ImpactGo

DPIA: Beneficiary Data

Version 1.0

Date: 7 July 2024

Table of Contents

1. Controller Details

Name of Controller	Craig Russell
Title of DPO	Director, CM Russell Limited
Controller Contact	craig.russell@impact.go.uk

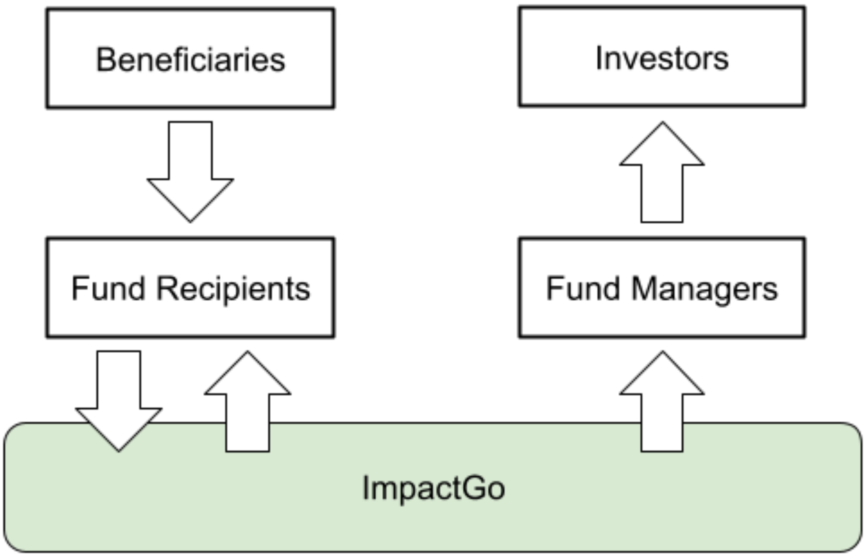
2. The Need for a DPIA

2.1. Social Impact Investment is a financial services industry focussed on providing funding to organisations that provide direct benefit to the most vulnerable people in society. While definitions differ, Big Society Capital (BSC) uses the following:

> *"Social impact investment provides repayable finance to enterprises with a social purpose like charities and social enterprises. The investment enables them to deliver products or services that create measurable, lasting social impact that improves people's lives. It also aims to make a financial return for investors"* [1]

2.2. ImpactGo is a web-based software-as-a-service (SaaS) application sitting between Fund Managers and Fund Recipients allowing them to safely share social impact data, such that Fund Managers can analyse and report on the social impact of their investments.

2.3. ImpactGo allows Fund Recipients to share data about their Beneficiaries (who for the purposes of this DPIA are the data subjects) with Fund Managers while enforcing the rights of the Beneficiary. By ensuring that Beneficiary data is collected anonymously (no Beneficiary personal data is held in the platform) ImpactGo protects the rights of the data subject, reduces the administrative overhead for the Fund Recipient, and enables a range of meaningful analysis and reporting for Fund Managers.



2.4. The ICO has issued guidance on the use of anonymisation for data disclosures. [2]

a. The guidance is clear that anonymous data is not personal data.

> *"Data protection law does not apply to data rendered anonymous in such a way that the data subject is no longer identifiable."*

b. And goes on to say that

> *"…anonymisation safeguards individuals' privacy and is a practical example of the 'privacy by design' principles that data protection law promotes."*
>
> *"…There is clear legal authority for the view that where an organisation converts personal data into an anonymised form and discloses it, this will not amount to a disclosure of personal data."*

2.5. ImpactGo is designed to minimise the risk that Fund Recipients provide Beneficiary personal data to ImpactGo, either accidentally or maliciously. This principle, carefully enforced, ensures that Beneficiary data is held anonymously in ImpactGo, and therefore is not personal data, and that the re-identification risk from the processing of this data is sufficiently remote to protect the data subject.

2.6. To provide a structured assessment of this project, the UKAN Anonymisation Decision Making Framework[3] ("ADF") is used to asses risk of, and design mitigations for, the processing and ongoing security of anonymous Beneficiary data. A Data Situation Evaluation, following the process defined in the ADF, indicates that this project does not require a risk assessment (Appendix A), however given the nature of the project (specifically the use of anonymisation) it is felt that conducting a DPIA is appropriate.

2.7. The need for a DPIA has been identified to ensure that appropriate safeguards are in place to minimise the risk of Beneficiary personal data being submitted to the service, and minimise the risk of Beneficiary data being de-anonymised by CM Russell Limited, users of ImpactGo or a motivated intruder.

2.8. This DPIA explains the nature of processing, risk assessment, and risk mitigations, as this pertains to Beneficiary data, in the design of ImpactGo.

2.9. In the interest of building trust, transparency and accountability, this DPIA, and subsequent DPIAs, will be published on the ImpactGo website[4].

3. The Nature of the Processing

3.1. The project aims to achieve a reduction in risk when disclosing Beneficiary data between Fund Recipients and Fund Managers and third-parties, relative to the current state-of-affairs. The project also aims to achieve a reduction in the use of resources, on the part of Fund Recipients, in complying with impact reporting obligations, such that these resources can be better deployed to provide more support to vulnerable groups. Furthermore, the project aims to improve the utility of this data, such that Fund Managers, their investors, and other related parties, can better understand the impact of their investments, and consequently, allocate capital resources to effect a greater positive effect on society. These aims will be achieved by providing:

a. a secure data sharing portal which supports the lawful disclosure of Beneficiary data from Fund Recipients to Fund Managers, and minimises the risk of non-compliance issues;

b. a service that is intuitive to use, and supports the processes and workflows of users, such that they experience a reduction in reporting effort over the current state-of-affairs;

c. tools to support Fund Recipients (as data controllers) in their compliance obligations to data subjects;

d. analytics and reporting tools for Fund Managers to understand the impact of their investments on the Beneficiaries of those investments, such that they can provide effective support to their Fund Recipients, and guide decision making on future investments; and

e. report outputs to Fund Managers and Fund Recipients so that they can demonstrate the impact of their work to relevant third-parties, and attract further funding to their services.

3.2. The intended effect on data subjects is that they will have greater access to the support they need for their life challenges, and that this support will provide more effective intervention to them personally, and to the wider betterment of society.

