NHS
Counter Fraud Authority

Data Analytics Platform

Data Protection Impact Assessment

April 2023

V3.0 Published

NHS fraud.
Spot it. Report it.
Together we stop it.

Executive Summary

The document is deemed OFFICIAL and any information viewed/obtained within it should be treated in the appropriate manner as advised and set out in the Government Security Classifications (May 2018).

More information in relation to this data classification, including the requirements for working with these assets can be found here:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/715778/May-2018_Government-Security-Classifications-2.pdf

In particular please note that:

ALL routine public sector business, operations and services should be treated as OFFICIAL - many departments and agencies will operate exclusively at this level. There is no requirement to explicitly mark routine OFFICIAL information. Baseline security measures should be enforced through local business processes. A limited subset of OFFICIAL information could have more damaging consequences (for individuals, an organisation or government generally) if it were lost, stolen or published in the media. This subset of information should still be managed within the 'OFFICIAL' classification tier, but may attract additional measures (generally procedural or personnel) to reinforce the 'need to know'. In such cases where there is a clear and justifiable requirement to reinforce the 'need to know', assets should be conspicuously marked: '**OFFICIAL–SENSITIVE'**

Nothing contained herein should be construed as endorsing any particular Technical Solution to any United Kingdom Government Invitation to Tender.

Table of contents

Document Control

PM	Ref	Document owner	Version No	Issue Date	Amendments
Analytical Intelligence Lead	DPIA – Data Analytics Platform	DPO	V0.1	April 2022	Initial creation
Analytical Intelligence Lead	DPIA – Data Analytics Platform	DPO	V0.2	April 2022	Redraft and update
Analytical Intelligence Lead	DPIA – Data Analytics Platform	DPO	V1.0	June 2022	Finalised draft following comments from IA
Analytical Intelligence Lead	DPIA – Data Analytics Platform	DPO	V1.1	June 2022	Updated following review by Information Governance
Analytical Intelligence Lead	DPIA – Data Analytics Platform	DPO	V1.2	July 2022	Finalised with confirmation from SMT lead for mitigation
Analytical Intelligence Lead	DPIA – Data Analytics Platform	DPO	V2.0	July 2022	Published Version
Analytical Intelligence Lead	DPIA – Data Analytics Platform	DPO	V2.1	April 2023	Updated following software changes (ensuring remains relevant)
Analytical Intelligence Lead	DPIA – Data Analytics Platform	DPO	V3.0	April 2023	Final updates and redactions for publication

Prefix	
Reference:	DPIA Data Analytica Platform
Date:	April 2023
Author:	Analytical Intelligence Lead
Data Owner:	N/A (dependant on the data being processed)
Version:	3.0
Supersedes	2.1

Links & Dependencies

Document	Title	Reference	Date	POC
DPA	Data Protection Act	All	2018	HMG
EU GDPR	EU General Data Protection Regulation	All	2016	GDPR
FOI	Freedom of Information Act	All	2000	HMG
Government Security Classifications	Government Security Classifications	All	May 2018	Cabinet Office
HRA	Human Rights Act	All	1998	HMG
ISO/IEC 27000	Information security management systems Standards	ISO/IEC 27001:2013	Oct 2010	ISO
IS1P1 & P2	HM Government Infosec Standard No. 1, Part 1 – Risk Assessment and Treatment	P1 - Issue 3.5 P2 – Issue 3.5	October 2009 October 2009	CESG
IS2	InfoSec Standard 2	Issue 3.2	January 2010	CESG
PECR	The Privacy and Electronic Communications Regulations	All	2003	HMG

1. Data Protection Impact Assessment Requirement & Process

Introduction

1. The General Data Protection Regulation (GDPR) 2016 introduces a new obligation to undertake Data Protection Impact Assessments (DPIAs), before carrying out types of processing **'likely to result in high risk(s) to individuals' interests'**. DPIAs are now mandatory for certain types of processing and there are specific legal requirements for content and process. Where a DPIA identifies a 'high risk' that cannot be mitigated, the Information Commissioner's Office (ICO) must be consulted.

2. DPIAs provide a way to systematically and comprehensively analyse the intended processing and help to identify and minimise data protection risks. In addition to considering compliance risks, they should also consider broader risks to the rights and freedoms of individuals, including the potential for any significant social or economic disadvantage. The focus is on the potential for harm to individuals or to society at large, whether it is physical, material or non-material.

3. To assess the level of risk, a DPIA must consider both the likelihood and the severity of any impact on individuals. It does not have to eradicate the risks altogether, but should help to minimise them and assess whether or not remaining risks are justified. A DPIA may cover a single processing operation or a group of similar processing operations. For new technologies you may be able to use a DPIA done by the product developer to inform your own DPIA on your implementation plans.

