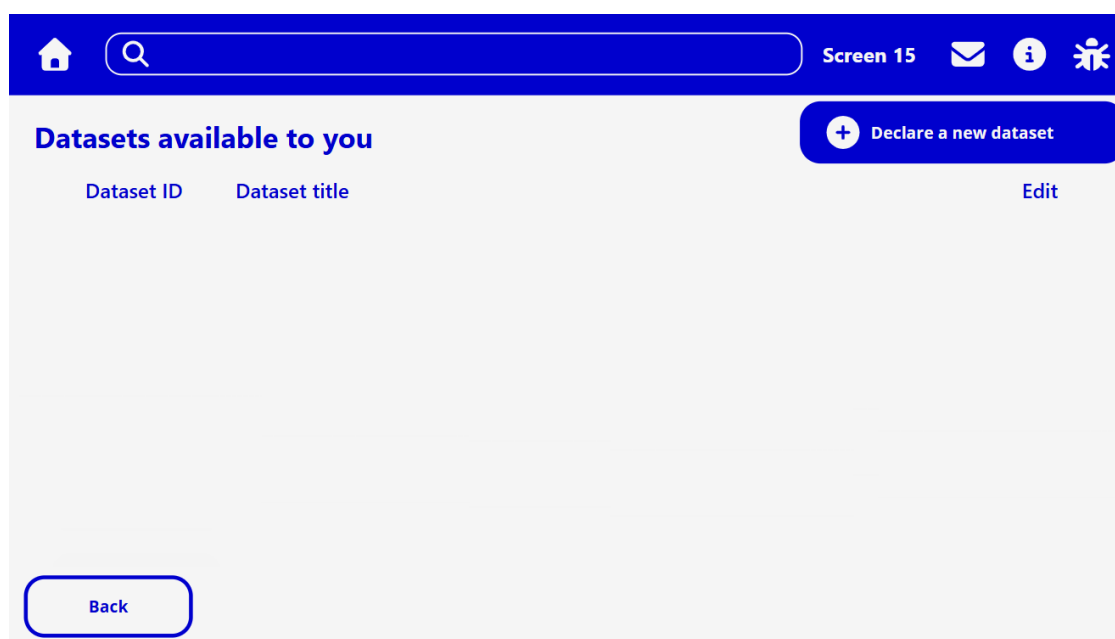
Declaring a New Dataset

Screen 13

24 For the purposes of logging a new dataset, click Option 2 'Declare or view existing datasets'.

Screen 15

25 On this screen you will see a list of;
- all 'public' datasets that are able to be viewed only;
- all datasets where you are named or have authored
- an option to 'Declare a new dataset' in the top right hand corner

26 For the purposes of logging a new dataset, click the button in the to right hand corner +Declare a new data set'

Screen 3 – Learning about your dataset

27 Upon starting a new dataset entry, a unique reference number will be listed in the bottom left hand corner of the screen (example DS-000000)

28 Please provide;
- **'Please enter title of your dataset'**
    - o If the dataset has an official name, it's fine to use that, but in general it's most helpful if the name briefly distinguishes the dataset from others in some way.
    - o Max 300 characters
- **'Please enter description of this dataset providing details regarding the data processing involved'**
    - o This field should include a general, high-level description of the dataset. More specific questions about storage and security will be addressed elsewhere.
    - o Max 10000 characters
- **'Information Asset Owner (IAO)'**
    - o The IAO is normally the PI or CI.
    - o Upon clicking on the 'find people' button, if you start to write either their first name or surname a list of possible personnel will be shown. Once identified, click on the name so it becomes visible and highlighted in the box.

o   To change the listed name, click on the box again and utilise the (x) listed next to the name and replace.

Screen 4 – Type of activity

29 You will be presented 2 options;

- Option 1 is listed as **'I am working on a research project'**

- Option 2 is listed as **'My work is not related to a research project'**

30 For the purposes of logging a research activity, click Option 1, then press the right pointing arrow.

Screen 5 – identifying possible data subjects

31 Click the option that best represents this activity;
- **This data could be used to identify:**
  o   Living person/s
  o   Dead person/s
  o   Both living and dead person/s
  o   None of the above

Whilst data protection legislation does not apply to the dead, we may still have reputational concerns or contractual responsibilities to adhere too. Furthemore, in the case of things like genomic data or family data, data collected from the dead can still tell us about the living (for example in the case of hereditary illnesses), which would make it personal data of others.