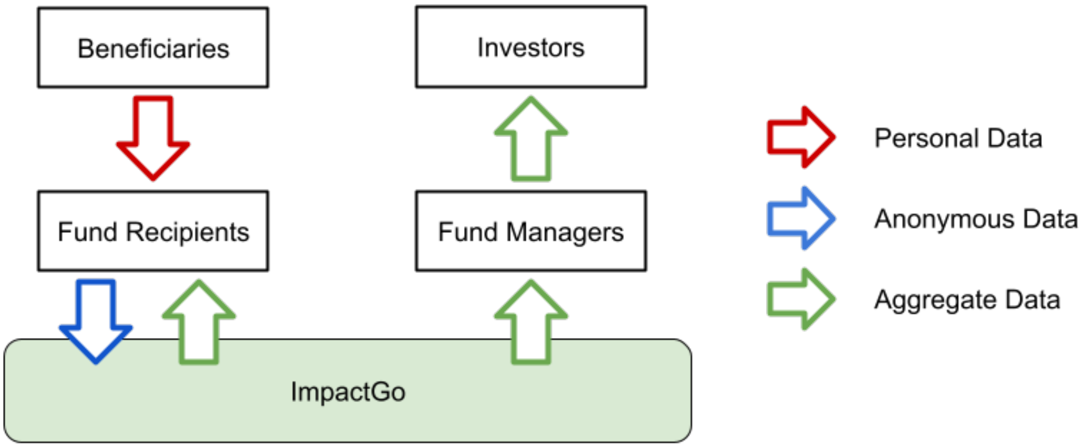
3.3. Data will be collected through ImpactGo, an online data sharing portal. Access to this portal will be restricted to authorised users only, who must accept the service terms and conditions[5] which bind the user to certain rights and obligations with regards to the disclosure and processing of Beneficiary data. These are summarised in Section 5.

3.4. Data will be stored in a secure database, managed by approved sub-processors under the terms of our DPA[6].

3.5. The source of the Beneficiary data are Fund Recipients, typically charities or other organisations who provide front-line services to support vulnerable individuals (“Beneficiaries”), who are customers of Fund Managers.

3.6. Fund Managers may also provide data about their organisation, or their Fund Recipients organisations, which may be associated with Beneficiary data in the normal use of the service.

3.7. The following provides an overview of the data flow between parties:



a. Beneficiaries will share their personal data with Fund Recipients, which may be conditional on the Beneficiary receiving support and other services from the Fund Recipient. Fund Recipients will hold this data in accordance with their own data processing policies.

b. Fund Recipients will disclose a controlled subset of Beneficiary data to CM Russell Limited via ImpactGo. ImpactGo will strictly limit the data permitted to be input into the portal to ensure the data is anonymous at the point of submission.

c. The service will process this data to provide reports, analytics and other aggregate data to Fund Recipients (limited to data they have provided) and Fund Managers (limited to data provided by their Fund Recipients). These reports are carefully designed to minimise the risk of re-identification of the data subject.

d. These reports, in whole or part, may be shared with third-parties, or published, in the course of their normal business operations. An example of this is the Sustainability

Reporting Standard for Social Housing[7], which incorporates aggregate Beneficiary metrics, alongside ESG metrics and others, to measure sector performance against targets.

3.8. Two aspects of the data processing have been identified as potential high risk to the data subject. ImpactGo is designed to support the Fund Recipient in the safe disclosure of Beneficiary data to their Fund Managers, and implements safeguards to minimise these risks as detailed in Sections 6 and 7.

a. When Fund Recipients submit Beneficiary data to ImpactGo, they may disclose data that leads to the data subject being identified, which in the hands of a motivated intruder, may disclose information about the data subject which they would not expect to be disclosed, and could lead to harm.

b. When reports or analysis are produced from Beneficiary data in the platform, it may be possible for a motivated intruder to re-identify data subjects from the dataset, either by "singling out" individuals to which the motivated intruder has prior knowledge, or through joining the dataset with some other data source. This may disclose information about the data subject which they would not expect to be disclosed, and could lead to harm.

3.9. Fund Recipients may provide the following data fields about Beneficiaries (more information provided in Appendix B):

a. Case Identifier for the Beneficiary in the Fund Recipients records

b. Local Authority in which they reside

c. Gender

d. Number of Dependents

e. Ethnic Group

f. Year of Birth

g. Categories indicating their vulnerability group

h. Date they started receiving support from the Fund Recipient

i. Date they finished receiving support from the Fund Recipient

j. Category indicating if the support they received was successful

3.10. Fund Managers may provide the following data to the platform, which may be directly, or indirectly, linked to Beneficiary records:

a. Name of Fund Recipients

b. Web address of Fund Recipients

c. Value of investments made to Fund Recipients

d. Date of investments made to Fund Recipients

3.11. The ADF provides a framework for classifying these variables the purpose of managing the risk of de-anonymisation. This is provided in Appendix C and summarised in the following statements. The data has:

a. one direct identifier, the Case ID which associates the record with the Case record held by the Fund Recipient;

b. seven indirect identifiers, which are demographic data; and

c. four target variables, which are are data about the support the Beneficiary received.

3.12. Fund Recipients will provide data about some/all Beneficiaries they support, for which they have received financial investment from a Fund Manager. Fund Recipients will submit and update Beneficiary data at their convenience using the data sharing portal. The number of individuals affected will vary, depending on the quantity of Beneficiary data provided by Fund Recipients. Data is expected to be reasonably up-to-date, accurate and generally free of errors.

3.13. Data will be retained for the duration of the Service Term, as set out in the Terms and Conditions[8], which are incorporated into a legal agreement between CM Russell Limited and Fund Managers and/or Fund Recipients, and in accordance with our DPA[9].

3.14. CM Russell Limited operates solely in the United Kingdom, and anticipate all customers, and users of the service to be based in the United Kingdom. CM Russell Limited does make limited use of sub-processors in other territories. The use of, and controls imposed upon, these sub-processors is set out in our DPA[10].

3.15. CM Russell Limited has no direct relationship with the data subjects. Data subjects will have a direct relationship with Fund Recipients, who provide them with support and other services.

3.16. Fund Recipients and Fund Managers will each have a direct relationship with CM Russell Limited, as users of ImpactGo, and will be required to accept our Terms and Conditions[11] to use the service.

3.17. Data subjects will have control over their data, as enforced by their Fund Recipients, under their data protection policies. ImpactGo is designed to support Fund Recipients in these obligations to their Beneficiaries as described in Section 5.