4. A DPIA must consider 'risks to the rights and freedoms of natural persons'. While this includes risks to privacy and data protection rights, it can also affect other fundamental rights and interests:

    a. "The risk to the rights and freedoms of natural persons, of varying likelihood and severity, may result from data processing which could lead **to physical, material or non-material damage**, in particular: where the processing **may give rise to discrimination, identity theft or fraud, financial loss, damage to the reputation, loss of confidentiality of personal data protected by professional secrecy**, **unauthorised reversal of pseudonymisation, or any other significant economic or social disadvantage**; where data subjects might be deprived of their rights and freedoms or prevented from exercising control over their personal data[1]…"

5. Under GDPR you must carry out a DPIA where for example you plan to:

    a. process special category or criminal offence data on a large scale.

6. The ICO also requires a DPIA to be undertaken for example, where you plan to:

    a. use new technologies;

    b. match data or combine datasets from different sources;

    c. collect personal data from a source other than the individual without providing them with a privacy notice ('invisible processing');

7. DPIAs are an essential part of the organisation's accountability obligations under GDPR and an integral part of the 'data protection by default and design approach'. An effective DPIA helps to identify and fix problems at an early stage, demonstrate compliance with data protection obligations, meet individuals' expectations of privacy and help avoid reputational damage which might otherwise occur.

[1] GDPR - Recital 75

8. Conducting a DPIA is a legal requirement for any type of processing. Failure to carry out a DPIA in required cases may leave the organisation open to enforcement action, including a **fine of up to €10 million**.

9. This DPIA is related to the NHSCFA Risk Assessments, which outline the threats and risks. The Risk Assessment document was developed in accordance with the requirements of NHSCFA and CESG HMG Infosec Standards 1 and 2.

Data Analytics Platform - General Description

10. This document contains information in relation to the NHSCFA data analytics platform and the associated processes and practices. This is designed to extent to any software package that is utilised by NHSCFA for data processing, analysis and dissemination purposes, particularly given the possibility of both replacements and additions to this software the focus will primarily be on the storage, processes, analysis and outputs as opposed to the software itself. Should any changes occur that apply to any issues of privacy, this DPIA will be updated.

11. As a number of individual assessments exist for the numerous NHSCFA data sources themselves, this DPIA is completed to consider the system itself and the manner that it may process data generally in the above manner. As such it does not concern itself with individual sources of data as these particulars are covered under more specific DPIAs

12. This is the first DPIA concerning the data analytics platform and has been carried out by the Information and Records Management Officer, in consultation with Analytical intelligence Lead and the Information Governance and Risk Management Lead.

13. The Data Analytics Platform, in addition to GDPR, is also required to comply with other relevant HMG legislation including where applicable the Data Protection Act 2018, Human Rights Act 1998 and Freedom of Information Act 2000.

Data Protection Impact Assessment

14. To ensure the databases and systems meets all legal requirements and the risks to personal data are identified and understood it is necessary to undertake a DPIA. This DPIA is based on the ICO's recommended template[2] comprised of seven steps:

    Step 1 - Identify the need for a DPIA

    Step 2 - Describe the processing

    Step 3 - Consultation process

    Step 4 - Assess necessity and proportionality

    Step 5 - Identify and assess risks

    Step 6 - Identify measures to reduce risk

    Step 7 - Sign off and record outcomes

[2] Version 0.3 (20180209)

STEP 1: Identify the need for a DPIA

Explain broadly what the system/project aims to achieve and what type of processing it involves. You may find it helpful to refer or link to other documents, such as a project proposal. Summarise why you identified the need for a DPIA.

The purpose of the Data Analytics Platform can be split into the following activities:

- Transforming and preparing data prior to building, configuring, administering, and deploying customisable dynamic web-based reports. These enable users to access their own data without the need for enhanced IT or analytical skills as many of these functions and calculations are prepared by Information Analytics Team (within the Performance and Analytics Unit) prior to release.

- Producing management information reports for strategic and tactical purposes to enable operational and management decision making to take place from an informed position.

- Querying, exploring and analysing data to support the strategic work within the NHS business plan and to answer parliamentary questions (PQs), Freedom of Information requests (FOIs) and other unplanned requests for information (RFIs) made by Ministerial, HMG, NHS, internal and external stakeholders. Examples of this include statistics for the Annual Report, Corporate and organisational dashboards.