Screen 6 – identifying data types

32 On this screen we ask which specific data types / fields are included;
- **'Please select all, if any, personal identifiers included in your data'**

33 Click all options that relate to the data involved in this activity

34 This screen lists a number of data types commonly collected as part of research projects, please click and highlight all that apply for this activity.

35 For any data types that are not listed, click on the '**Other**' option and then list them in the free text box which will appear.

Screen 7 – Risks

36 On this screen we ask what, if any, potential risks or limitations the data to be used within the activity represents to the university;
- **'This data poses contractual or regulatory risk or could be used for malicious intent at risk to the university'**

37 Click all options that relate to the data involved in this activity. You should consider any limitations of use, whether the content poses any reputational, financial, physical risks, where the data has come from and / or whether any controls are in place or been agreed as part of the data collection.

Screen 17 – About the dataset

38 On this screen we ask about the provenance of the dataset, associated risks and the systems / platforms / tools which are intended to be used / are being used for the purposes of processing the dataset.
- **Should this DART dataset registration be visible to all DART users?**

- o This option is turn off by default meaning if unchanged, the dataset **will not** become visible to other users of DART – including those not named in the activity nor dataset respectively. For most contained activities, limiting access is expected though for large / university wide activities or those which will be replicated across the organisation it may be prudent to make the dataset visible and allow people to link / utilise directly
- **Dataset Owning Department**
  - o The department should normally be the department with which the researchers are affiliated. There is a drop down which allows for specific teams / divisions / areas to be identified.
- **Please indicate if any person is identifiable in any part of your dataset**
  - o Pseudonymous datasets are datasets which are not identifiable on their own but would be if combined with other available data.
  - o Identifiable data would be for a dataset which includes direct identifiers (name, email address, IP addresses, phone numbers, ID numbers, etc.)
  - o Anonymous datasets are those whose data subjects cannot be identified even when combined with other available datasets
- **Impact of data release on the university or an individual**
  - o Low - This data would be unlikely to have a significant impact on data subjects. The last time a data subject had a cold would be a good example of low impact data.
  - o Medium – a risk set between low and high.
  - o High - This data would, if breached, have a significant impact on data subjects. Mental health data and HIV status are two good examples of high impact data.
- **Data Risk**
  - o Click all options that relate to the data involved in this activity. You should consider any limitations of use, whether the content poses any reputational, financial, physical risks, where the data has come from and / or whether any controls are in place or been agreed as part of the data collection.
- **Dataset sensitivity classification**
  You will be guided through the assessment of the data category by the assessor after submission. In line with university definitions This cannot be answered by the author of the Dataset.
  - o *Certified Environment Required:*
    Release of this data would have significant legal and reputational impact on the university. It may be significantly impactful on a large number of individuals.
  - o *Confidential Data:*
    Release of this data would have a high impact on either the university or the individual or it has significant legal restrictions
  - o *Restricted Data :*
    Release of this data would have a medium impact on either the university or an individual. It is not expected to have significant legal restrictions.
  - o *Unrestricted Data:*
    Release of this data would have a low impact on the university and it could not be tracked back to an individual. It will not have significant legal restrictions.
  - o *Public:*
    There are no impacts from release of this data to the public domain.
  - o *Non-sensitive organisational data*
  - o This is data pertaining to Imperial which may or may not be published by default, but may be disclosed via freedom of information requests subject to legal advice.
- **Please select origins of this dataset**
  - o Generated as part of university activity – Select this if the data is generated as part of other ordinary university activities, for example if you're using HR data or student data.
  - o NHS Trust – ICHT – Choose this if this data is being shared with us by Imperial College Healthcare trust (ICHT).