3.18. Beneficiaries would reasonably expect their Fund Recipients to use their data in the delivery of their services, which includes securing funding for those services. This has been confirmed in consultation with Fund Managers, Fund Recipients and other parties with knowledge of the social impact investment industry.

3.19. As an aim of the service is to measure, and report on, the impact of support provided to vulnerable groups, it is understood that, consequently, all data subjects are recognised as vulnerable individuals, and the design of the service respects them as such.

3.20. Fund Recipients currently provide similar data to their Fund Managers, usually as a condition of receiving funding from the Fund Manager. Typically, this data is shared in spreadsheets over e-mail, or through in-house portals. ImpactGo aims to improve on this state-of-affairs by providing a secure data sharing portal such that all parties can reduce the risk, better meet compliance obligations, improve operational efficiency, and increase utility of the disclosed data.

3.21. With regards to the use of web-based portals for disclosing data between organisations, the current state of technology is well established, and not considered to be novel. Use of similar technologies to reduce risk, and increase compliance efficacy is common, notably in the ESG reporting industry.

3.22. The state-of-the-art for secure data anonymisation is still evolving, ImpactGo will adopt current best-practices as described in guidance published by the ICO[12] and supported by the decision making framework published by UKAN[13]. Approaches and techniques used in the design of ImpactGo will be regularly reviewed, and updated, as these best-practices evolve.

4. Consultation Process

4.1. CM Russell Limited has, and will continue to, consult with various individuals in the social impact investment industry, who are representative stakeholders of the project. This includes, but is not limited to:

a. Fund Recipients, registered providers[14], charities and other bodies delivering front-line services

b. Fund Managers of social impact funds

c. Investors in social impact funds

4.2. CM Russell Limited has, and will continue to, consult with information security experts seeking guidance and advice on the project. Including but not limited to:

a. The UK Anonymisation Network[15], who provide guidance in the use of anonymisation techniques

b. The ICO[16], who will be consulted on this DPIA

4.3. In the interest of building trust, transparency and accountability, this DPIA, and subsequent DPIAs, will be published on the ImpactGo website[17].

5. Necessity and Proportionality

5.1. In accordance with the ICO guidance, the lawful basis for holding and processing this data it that the data, in being anonymous, is not personal data, and therefore not subject to the Regulation (EU) 2016/679 of the European Parliament and of the Council ("GDPR"). However, other legal restrictions and obligations do still apply, though this is out of scope for this DPIA.

5.2. This processing does achieve the purposes set out in section 3.1. Anonymisation of the Beneficiary data protects the rights of the data subjects, simplifies compliance for the Fund Recipients, while still allowing meaningful analysis for the Fund Recipients.

5.3. It may be possible to achieve the same outcome without anonymising the data. This would require identifying a lawful basis under Article 9 of the GDPR, as some of the data processed may be considered to be special category data.

a. It could be argued that personal data can be processed under Article 9.2(g), "substantial public interest", but this may have to be justified for each customer of the platform, which for a project of this type may be detrimentally complex to implement.

b. Additionally, personal data could be processed under Article 9.2(a) "explicit consent", but as CM Russell Limited has no direct relationship with data subjects, obtaining consent would likely not be practical.

5.4. In either case, processing personal data under GDPR Article 9 would not necessarily yield the same level of protection for the data subject as would processing the data anonymously, neither would processing personal data yield a greater analytical benefit over processing anonymous data, therefore data anonymisation has been pragmatically selected for ImpactGo. Furthermore, this in alignment with the 'privacy by design' principles that data protection law promotes.

5.5. This DPIA assesses the risk, and identifies risk mitigations, at the onset of this project. These considerations have deeply informed the design of the service, and (in consideration of ongoing risk mitigation) impose restrictions on future capabilities. Follow on work, which may be functional enhancements to the project or it's application in a new business domain, may require additional DPIAs to be undertaken. The circumstances requiring additional DPIAs include, but are not limited to:

a. Changes to the data stored and processed about data subjects, which may be modifications to existing fields or introducing additional fields;

b. Introducing new data sources into the service, to be joined with Beneficiary data, for the purposes of extending the reporting capabilities;

c. Changes to user access control or authentication policies as enforced by the service;

d. Changes to the reporting capabilities to provide new analyses from existing data;

e. Introducing a new "type" of organisation to the service, in addition to Fund Managers and Fund Recipients;

f. Changes to relevant internal policies and processes;

g. Changes to the legal status of CM Russell Limited in relation to other companies (e.g. merger or acquisition).

5.6. The service ensures data quality by limiting the set of input fields to the minimum required to provide the necessary functionality. These fields have been designed to accept restricted input, such that the data provided to ImpactGo is well structured and well defined (Appendix B). Additionally, the service provides Fund Recipients with reports, produced from their submitted data. It is expected that Fund Managers will find value in using the platform, and consequently, will be incentivised to provide timely and accurate data.

5.7. Data subjects are not direct users of the platform and will not be directly provided information on the use of their data. However, Fund Recipients are direct users of the platform, and, as data controllers, have responsibilities towards their data subjects. Fund Recipients will be provided resources explaining how data is used (including this DPIA) which they can pass onto the data subject as required.

5.8. CM Russell Limited supports Fund Recipients, in their role as data controllers, by:

a. providing software tools to manage (create, edit, delete) Beneficiary data;

b. auditing user access to, and management of, Beneficiary data;

c. limiting the use of Beneficiary data as set out in the Terms and Conditions; and

d. providing relevant support documentation and resources.