- Exploring, analysing and profiling (understanding through review) current and historical data to identify statistical trends, patterns and anomalies not identifiable through normal methods of examination used to proactively identify fraud and where needed the ability to amend predefined parameters as well as code. In additions, tools and unsupervised learning techniques are used where the problem relevant data is suitable to do so.

- Transformation including matching, merging, filtering, extraction, analysis, profiling and mining of data for intelligence assessments and individual fraud investigations producing meaningful summaries as well as granular pattern level extraction output using large datasets.

- Designing, developing, transforming, testing, deployment, maintenance and access control of NHSCFA databases.

- Deployment, configuration, maintenance, troubleshooting and access control (both internal and external) for the Business Intelligence software enabling analysis to be undertaken and appropriate dissemination of information and intelligence.

- Importing, transforming, storing, maintaining, securing and updating datasets dynamically and making them available for analysis.

- Building, configuring and deploying secure web based dynamic reports for external NHS users.

- Distributing reports with Row Level Permissions to external stakeholders to ensure appropriate but limited access to individual tables and the data contained therein.

--

Because the nature of the platform and the wide-ranging nature of data types that can be processed, it has been considered necessary to consider the platform itself as within the DPIA framework as

a) There is the possibility from some data sources that personal data may be processed
b) There are specific risks and mitigations that must be considered separately from the data sources themselves and warrant special consideration and recording within a DPIA format

STEP 2: Describe the processing

Describe the nature of the processing:

1. How will you collect, use, store and delete data?

2. What is the source of the data?

3. Will you be sharing data with anyone (consider using a flow diagram or other way of describing data flow)?

4. Why types of processing identified as 'likely high risk' are involved?

1. How will you collect, use, store and delete data?

Data utilised by the analytics platform will be collected from data sources either within NHSCFA or from external data sources. In the case of external data sources, NHSCFA would coordinate the particulars of this extract with the data owner directly to coordinate and manage the request, stemming from the requirements through to the data share itself. In either circumstance, a NHSCFA Database -Specialist or Information Analyst/ Information Specialist would import the data onto the platform and create a view or dataset for analysts to commence exploration and transformation.

How the data will be used would be project specific, however generally speaking the outcomes would be either

a) Information Analysts will produce rule-based analysis to determine outliers. The outcomes of which would be summarized in a written report that describes the extent of the fraud risk.

   Depending on the findings themselves, the basis for sharing and processing the data and the strength of the outliers, it is possible that the outliers may be shared with either the external data providers or NHSCFA's own investigative services for further action.

b) Information Analysts produce a data-based output, for example an electronic dashboard or statistical physical report detailing output and findings, that is utilised to support NHSCFA activity and, where necessary, is shared with the wider NHSCFA, wider NHS or any external stakeholder (example range in response to an Freedom of Information enquiry , benchmarking data provided to support counter fraud activity to wider NHS organisations or a physical report focused on a specific problem and detailing the analysing and findings found)

The storage of the data is on the NHSCFA secure cloud, following migration in January 2022 to a cloud-based solution hosted on Microsoft Azure.

Retention is managed via the NHSCFA information register, and, through it, schedules are set for deletion and removal. NHSCFA also proactively review their data and will remove records once there is no longer a clear and specific purpose for their continued use.

2. What is the source of the data?

The sources of the data would be project and purpose driven and this covers a wide range of data sources too numerous to name. However, there are a number of key data systems and data owners that are internal and external to the NHCFA and would be considered fundamental to expected / business as usual activities for the software. These are:

Internal:

- NHSCFA Case Management System (CLUE), as well as legacy systems (FIRST)
- NHSCFA Intelligence Database (IBase)

- Performance Reporting Database (Verto) as well as legacy systems (MRT)
- The LCFS/DoF nominations database (CPOD)

External

- The NHS provider and commissioning sectors via the NHSCFA Data Capture System (DCS) as calibrated for specific data capture exercises
- NHS staff and the general public via the NHSCFA Fraud and Corruption Reporting Tool (FCROL) –n.b This data is gathered indirectly, through records captured via CLUE and Ibase
- NHSBSA Dental and Pharmaceutical services (project specific)
- NHSBSA Pharmaceutical routine data source e.g. monthly payment schedules
- HM Cabinet Office via National Fraud Initiative (NFI) data
- External open-source data (e.g. Companies House, Ordnance Survey etc)

3. Will you be sharing data with anyone (consider using a flow diagram or other way of describing data flow)?

The purpose and application of data that is processed by the Data Analytics Platform is project driven and may include the following:

- Sharing findings (outliers) internally for the purposes of intelligence, loss measurement and investigation and/or to support fraud prevention activity to mitigate identified fraud risks
- Sharing findings externally with the data controllers to support their own remedial activity and fraud prevention
- Provision of summary data internally and externally to support a variety of performance reporting techniques that measure the impact of activity within NHSCFA and across the wider NHS

4. Why types of processing identified as ‘likely high risk’ are involved?

Although it is impossible to be specific as each type of data processed will pertain to the requirements and defined objectives of the project that warrant it, there is some potential for the use of high risk data. However, the nature of NHS fraud concerns deceptive and dishonest activity concerning the treatment of patients and the services and support activity that support the NHS in delivering this function. The data that identifies this fraud therefore necessarily pertains to these activities and, in doing so, may contain information about associated persons, for example alleged perpetrators and witnesses – which may include NHS patients and the members of staff.

In particular, outlier detection that relates to the treatment of patients may necessitate details of individual steps of clinical activity provided to individual patients in order to determine courses of linked treatment.

Finally, the case management system contains information about ongoing and concluded investigations, including outcomes of sanctions.

Describe the scope of the processing:

1. What is the nature of the data and does it include special category or criminal offence data?

2. How much data will you be collecting and using?

3. How often?

4. How long will you keep it?

5. How many individuals are affected?

6. What geographical area does it cover?

1. What is the nature of the data and does it include special category or criminal offence data?

These would be specific to each project and the data sources being applied and therefore would defer to the DPIA and other data usage considerations being applied for these.

2. How much data will you be collecting and using?

These would be specific to each project and the data sources being applied and therefore would defer to the DPIA and other data usage considerations being applied for these.

3. How often?

These would be specific to each project and the data sources being applied and therefore would defer to the DPIA and other data usage considerations being applied for these.

4. How long will you keep it?

These would be specific to each project and the data sources being applied and therefore would defer to the DPIA and other data usage considerations being applied for these. NHSCFA maintains a register of data assets which is used to consider and schedule retention of data and these would be applied to the appropriate data source

5. How many individuals are affected?

These would be specific to each project and the data sources being applied and therefore would defer to the DPIA and other data usage considerations being applied for these.

6. What geographical area does it cover?

Although these would be specific to each project, the NHSCFA only provide services to the NHS in England and Wales and so it would not be expected in any circumstances to be processing data from outside this catchment area.

Describe the context of the processing:

1. What is the nature of your relationship with the individuals?

2. How much control will they have?

3. Would they expect you to use their data in this way

4. Do they include children or other vulnerable groups?

5. Are there any prior concerns over this type of processing or security flaws?

6. Is it novel in any way?

7. What is the current state to technology in this area?

8. Are they any current issues of public concern that you should factor in?

9. Are you signed up to any approved code of conduct or certification scheme (once any are approved)?

1. What is the nature of your relationship with the individuals?

Although these would predominantly be specific to each project and the data sources being applied, the category of individual can be summarised in the following manner:

- Patients and members of the public
- NHS Employees (including clinicians as well as other support workers and administrative staff)
- Person(s) otherwise concerned with NHS fraud cases

However in terms of the specifics this would defer to the DPIA and other data usage considerations being applied for each project or data source.

2. How much control will they have?

These would be specific to each project and the data sources being applied and therefore would defer to the DPIA and other data usage considerations being applied for these.

3. Would they expect you to use their data in this way

These would be specific to each project and the data sources being applied and therefore would defer to the DPIA and other data usage considerations being applied for these.

4. Do they include children or other vulnerable groups?

These would be specific to each project and the data sources being applied and therefore would defer to the DPIA and other data usage considerations being applied for these.

5. Are there any prior concerns over this type of processing or security flaws?

Although these would predominantly be specific to each project and the data sources being applied, the nature of NHS fraud and the data that exists in relation to it (in any context) it would be expected that this would have potential to be contentious as it is likely to concern the treatment of patients and may require information about the recipient of treatments or NHS employees who provided or supported this work

6. Is it novel in any way?

This would be dependent on each project and the data sources being used and techniques being applied.

7. What is the current state to technology in this area?

The current NHSFCA data platform is a cloud based analytical system that utilises the SAS software package, It was procured on 31st March 2021 for a 3 year licence following a tender exercise, fitting the specifications of the organisation. The tool is accompanied by use of PowerQuery and PowerBI for external reporting transformation, report creation and dissemination of data.. Key elements of the tools used include:

- Extract, Transform and Load (ETL): Pulling data from one source, transforming the data into a standardised format and placing it into another database.
- Analytics / Data Mining: Querying both structured and unstructured data to provide information to internal and external stakeholders, analysing data to identify trends and create rules-based processes to highlight anomalies as well as utilising data mining techniques both supervised and unsupervised to pro-actively identify anomalies or fraud patterns of criminality which are indicative of fraud.
- Reporting & Distribution of Information software: Providing secure static and dynamic management reports for all units within the business and to our external stakeholders via a web portal platform.