- o NHS Trust – Other - Choose this if this data is being shared with us by a trust which isn't ICHT.
  - o NHS Digital – Choose this if the data is coming from NHS Digital.
  - o ONS – ONS stands for the Office of National Statistics.
  - o CPRD – For where data comes from the Clinical Practice Research Datalink (CPRD)
  - o Collected as part of trial / volunteer participation – Choose this if the source of this data is directly from data subjects recruited as part of a trial or research project.
  - o Acquired from another 3rd party – If we're receiving the data from another party but none of the other choices apply, choose this one.
  - o Publicly available data (i.e. accessible without a password or similar protection) -
  - o Other – If none of the above apply choose this and explain in the additional information question.
  - o N/A – highly unlikely, if unknown then this will need to be investigated.
- **Please indicate any systems within which data will be processed (select all that have been used. or 'None of the above)'**
  This is a long, searchable list which includes almost all systems in use by Imperial. If you can't find it at first search, try alternative versions of the name - for example, most Office 365 software is included under "Microsoft 365" rather than something like "Excel Online". If you are using a system that's not approved for use within Imperial or just one that isn't in the list, please list it and a contact for it here. This is likely to be the case if you're using a collaborator's system
- **Dataset Status**
  - o Active - Active datasets are ones that are being collected or processed. If you are collecting
  - o Archived - The dataset has been placed in long-term storage for the rest of the retention period.
  - o Destroyed - This dataset has reached the end of its retention period and been destroyed. This may involve destruction of physical storage, overwriting storage locations, etc. depending on current Imperial or collaborator guidelines
  - o To be transferred - The data will be transferred between us and another institution in either direction.
  - o Transferred - The data has been transferred between us and another institution in either direction.
  - o Other
- **Information Asset Administrator (IAA)**
  - o the IAA is the staff member or staff members (you can choose several) who will be the first point of contact for questions about the dataset.
- **Location / country of origin**
  - o This is a list of all countries. Please select as many as apply.
- **Access controls**
  - o If access is controlled by an Imperial distribution lists please select them
- **Data controllership**
  - o Imperial College London – If we're the sponsor, we are almost certainly the controller.
  - o Another Organisation – If another organisation is hiring us to process this dataset, they would be the controller.
  - o Joint controller including Imperial College London – If we're one of two or more organisations collaborating on a project, and more than one organisation has significant input into the purpose and means of processing then they may be joint controllers.
  - o Don't Know – to be identified
- **Retention**
  - o For most research projects, the retention period is 10 / 15 years, but there are exceptions, especially for clinical trials.

We have a general obligation to destroy data after we've satisfied the purpose for which it was collected. For research, that purpose includes a period of archiving for regulatory reasons and so that research can be checked later on. Storing data beyond appropriate retention periods is a data breach. This means that with very few exceptions datasets should have a retention period recorded. If this is research data that is going to be stored for more than 10 / 15 years or has an indefinite retention period, we will need to know and record why.

- **Additional mitigations**
  o It may be useful to highlight any additional system protections you have put in place such as encryption, access controls, addition security measures.

Screen 18 – Data Questions

39 On this page, we ask questions about the dataset including specific data subject types, volume of records and data points are included and how many data subjects are included.
- **Data Subject Type**
  o Advisers, consultants, other professional experts – These are people who are working for Imperial but who aren't staff.
  o Alumni – Former Imperial students
  o Customers, clients – Choose this in case this dataset includes the personal data of people we're providing a service to or the employees of organisations we're providing services to.
  o External partners / associates – If this dataset includes the personal data of people who work for collaborating institutions, choose this one.
  o Former staff – Former Imperial staff
  o Members of the public – This one is used where you're interacting with members of the public. Examples would be things like Imperial-run events and public outreach.
  o Not personal data – If the data isn't about people or the FoM GDPR team have already confirmed this dataset is anonymous, this one may be applicable.
  o Patients – If this data is being collected as part of care being provided, then this may be relevant.
  o Records relating to those under 13 years old
  o Records relating to vulnerable individuals – If the people whose data is being collected could be classed as vulnerable, please choose this one.
  o Service users
  o Staff (includes volunteers, agents, temp and casual) – Current Imperial staff
  o Students – Current Imperial students
  o Study Participants – If the personal data being collected or received is to be used in a study, then choose this one in addition to those.
- **How many data subjects does the dataset concern?**
  o Please select the option which corresponds to the number of data subjects in this specific dataset.
- **Does the dataset contain any of the below (select as many as apply)**
  o This drop down menu lists a number of data types commonly collected as part of research projects. Many of these are special category data types. Please choose all that apply.
40 Upon completion of the above questions, you will have completed the dataset entry where you will be presented the following message and 2 options;
- 'You have reached the end of the dataset declaration. You can return to this declaration as many times as you need.'
  Clicking one of the buttons below is not final and doesn't send anything for review.
- Option 1 is listed as **'Datasets available to me'**