6. Risk Assessment

Describe source of risk and nature of potential impact on individuals. Include associated compliance and corporate risks as necessary.	**Likelihood of harm**	**Severity of harm**	**Overall risk**
Identifying data provided in Case Id field A Fund Recipient may submit identifying information into the Case Id field, this would allow the Beneficiary to be identifiable from this data. This would be a breach of CM Russell Limited compliance obligations, as we have no lawful basis for holding personal data about Beneficiaries. The risk of harm to the data subject is likely to be minimal, as this data is not shared with other customers.	Possible	Minimal	Medium
Intruder gains access to Fund Recipient account A motivated intruder may gain unauthorised access to a Fund Recipient user account. This scenario is discussed in depth in Appendix D: Scenario A1.2	Remote	Significant	Low
Intruder gains access to Fund Manager account A motivated intruder may gain unauthorised access to a Fund Manager user account. This scenario is discussed in depth in Appendix D: Scenario A1.2	Remote	Minimal	Low
Intruder gains access to ImpactGo platform A motivated intruder may gain unauthorised administrator access to a ImpactGo services, infrastructure and/or database. This scenario is discussed in depth in Appendix D: Scenario A1.2	Remote	Significant	Low
Intruder attempts to re-identify individuals from publicly disclosed output This scenario is discussed in depth in Appendix D: Scenarios A3.1, B1.2, B3 and B7.2	Remote	Significant	Low

Describe source of risk and nature of potential impact on individuals. Include associated compliance and corporate risks as necessary.	**Likelihood of harm**	**Severity of harm**	**Overall risk**
False data provided and/or data incorrectly deleted Proving false data or deleting data is possible, in fact a certain degree of user error is expected in a system of this type. The risk of harm to data subjects is minimal as minor errors in the data set, and by extension reporting outputs, are highly unlikely to lead to consequences for the data subject.	Possible	Minimal	Medium
Fund Manager singles out individual in report output Report output may include low number values for some aggregated metrics (e.g. one Beneficiary with addiction issues), combinations of low number values across segments, may provide information about individuals. The risk of harm to data subjects is low. Fund Managers often have face-to-face exposure with some Beneficiaries as part of their relationship with Fund Recipients. Should a Fund Manager be motivated to cause harm, they likely have access to other opportunities to do so, outside of ImpactGo.	Remote	Minimal	Low
Data subject unable to exercise their rights over their data ImpactGo provides tools to support Fund Recipients in their obligations to their data subjects. If a Fund Recipient is unable to act upon these requests, this may be frustrating for the data subject, but is not expected to cause them harm.	Remote	Minimal	Low
Internal data handling processes not followed CM Russell Limited has certain obligations and responsibilities, as set out in our terms and conditions. If, for whatever reason, these obligations are not followed, data subject’s rights may not be as robustly enforced as they otherwise would be. However, as ImpactGo is a “privacy by design” service, the risk of harm to the data subject is low.	Remote	Minimal	Low

Describe source of risk and nature of potential impact on individuals. Include associated compliance and corporate risks as necessary.	**Likelihood of harm**	**Severity of harm**	**Overall risk**
Customer obligations not followed Customer's obligations and responsibilities are set out in our Terms & Conditions, which all users of ImpactGo must sign up to. If, for whatever reason, a customer does not meet these obligations, data subject's rights may not be as robustly enforced as they otherwise would be. However, as ImpactGo is a "privacy by design" service, the risk of harm to the data subject is low.	Remote	Minimal	Low

7. Measures to Reduce Risk

Risk	Options to reduce or eliminate risk	Effect on risk	Residual risk	Measure approved
Identifying data provided in Case Id field	The user interface of ImpactGo has been designed to detect possible identifying data in this field and prompt the user if this is the case (see Appendix B). Our T&Cs permit CM Russell Limited to audit beneficiary data for suspected identifying data, and to remove any found. Should an investigation be required, audit logs are able to ascertain when the data was provided to ImpactGo, and which users have accessed it.	Reduced	Low	Yes
Intruder gains access to Fund Recipient account	This scenario is discussed in depth in Appendix D: Scenario A1.2	Reduced	Low	Yes
Intruder gains access to Fund Manager account	This scenario is discussed in depth in Appendix D: Scenario A1.3	Reduced	Low	Yes
Intruder gains access to ImpactGo platform	This scenario is discussed in depth in Appendix D: Scenario A1.4	Reduced	Low	Yes
Intruder attempts to re-identify individuals from publicly disclosed output	This scenario is discussed in depth in Appendix D: Scenarios A3.1, B1.2, B3 and B7.2	Reduced	Low	Yes
False data provided and/or data incorrectly deleted	We accept that a certain degree of incorrect data is expected for a product of this type. As the risk of harm to data subjects is minimal, we accept this risk.	Accepted	Low	Yes

Risk	**Options to reduce or eliminate risk**	**Effect on risk**	**Residual risk**	**Measure approve d**
Fund Manager singles out individual in report output	We recognise that there may be some circumstances where individuals may be singled out in aggregate reports. Strategies against this might be to round up values below 10, though this would impact the utility of the report output. As this constitutes a negligible risk of re-identification and the risk of harm to data subjects is minimal we accept this risk.	Accepted	Low	Yes
Data subject unable to exercise their rights over their data	As CM Russell Limited has no direct relationship with data subjects, it is difficult to offer direct support on this issue. However, as the Fund Recipients are strongly motivated to support Beneficiary's in their need, we find this scenario to be highly unlikely to cause harm and accept this risk.	Accepted	Low	Yes
Internal data handling processes not followed	CM Russell Limited is a small business, with low staff numbers, as consequence employees are heavily motivated in their responsibilities towards the business and it's customers. Future employees will be trained on these responsibilities, policies will be reviewed and updated, and internal access controls put in places as appropriate.	Accepted	Low	Yes
Customer obligations not followed	Customers are strongly motivated to support Beneficiaries, including respecting their rights. However, should mistakes occur, CM Russell Limited has auditing procedures in place to identify and address issues.	Reduced	Low	Yes

8. Sign Off and Outcomes

Item	Name/position/date	Notes
Measures approved by:	Craig Russell	
Residual risks approved by:	Craig Russell	
DPO advice provided:	Craig Russell	
Summary of DPO advice:		
DPO advice accepted or overruled by:		
Comments:		
Consultation responses reviewed by:		
Comments: This DPIA submitted to ICO for review. ICO have confirmed that no formal review is required as there are no unmitigated high risks.		
This DPIA will kept under review by:	Craig Russell	

Appendix A: ADF Data Situation Evaluation

A. Agreement Sensitivity	
1. Are the data subjects aware that their data have been collected in the first place?	Yes
2. Have the data subjects agreed (explicitly or implicitly) to the collection of their data?	Yes
3. Were the data subjects completely free to agree to the collection of their data (or have they agreed to collection because they want something (a good or service) for which are required to hand over some data before they can obtain it)	Yes
4. Are the data subjects aware of the original use of their data?	Yes
5. Have the data subjects agreed (explicitly or implicitly) to the original use of their data?	Yes
6. Have the data subjects agreed in general to the sharing of a functionally anonymised version of their data?	Not Known
7. Are the data subjects aware of the specific organisations that you are sharing a functionally anonymised version of their data with?	No
8. Have they agreed to your sharing their data with those organisations?	No
9. Are the data subjects aware of the particular use to which their functionally anonymised data are being put?	No
10. Have they agreed to those uses?	No
A: Count the Number of No's	**5**

B. Expectation Sensitivity	
1. Does your organisation have a relationship with the data subjects such that a reasonable data subject would expect you to have access to their data?	No
2. Does the receiving organisation have a relationship with the data subjects such that a reasonable data subject would expect them to have access to their data?	No
3. Is the receiving organisation a government or commercial entity?	Yes
4. Is your organisation's area of work one where trust is operationally important (e.g. health or education)?	Yes
5. Will you receive financial or commercial benefit from the data share?	Yes
6. Is there an actual or likely perceived imbalance of benefit arising from the proposed share or release? e.g. is the data controller benefiting but the data subjects not?	No
B. Add the Number of Yes's to questions 3-6 and the number of no's to questions 1 and 2 and then multiply by 2.	**(3 + 2) x 2 = 10**

C. Data sensitivity	
1. Are some of the variables sensitive?	Yes
2. Are the data about a vulnerable population?	Yes
3. Are the data about a sensitive topic?	Yes
4. Is the use of the data likely to be considered sensitive?	Yes
5. Do you have reason to believe that the intended use of the data might lead to discrimination against the data subjects or a group of which they are members?	No
C. Number of Yes's multiplied by 2	**4 x 2 = 8**

D. Desensitising Factors	
1. Will there be some public benefit arising from the downstream use of the data? (Yes = -3, No =0)	-3
2. Have you carried consultations with groups of stakeholders (particularly the general public and/or data subjects)? (Yes = -3, No =0)	-3
3. Have you carried consultations with groups of stakeholders (particularly the general public and/or data subjects) and implemented the recommendations arising there from? (Yes = -10, No =+3)	-10
4. Does your communication plan engender trustworthiness through transparency (sufficient to offset adverse responses in the expectation sensitivity section)? (Yes = -5, No =0).	-5
Desensitising Factors Score	**-21**

A: Agreement Sensitivity Score	5
B: Expectation Sensitivity Score	10
C: Data Sensitivity Score	8
D: Desensitising Factors Score	-21
Total Data Situation Sensitivity A+B+C+D	**2 (LOW)**

Summary Risk	**points**
1. Are the data of high quality?	
a. Yes, the data are clean and contain no or minimal errors and no or minimal missing data. (2 points)	2
b. The data contained errors but have been cleaned. (1 point)	
c. The data contains some errors and or missing data. (1 point)	
d. The data are dirty - they contain many errors and missing data issues. (0 points)	
2. How old are the data?	
a. Less than 1 year. (5 points)	5
b. 1-5 years. (4 points)	
c. 5-10 years. (3 points)	
d. 10-20 years. (2 points)	
e. More than 20 years old. (0 points)	
3. Do the data constitute a whole population or a sample?	
a. Population. (5 points)	
b. Sub-Population (4 points)	4
b. Sample. (0 points)	
4. How many variables are there that fall within the standard key variable sets?	
a. 0. (0 points)	
b. 1-4. (1 point)	
c. 5-9. (4 points)	
d. 10+. (5 points)	5
5. Which of the following best describes the data?	

a. A single aggregate output. (0 points)	
b. A set of aggregate outputs that do not overlap. (1 point)	
c. A set of aggregate outputs that do overlap. (4 points)	
d. Flat microdata. (4 points)	4
e. Hierarchical but not longitudinal microdata. (7 points)	
f. Longitudinal but not hierarchical microdata. (7 points)	
g. Hierarchical and longitudinal microdata. (10 points)	
6. Do the data include any data types that present particular reidentifiability challenges (e.g. genomics data, photographs, significant text narratives, timestamped location data or other timestamped sequences)?	
No/Yes (0/10 points)	0
7. Now considering the details of the focal environment, which of the following best describes that environment?	
a. It is a remote analysis server where users may submit code for analyses but are not able to directly access the data. Code and output are checked before the outputs are released to the user. (-25 points)	
b. It is a secure facility with on-site access with limited personnel being able to access the data that is housed within the data controller’s infrastructure (-25 points)	
c. It is a secure facility owned by the user. (-20 points)	
d. It is a remote access server with controls on the who and how of access. Users will be able to interact with the data but do not have a copy themselves (so are prevented from linking to other datasets). How users access and work with the data is pre-specified. (-20 points)	-20

e. It is a point-to-point data share based on a bespoke data sharing agreement(s) with purpose limitations, data minimisation, and specific named users. Some auditing for compliance is in place. (-5 points)	
f. It is a licensing environment. Users sign a license agreement to access the data and then are able to download them. (-2 points). Restrictions and policing of secondary use are limited.	
g. The environment is open or quasi-open (with minimal sign up conditions). (0 points)	
8. Are there data in - or which could be moved into - the focal environment that could be used to re-identify any data subjects in the data?	
Yes/No/don't know (10/0/10)	0
Summary Risk Score	**0 (Negligible)**

The Data Situation Sensitivity Score is **LOW**. The Summary Risk Score is **Negligible**. Under the Anonymisation Decision Making Framework, this project does not require a disclosure risk assessment.

		Data situation sensitivity		
		Low	Medium	High
Summary risk	High	🟥	🟥	🟥
	Medium	🟨	🟥	🟥
	Negligible	🟩 ☆	🟨	🟨

🟥 Essential	The summary risk and/or the data situation sensitivity are elevated and therefore risk assessment and control processes are going to be needed.
🟨 Borderline	You should consider carrying out a more detailed disclsoure risk assessment (and/or taking action to reduce the data situation sensitivity). Go to compoent 7.
🟩 Unnecessary	No imediate further action is necessary now but continue to monitor the data sitauation. Go to component 8.

Appendix B: Beneficiary Data Fields

This section describes the data collected about Beneficiaries, explaining how each field has been carefully designed to minimise re-identification risk.