- o Clicking this option will return you to Screen 15, which allows you to either create a new dataset, view/edit previously completed datasets – including the most recent and view public / visible datasets entries.
- Option 2 is listed as '**My data activity risk-assessments'**
  - o Clicking this option will return you to Screen 2, which allows you to either create a new activity entry and view/edit previously completed datasets.
41 For the purposes of linking and submitting a previously completed Activity entry and Dataset entry, please click option 2 and enter the activity you wish to link a dataset too.
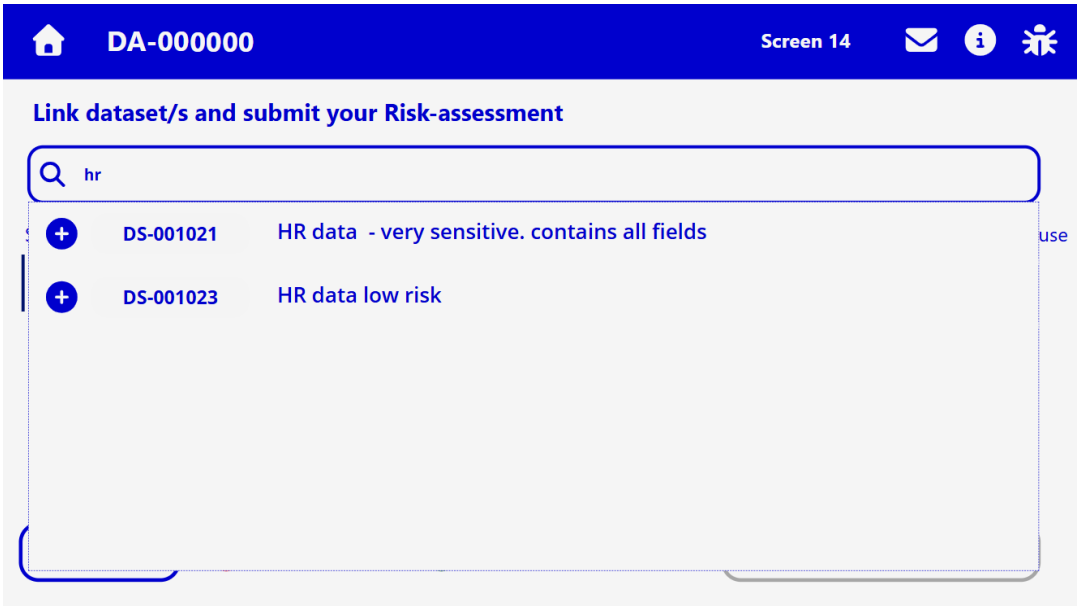
Link Data and Submit

42 Upon opening the relevant Activity entry you want to link to a Dataset too, progress to Screen 13, either by using the arrows on the right hand side of the screen or directly opening the Activity entry

43 On Screen 13 you are presented 3 options;
- My Data activity Risk-assessments
    o Click on this option to return to Screen 2 and review your activity list.
- Declare or view existing datasets
    o Click on this option to view all datasets available to you and/or create a new dataset
- Link & Submit to this Risk-assessment
    o Click on this option if the necessary dataset has already been created and you are therefore ready to link the data set and the activity registration for submission and review
- For the purpose of linking an Activity and a Dataset, please click option 3 – '**Link & Submit to this Risk-assessment'**.

Screen 14

44 On this screen you will be presented a search bar where upon clicking you can search for the name of the data set you wish to link to the activity. You can search for either the unique Dataset number (DS-000000) or the name of the Dataset (including partial spellings, individual words in the title etc.)

45 Once you have identified a dataset you want to link, please use the (+) visible to the left hand side of the name

46 You are able to link as many Datasets as required, including those you have previously created / been involved in the development of as well as any datasets that have been made public / visible by others.

Searching for datasets

Linked datasets

Status	Linked datasets	Unlink	Declare use

Link dataset/s and submit your Risk-assessment

🔍 Your linked datasets will appear here. If you know the name of the dataset you need - use the search bar to find it and lin

🔴 **DS-001021 - HR data - very sensitive. contains all fields** ✖ ›

🔴 DS-001023 - HR data low risk ✖ ›

Back   🔴 Use not declared 🟢 Use declared   Declare use to submit

47 To remove / delink a dataset from an activity please use the (X) listed next to the relevant title.

48 To begin the process, pick a dataset you want to attach/declare and press the 'Declare Use' arrow.

49 Please note, the option to 'Submit my declaration for approval' will remain unavailable until all datasets are declared.