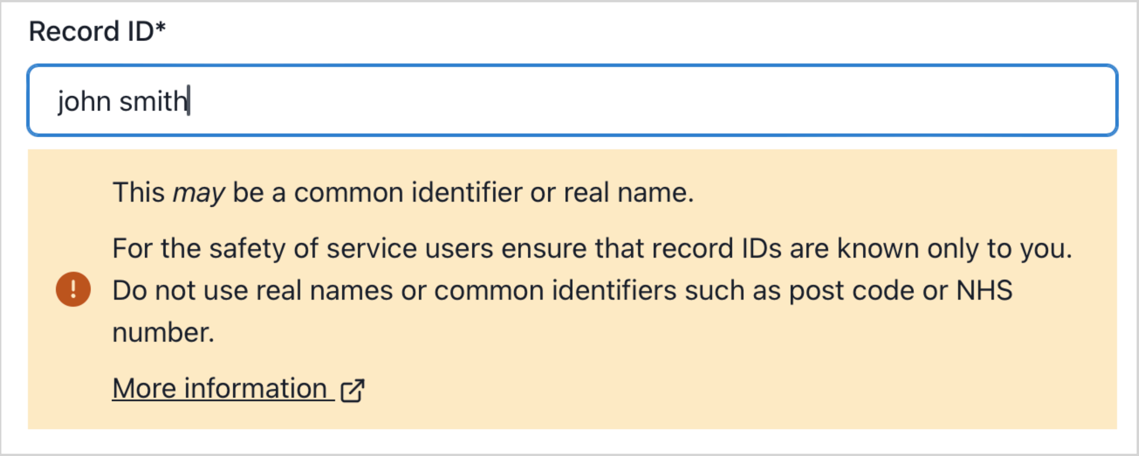
All fields are optional, except where stated.

Case identifier for the Beneficiary in the Fund Recipients records

The case identifier is a free text field for Fund Recipients to provide some internally recognised identifier (to their organisation) for the Beneficiary. This allows them to visually associate the data held in ImpactGo with Beneficiary data held internally. This is required for the ongoing management of Beneficiary data.

Guidance provided to the Fund Recipient in ImpactGo will explain that it is good practice to use identifiers that are known only within their organisation, and not to use identifiers that are found in public data sets. Furthermore, the guidance will explain that when providing data to multiple external third-parties, it is best practice to use different identifiers for each party. Fund Recipients are expected to follow their own data protection policies in this regard.

As a free text field, there is a risk that Fund Recipients may provide information here which could disclose the identity of the Beneficiary either directly (e.g. their name) or indirectly (e.g NI number). ImpactGo mitigates this risk by validating the field value against common public identifier formats (e.g. NI Number, NHS Number, Post Code, commonly used names) warning that these could risk re-identifying the Beneficiary.



Record ID*

john smith

This *may* be a common identifier or real name.

For the safety of service users ensure that record IDs are known only to you. Do not use real names or common identifiers such as post code or NHS number.

As this field is used to associate the anonymised data in ImpactGo with (potentially) identifying data held in by the Fund Recipient, there is a risk that if both data sets were made public a malicious third-party could link them together and identify the Beneficiary. It is felt that this is a minimal risk as this would require data held in two disconnected systems, in different organisations, to be compromised, which is assessed to be a low likelihood.

ImpactGo ensures this field is visible and editable by the Fund Recipient only and is not included in any data export or analysis output that is accessible to users outside of their organisation (excluding ImpactGo staff).

Identifier is a required field to ensure that Fund Recipients can manage records in ImpactGo as they relate to their internal case management system.

Local authority in which they reside

The local authority ("LA") code identifies the LA in which the Beneficiary lives.

Available options are taken from the UK Government Open Geography Portal[18]. These values are a recognised standard for identifying LAs and can be used to link data across data sets.

Location data of finer granularity (e.g address, postcode) is not collected as it might be possible to re-identify the Beneficiary if this were combined with other publicly accessible data sets, which is an unacceptable risk.

Number of dependents

This field is used to provide the number of dependents of the Beneficiary.

ImpactGo uses a loose definition of dependents which may include any individual for who is generally considered to be "dependent" on the Beneficiary. The ONS defines dependents as children of a certain age[19], but the metric used in ImpactGo may include individuals outside of this definition.

No further information about dependents is collected (e.g. age, gender) as, if the data were compromised, it might be possible to use this data in combination with others to re-identify the Beneficiary if this were combined with other publicly accessible data sets, which is an unacceptable risk. Furthermore, this would add little analytical value to the data set.

Gender

The gender code is used to provide the gender of the Beneficiary. Possible responses are:

- Male
- Female
- Prefer not to say
- Other

This field holds data on gender, not sex, in accordance with the UK Government Design System, which advises:

> *If you do need to ask, use 'sex' when you need biological data (for example, if you're providing a medical service). In all other cases, use 'gender'.*[20]

A free text field to provide more detail for the "Other" category is not provided as, if the data were compromised, it might be possible to re-identify the Beneficiary if this were combined with other publicly accessible data sets, which is an unacceptable risk. Furthermore, this would add little analytical value to the data set.

Year of birth

The year of birth is used to indicate the approximate age of the Beneficiary.

The full date of birth is not collected as it might be possible to re-identify the Beneficiary if this were combined with other publicly accessible data sets, which is an unacceptable risk. Furthermore, this would add little analytical value to the data set.

ImpactGo implements further protection of this data by banding aggregated values into age ranges when included in reports and analytics ("Under 16", "16-24", "25-50", "Over 50").

Ethnic group

This field is used to provide the ethnic group of the Beneficiary.

The options available are taken from the values used in the UK & Wales 2021 census[21].

In the database, these options are encoded as key values unique to ImpactGo.

Categories indicating their vulnerability group

This field is used to provide data on the vulnerability group of the Beneficiary.

While there is no recognised standard for categorising vulnerability groups, the options available are taken from the Good Finance Measuring Social Impact Outcomes Matrix[22], which has broad recognition in the sector.

- People experiencing long term unemployment
- People experiencing homelessness
- People living in poverty and/or financial exclusion
- People dealing with addiction issues
- People with long-term health conditions/life threatening or terminal illness
- People with learning disabilities and other neurodivergence
- People with mental health needs
- People with physical disabilities or sensory impairments
- Voluntary carers
- Vulnerable parents
- Vulnerable children
- Vulnerable young people
- Older people
- Ex/Offenders and prisoners
- People who have experienced crime or abuse
- Refugees, asylum seekers, undocumented and other migrants

In the database, these options are encoded as key values unique to ImpactGo.

Date they started receiving support from the Fund Recipient

This field is used to provide the date that the Beneficiary began receiving support from the Fund Recipient's service.

This field is required if the service end date is provided.

Date they finished receiving support from the Fund Recipient

This field is used to provide the date that the Beneficiary last received support from the Fund Recipient's service.

Category indicating if the support they received was successful

This field is used to provide information about the reason why the Beneficiary stopped receiving support from the Fund Recipient's service.

This field is required if service end date is provided.

The available options are:

- Positive
- Negative
- Indeterminate

The field is designed to capture an indication of the circumstances in which the Beneficiary left the service, without holding specific data that might risk re-identification.

For example, a Beneficiary who left in agreement with the Fund Recipient might be considered a "positive" move on. Whereas a Beneficiary who left without the agreement of the Fund Recipient might be considered a "negative" move on. The "indeterminate" option is provided for situations where the circumstances may be unclear.

Appendix C: ADF Variable Classification

Variable	Direct Identifier	Indirect Identifier	Target Variable	Notes
Case Identifier	Yes			Uniquely associates a record in ImpactGo with a Fund Recipient's case management records
Local Authority		Yes		May be a unique identifier in combination with other indirect identifiers
Gender		Yes		May be a unique identifier in combination with other indirect identifiers
No. Dependents		Yes		May be a unique identifier in combination with other indirect identifiers
Ethnic Group		Yes		May be a unique identifier in combination with other indirect identifiers
Year of Birth		Yes		May be a unique identifier in combination with other indirect identifiers
Vulnerability Group			Yes	May expose the nature of the support the data subject received e.g. drug rehabilitation
Support Start Date		Possibly		May be a unique identifier in combination with prior knowledge and/or other indirect identifiers
Support End Date		Possibly		May be a unique identifier in combination with prior knowledge and/or other indirect identifiers
Support Outcome			Possibly	May indicate the post-support state of the data subject e.g. returned to drug use
Fund Recipient Name			Yes	May expose the nature of the support the data subject received e.g. drug rehabilitation
Fund Recipient Web Address			Yes	May expose the nature of the support the data subject received e.g. drug rehabilitation
Investments Value				No concern
Investments Date				No concern

Appendix D: ADF Re-Identification Scenario Analysis

The ADF specifies a scenario based analysis for the risk of re-identification from the disclosed data set. Making use of the ADF process for building disclosure scenarios[23] and the standard key variables sets[24], the following scenarios were assessed for the risk of re-identification.

NB: Introduction paragraphs and attacker profiles (*emphasised*) are quoted directly from UKAN resources.

Scenario A1.2 Restricted access database cross match (general, extended).

This scenario is based upon an analysis of the information commonly available in restricted access databases. Attacker Profile: Person with access to restricted access dataset or hacker able to obtain such access.

This scenario assumes the following inputs:

- **Motivation:** We assume a motivated intruder seeking to gain compromising information about individuals with a view to exploiting them for financial gain or to deliberately cause repetitional harm.
- **Means:** We assume that the intruder has access to some restricted access dataset, and the skill to gain unauthorised access to ImpactGo (i.e a hacker).
- **Opportunity:** The intruder may gain access to ImpactGo data through unauthorised use of a legitimate user account (either by stealing the login credentials or gaining physical access to the users computer), or they gain unauthorised access to ImpactGo services and database.
- **Target Variables:** If the intruder were able to match the source data with records in ImpactGo, they may be able to learn the name of the service which provided support to a Beneficiary and/or the vulnerability groups describing their need. This could reveal something of the nature of the support they received e.g. drug rehabilitation. This would not explicitly re-identify the individual, but could, in combination with other identifying sources, lead to this information being exposed.
- **Goals achievable by other means?** Possibly, Fund Recipients often publish case studies of individuals who have received their support. These sometimes use redacted information,

or false names, but may not, and sometimes include photographs of the individual. These disclosures are always made with the individual's consent, which invalidates the intruders motivation.

- **Effect of Data Divergence:** While it is expected that data should be reasonably complete and accurate, ImpactGo would not contain records for the whole population of vulnerable individuals in the UK. Consequently, matches between data in ImpactGo and other sources could be considered to be "suggestive" at best, and certainly not guaranteed. Furthermore, if the intruder gained access via a user account, the set of data available to them would be limited to the system access of the user, which is much reduced in scope compared to the whole database.

This scenario assumes the following intermediate inputs:

- **Attack Type:** With the key variables below, the intruder may attempt to find records with matching these characteristics, or to look for records with unique combinations of these characteristics.
- **Key Variables:** Age, Sex/Gender, Number of Dependents.

A risk analysis of this scenario produces the following results:

- **Likelihood of Attempt:** An attack of this type is highly unlikely as an intruder would require a sophisticated skillset in order to exploit the service and gain access to the data. Furthermore the intruder would need to be highly motivated to conduct such an attack, as this would require sustained effort over a period of time, with a low chance of success.
- **Likelihood of Success:** An attack of this type is highly unlikely to succeed. An intruder would first have to gain unauthorised access to the service, which would require sustained effort on their part and a high chance of failure. Security policies and monitoring of the service (e.g. MFA authentication) would likely prevent such an attack from succeeding.
- **Consequences of Attempt:** If the intruder were successful in gaining access to the data, the likelihood of them re-identifying individuals is very low. If the intruder was successful in re-identifying individuals in the dataset, this may expose a general indicator of the support they received and the service who provided it, which is of limited impact to the data subject, and in some cases, is already in the public domain.

- **Effect of Variations in the Data Situation:** We are satisfied that the data situation is sufficiently robust to adequately mitigate the risk of this scenario. However, we will keep our security policies under constant review and enhance these where required.

Scenario A3.1: Restricted database cross match (personnel)

This scenario is based on information commonly held in personnel databases. Typically this includes considerable detail on economic characteristics such as occupation, industry, economic status, basic physical characteristics (such as age, sex and ethnic group) and some information on personal circumstances (area of residence, long term illnesses, marital status and number of children). Attacker Profile: Person working in personnel office of large organisation.

This scenario assumes the following inputs:

- **Motivation:** We assume an intruder with access to personnel information, who may work for an employer of a Beneficiary, seeking to discover compromising information about an employee.

- **Means:** We assume that they have the skills to access online information, and analyse this using spreadsheet software.