Declaration of dataset usage

How are you using the data

50 On this screen you will be presented a series of questions which address particular risks regarding how the data will be kept secure throughout its data lifecycle using either drop down answers or free text boxes. For the purposes of completing your DART you will need to repeat this process for each dataset you want to attach/declare relating to your activity;

About
- **Dataset usage department**
    - o Please name the department which is utilising the dataset
- **Details of mitigations to risk relating to a specific usage of a dataset (additional to any mitigations of the dataset risks themselves)**
    - o If there is anything missed or you would like to add in addition to your answers then do so here.
- **Please describe any anonymisation methods you have implemented on the source dataset to further reduce identifiability for this use (E.g. for publication)**
    - o There are some specific methods which can be applied to some types of data to reduce identifiability. In the case of image data, for example, facial features may be blurred. If you are applying any of these techniques, please describe them here and see the Anonymisation Form which will need to be completed and uploaded in the Document Upload section.

Location
- **Where do you store this dataset? (Please tick all applicable options)**
    - o Please choose all that apply.
    - o a. ICT (Imperial) Managed Server (e.g. ICL database, H: drive, hosted at ICL etc.) - This would include the Group space, Shared Drives, Secure Enclaves, BDAU or other systems that are owned by Imperial and managed by ICT.
    - o b. Non-ICT (Imperial) Managed serve in Imperial building – Some systems which are owned by Imperial are managed by teams outside of ICT such as by research groups
    - o c. Outside Imperial server – If you're storing data on ICHT systems, Qualtrics, OpenClinica or another system which uses servers managed by a third party, choose this.
    - o d. Imperial 365 (e.g. SharePoint, email OneDrive) – This refers to all Microsoft cloud systems. So, this would include things like Excel files stored in OneDrive or SharePoint but not Excel files stored on the shared drives or on your PC.
    - o e. other cloud (e.g. Dropbox) – Examples would include Google Drive, Dropbox, or non-Imperial 365 instances. We generally recommend against using these systems.
    - o f. PC Imperial – If data is being stored on an Imperial computer, please choose this.
    - o g. PC (non-Imperial) – If data is being stored on a non-Imperial computer (such as your personal computer) choose this.
    - o h. Portable – If data is stored on devices such as USB Keys, memory cards, portable storage devices, DVD's etc.
    - o i. Imperial Archives and Corporate Records Unit (ACRU)
    - o j. Other
- **In which countries is the dataset stored? (Please tick all applicable options)**
    - o Please choose all countries in which this is stored.
- **Please provide any additional information about locations and storage arrangement for processing of this dataset**
    - o In this field, please detail any other relevant information about the systems used. It may be useful to highlight any additional system protections you have put in place such as encryption.

- **If you store any aspect of this dataset in a non-electronic format (e.g. consent forms) please select relevant locations/formats from the drop down below**
  - o This is most relevant for paper CRFs or consent forms.

Systems
- **Please indicate any systems within which data will be processed (select all that have been used. or 'None of the above')**
  - o This is a long, searchable list which includes almost all systems in use by Imperial. If you can't find it at first search, try alternative versions of the name - for example, most Office 365 software is included under "Microsoft 365" rather than something like "Excel Online".
- **If you are using any other system(s) not already registered in the list above, please provide a name and contact email for the system owner of each unregistered system**
- **For all unregistered systems, please also provide the system names here, with details of what they do and how you are using them.**
  - o By answering these 2 questions will allow the team the ability to investigate any checks that may have been undertaken and/or ensure relevant additional support is engaged to review the propose system(s) to be used.