- **Opportunity:** Output from ImpactGo may be published online by Fund Managers or their Investors, as part of public disclosures declaring the impact of their investments. This presents a negligible opportunity for data to be exploited as the output available to Fund Managers is are aggregated, Fund Managers have no access to individual Beneficiary records in ImpactGo.

- **Target Variables:** Published output may include the names of Fund Recipients, or the aggregate totals of vulnerability groups. This is unlikely to expose the data subject to greater risk than the information already available to the intruder.

- **Goals achievable by other means?** As the intruder has a professional relationship with the individual, they are more likely to use organisational resources to find compromising information (e.g e-mails, chat logs) than attempt to exploit published information sourced from ImpactGo.

- **Effect of Data Divergence:** Minor errors in aggregate metrics would introduce uncertainly and reduced confidence in suspected matches.

This scenario assumes the following intermediate inputs:

- **Attack Type:** An intruder would attempt to identify specific individuals in the dataset, based on some prior knowledge.
- **Key Variables:** Address, Age, Sex/Gender, Ethnic Group, Number of Dependents, Long-term Illness.

A risk analysis of this scenario produces the following results:

- **Likelihood of Attempt:** An attack of this nature is highly unlikely. In this scenario an intruder would be much more likely to make use of other sources to discover compromising information about the data subject. Furthermore, they would have to be aware of, and source, published resources having suspected these resources to contain information about the data subject, which is a very remote possibility.
- **Likelihood of Success:** If such an attack were to take place, the likelihood of success is very remote. Data output from ImpactGo uses aggregate metrics only, not individual Beneficiary records, limiting the ability to identify individuals within the data set. Output may specify single digit records for a category (e.g. one Beneficiary with addiction issues) but this cannot be linked with other single digit categories (e.g. one female Beneficiary) preventing individuals being singled out.
- **Consequences of Attempt:** Likely none, as a successful attack is unlikely to yield information not more readily available to the intruder.
- **Effect of Variations in the Data Situation:** We are satisfied that the data situation is sufficiently robust to adequately mitigate the risk of this scenario.

Scenario B1.2 Commercial database cross match (superset, resource cost high)

This scenario is based upon an analysis of the information available in commercial databases. This is effectively a superset of available variables which could be exploited by a well-resourced attacker who links multiple data sources together. Attacker Profile: Person or organisation with sufficient resources to purchase multiple lifestyle databases.

This scenario assumes much of the same inputs as Scenario A3.1, except for a reduced set of target variables. Consequently, it is assessed that there is a lower likelihood of success than A3.1 and we conclude that the data situation is sufficiently robust to mitigate the risk of this scenario.

Scenario B3: Extended local search

This scenario corresponds to what might be obtained through estate agent details combined with the electoral register. The variables (new voter/adult) and ethnic group that could be used in a crude form from the electoral register are included in this variant. Attacker Profile: anyone.

This scenario assumes much of the same inputs as Scenario A3.1, except for a reduced set of target variables. Consequently, it is assessed that there is a lower likelihood of success than A3.1 and we conclude that the data situation is sufficiently robust to mitigate the risk of this scenario.

Scenario B7.2: Combined public, visible and commercial sources

This implies a very well-resourced attacker who is carrying out a deep information gathering exercise on a small targeted population. Note the list of variables is more extensive than might be obtained on any restricted access database. Attacker Profile: anyone.

This scenario assumes much of the same inputs as Scenario A3.1, except for a reduced set of target variables. Consequently, it is assessed that there is a lower likelihood of success than A3.1 and we conclude that the data situation is sufficiently robust to mitigate the risk of this scenario.

References

[1] Big Society Capital, **Our Approach** - https://bigsocietycapital.com/our-approach/

[2] ICO, **Anonymisation: Code of Practice** - https://ico.org.uk/media/for-organisations/documents/1061/anonymisation-code.pdf

[3] UKAN, **The Anonymisation Decision Making Framework** - https://ukanon.net/framework/

[4] ImpactGo, **DPIAs** - https://impactgo.uk/legal/dpia

[5] ImpactGo, **Terms and Conditions** - https://impactgo.uk/legal/service-terms-and-conditions/

[6] ImpactGo, **DPA** - https://impactgo.uk/legal/dpa/

[7] The Good Economy, **Sustainability Reporting Standard for Social Housing** - https://thegoodeconomy.co.uk/sustainability-reporting-standard-for-social-housing

[8] ImpactGo, **Terms and Conditions** - https://impactgo.uk/legal/service-terms-and-conditions/

[9] ImpactGo, **DPA** - https://impactgo.uk/legal/dpa/

[10] ImpactGo, **DPA** - https://impactgo.uk/legal/dpa/

[11] ImpactGo, **Terms and Conditions** - https://impactgo.uk/legal/service-terms-and-conditions/

[12] ICO, **Anonymisation: Code of Practice** - https://ico.org.uk/media/for-organisations/documents/1061/anonymisation-code.pdf

[13] UKAN, **The Anonymisation Decision Making Framework** - https://ukanon.net/framework/

[14] Regulator of Social Housing, **Registered providers of social housing** - https://www.gov.uk/government/publications/registered-providers-of-social-housing

[15] UKAN - https://ukanon.net/

[16] ICO, **DPIA consultation** - https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/accountability-and-governance/data-protection-impact-assessments-dpias/do-we-need-to-consult-the-ico/

[17] ImpactGo, **DPIAs** - https://impactgo.uk/legal/dpia

[18] MySociety, **UK Local Authorities** - https://geoportal.statistics.gov.uk/datasets/ons::local-authority-districts-april-2023-names-and-codes-in-the-united-kingdom/explore

[19] ONS, **Families and households statistics explained** - https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/articles/familiesandhouseholdsstatisticsexplained/2019-08-07#families-and-households-definitions

[20] UK Government, **Design System: Gender or sex** - https://design-system.service.gov.uk/patterns/gender-or-sex/

[21] UK Government Cabinet Office, **List of ethnic groups** - https://www.ethnicity-facts-figures.service.gov.uk/style-guide/ethnic-groups

[22] Good Finance, **Measuring Social Impact Outcomes Matrix** - https://www.goodfinance.org.uk/measuring-social-impact/outcomes-matrix

[23] UKAN, **Building Disclosure Scenarios** - https://ukanon.net/framework/

[24] UKAN, **Standard Key Variables** - https://ukanon.net/framework/