Access
- **Who controls access to this dataset**
  - o Imperial Staff – This would be anyone with an Imperial staff contract.
  - o Imperial Honorary Staff – This would be staff who have honorary contracts with Imperial and so are bound by our policies.
  - o Non-Imperial Individuals – This is most likely to be the case for datasets that are being shared or whose processing is being outsourced.
  - o Students – Imperial students. If students from another institution, please choose non-Imperial individuals.
  - o Visiting Researchers – Visiting researchers from other institutions.
  - o Other – Please clarify in one of the text boxes below this question.
- **Who has access**
  - o Information Asset Owner – This is probably the PI/CI.
  - o Information Asset Administrator – This is probably you.
  - o Delegated by IAO but outside Imperial – There will probably need to be some agreement in place which governs this person's role.
  - o System manager – This will normally be someone with a role in technically managing the system.
  - o Other – Please explain in one of the text boxes below.

Data Transfers between Organisations
- **Is this specific dataset shared with third party controllers outside of Imperial?**
  - o If you are sharing data with a collaborator or if you have specific plans to place this dataset in an external repository, please choose "Yes"
- **Are there data sharing agreements in place for this specific dataset?**
  - o Data sharing agreements will be required whenever personal data is being shared between Imperial and another party. We recommend you contact the research contracts team to begin getting these in place as soon as possible.
- **Will data be transferred outside of the UK at any time, for any purpose?**
  - o Special considerations may apply where data is leaving the UK.
- **Do you outsource processing for this specific dataset?**
  - o If you are engaging parties external to Imperial to undertake some processing as part of this project then choose "Yes".
- **Please tick all applicable common mitigations that are in place to protect this dataset?**
  - o If you're unsure what any of these mean, don't choose them.
    - ▪ **Dedicated secure environment / toolkit** – This would normally refer to an environment with DSPT certification. NHS England maintains a register

of systems which have this in place. If you aren't sure but believe the system you're storing this dataset in has DSPT, you can search there.

- **ISO27001 certified environment** – This refers to a specific ISO security certification. If an environment has this, it will likely be publicised or you will be able to confirm this with the system manager.
- **ISO 27017 certified environment** - This refers to a specific ISO security certification for cloud service providers. If an environment has this, it will likely be publicised or you will be able to confirm this with the system manager.
- **ISO27018 certified environment** - This refers to a specific ISO privacy certification for cloud computing services. If an environment has this, it will likely be publicised or you will be able to confirm this with the system manager.
- **ISO9001 certified environment** - This refers to a specific ISO security certification. If an environment has this, it will likely be publicised or you will be able to confirm this with the system manager.
- **Cyber Essentials** – This refers to a UK government cyber security certification. Again, this will either be publicised or you will be able to confirm this with the system manager.
- **SOPs for data management, data access, etc**. - If there are written standard operating procedures for the handling of data associated with this dataset, choose this. If you don't have access to these SOPs, don't choose this.
- **Encryption** – Choose this if this dataset stored on an encrypted system, in an encrypted location.
- **Physical isolation** – This may be physical isolation of paper files, of physical hard drives or of a server. This could include storing them in a secure facility, in a locked office, or in a locked filing cabinet.
- **Robust pseudonymisation** – If the data has been stripped not just of identifiers but also fields which would meaningfully contribute to making the dataset identifiable (date of birth, for example) choose this one.
- **Other** – If you're aware of other mitigations in place, please describe them in the additional fields on this page.

Archive details
- **What archiving arrangements do you have in place?**
  - **Via Imperial Archives and Corporate Records Unit (ACRU)** - This should be your default choice for datasets stored in Imperial systems. Note that ACRU guidance may change between project planning or data collection and project completion, so please check what the most up-to-date guidance is when your project gets close to completion. Update this field when you do.
  - **Via local arrangements (please provide details)** - Especially in the case of datasets shared with collaborators or stored in collaborator's environments, archiving may be handled according to their guidance. This may also be the case if you are using a local secure environment such as the BDAU, Secure Enclaves or other specialised local research teams.
  - **Via ICT WEEE regs disposal of equipment** - Some datasets won't need to be archived or won't need to be archived any further. Where such data is stored on locally owned storage, the best thing to may be to physically destroy the physical storage medium. Where this is the case, this must be done according to ICT WEEE regs.
  - **College provided confidential waste bins**
  - **Don't know**
  - **N/A** – if the intention is the data will be retained indefinitely then please explain
- **Please provide any additional archiving details**

- o Please provide any information you have about long-term storage arrangements and what environments will be used. We appreciate that you may not know how data will be archived at the start of the project, but if you don't record this at initial registration, please remember to return to the registration and record archiving arrangements before the project ends. It's a good idea to contact ACRU to check our most up-to-date guidance on archiving when you do.
  Some datasets, especially those sourced from third parties, may place certain requirements on how data is to be archived. If that's the case (please check any agreements with providers you have), you should follow their instructions and record them here. Other datasets, such as those which are processed as part of a collaboration, may have archiving arrangements which are agreed with a collaborator. Please record whatever is agreed.
- **Please provide details about data destruction procedures or advise if they are not in place**
  - o We appreciate that for many projects you may not know the answer to this question yet. If the project data will be stored within Imperial, we will follow whatever Imperial's solution is at the end of the retention period. Some datasets, especially those sourced from third parties, may place certain requirements on how data is to be destroyed and / or how data destruction is to be recorded. If that's the case (please check any agreements with providers you have), you should follow their instructions and the process should be recorded here.
- **Please advise what data arrangements you have in place to ensure personal data is protected while being sent or received (e.g. for bringing dataset in and out of its main storage location)**
  - o Please describe how you're transferring data – if you're transferring to or from a collaborator or Imperial has certain recommend tools for transferring data. In general, transfers should be encrypted. Please describe any protections in place

51 Following completion of the last question you will able to 'Save and return' where, if completed fully, the Status for the dataset will turn Green.

52 Once you have repeated this process for all datasets, press 'Submit my declaration for approval' which should now be visible/able to be pressed.

53 Upon submitting a final message will appear advising you DART has been completed and sent for triage and assessment by the relevant team.

54 A confirmation email will be sent to all named parties which confirms the DART has been submitted

Assessment Process

55 Following receipt, the DART will be assigned to an assessor for review. As part of this process they will engage with you, provide feedback / guidance / suggestions and upon completion the DART will be finalised / signed off and a notification to confirm completion will be sent

Additional Resources

Forms;

- Template – Anonymisation Form https://imperiallondon.sharepoint.com/:w:/r/sites/cf/DPIA/_layouts/15/Doc.aspx?sourcedoc=%7B0CEF1415-CBD8-4A60-BEA6-EB91D7021841%7D&file=Template%20-%20Anonymisation%20Form.docx&action=default&mobileredirect=true&CID=2D202C31-FB8A-4DBF-9D3E-D8F1912FC85C
- Template – Legitimate Interests Assessment (LIA) / https://www.imperial.ac.uk/media/imperial-college/administration-and-support-services/legal-services-office/internal/Legitimate-Interest-Assessment_Updated-(2).docx

Website

- Main website / www.imperial.ac.uk/data-aware
- DART Support Site / https://www.imperial.ac.uk/admin-services/governance/policies-and-guidance/data-assessments/

Policy, Procedures and Code of Practice

- Code of Practice 05 – Information Asset Register / https://www.imperial.ac.uk/media/imperial-college/administration-and-support-services/legal-services-office/public/data-protection/DPA-CoP-05---Information-Asset-Register.pdf
- Code of Practice 07 – Data Protection Impact Assessment / https://www.imperial.ac.uk/media/imperial-college/administration-and-support-services/legal-services-office/public/data-protection/DPA-CoP-07---Data-Protection-Impact-Assessment.pdf
- FoM Code of practice – Handling of Health and Social Care Research Data / https://imperiallondon.sharepoint.com/sites/fom/operations/fomig/Shared Documents/Forms/AllItems.aspx?id=%2Fsites%2Ffom%2Foperations%2Ffomig%2FShared%20Documents%2FFoM-CoP-IG-02%2FFoM%20Code%20of%20Practice_HandlingHealthSocialCareData.pdf&viewid=f86ea4d5-9764-4cc2-bb0b-f70d88fa84e4&parent=%2Fsites%2Ffom%2Foperations%2Ffomig%2FShared%20Documents%2FFoM-CoP-IG-02
- FoM Information Governance Code of Practice / https://imperiallondon.sharepoint.com/sites/fom/operations/fomig/Shared Documents/Forms/AllItems.aspx?id=%2Fsites%2Ffom%2Foperations%2Ffomig%2FShared%20Documents%2FFoM-CoP-IG-01%2FFoM%20Code%20of%20Practice%201_IG.pdf&parent=%2Fsites%2Ffom%2Foperations%2Ffomig%2FShared%20Documents%2FFoM-CoP-IG-01