8. Are there any current issues of public concern that you should factor in?

NHS fraud is, in itself, a matter of public concern and the use of personal data to combat it, particularly in ensuring that it is applied in an appropriate and proportional matter, are all issues of public concern.

9. Are you signed up to any approved code of conduct or certification scheme (once any are approved)?

NHSCFA codes of conduct are summarised in the following documents:

Information Governance:

- Information Governance Policy
- Information breach Reporting Policy
- Data Quality Policy
- GDPR – Data Protection Policy

Information Security:

- Information Security Policy
- NHSCFA Acceptable use Policy

The NHSCFA has an ISO ISO27001:2013 certification on information security which covers information processed within the NHSCFA network.

Describe the purposes of the processing:

1. What do you intend to achieve?

2. What is the intended effect on individuals?

3. What are the benefits of the processing, for you and more broadly?

1. What do you intend to achieve?

These would be specific to each project and the data sources being applied and therefore would defer to the DPIA and other data usage considerations being applied for these.

2. What is the intended effect on individuals?

These would be specific to each project and the data sources being applied and therefore would defer to the DPIA and other data usage considerations being applied for these.

3. What are the benefits of the processing, for you and more broadly?

The benefits of any processing would be specific to each project. However briefly summarised they would be in support of fraud detection activity or in support the organisation in meeting these aims.

STEP 3: Consultation process

Consider how to consult with relevant stakeholders:

1. Describe when and how you will seek individuals' views or justify why it's not appropriate to do so?

2. Who else do you need to involve within your organisation?

3. Do you need to ask your processors to assist?

4. Do you plan to consult information security experts or any other experts?

1. Describe when and how you will seek individuals' views or justify why it's not appropriate to do so?

This would be specific to each project and the data sources being applied and the nature of the data and who it pertains to. Informing stakeholders via the NHSCFA website or individual engagement opportunities or exercises therefore would defer to the individual project and the DPIA and other data usage considerations being applied for these will consider these options.

2. Who else do you need to involve within your organisation?

This would be specific to each project and the data sources being applied and the nature of the data and who it pertains to and how it is being applied.

3. Do you need to ask your processors to assist?

There are no identified circumstances in business as usual activity where non NHSCFA processors would directly assist in the handling and processing of data - the extent of external parties involvement would be limited to the preparation of data for sharing with NHSCFA prior to the receipt and processing of data and the sharing of results following the processing. In both cases, this activity would be specific to the project but would also constitute compliance with the high-end aims of that particular piece of work, and therefore would be a key part of the initial considerations when commencing this course of work. Therefore, this would primarily be specific to each project and the data sources being applied and the nature of the data and who acts as the data controller

4. Do you plan to consult information security experts or any other experts?

This would be specific to each project and the data sources being applied and the nature of the data and who it pertains to.

Describe compliance and proportionality measures, in particular:

1. What is your lawful basis for processing?

2. Does the processing actually achieve your purpose?

3. Is there another way to achieve the same outcome?

4. How will you prevent function creep?

5. How will you ensure data quality and data minimisation?

6. What information will you give individuals?

7. How will you help to support their rights?

8. What measures do you take to ensure processors comply?

9. How do you safeguard any international transfers?

1. What is your lawful basis for processing?

This would be primarily specific to each project and the data sources being applied and the nature of the data and who it pertains to and how it is being applied. However, the overarching basis for processing within the NHSCFA remit is linked to the provisions of Article 6 of the General Data Protection Regulations (GDPR), namely linked to “the processing (being)s necessary for you to perform a task in the public interest or for (an) official functions, and the task or function has a clear basis in law.

2. Does the processing actually achieve your purpose?

This would be specific to each project and the data sources being applied and the nature of the data and who it pertains to and how it is being applied.

3. Is there another way to achieve the same outcome?

This would be specific to each project and the data sources being applied and the nature of the data and who it pertains to and how it is being applied.

4. How will you prevent function creep?

Although this would be specific to each project and the data sources being applied and the nature of the data and who it pertains to and how it is being applied, the parameters of each data application would be linked to the objectives of each data source and/or application and this would be outlined in the appropriate documentation – whether it be linked. The NHSCFA has a range of oversight tools, ranging from Project Boards to the centralised Data Strategy Group, and supporting governance and assurance processes, which can manage and mitigate this risk (this is also reflected in key roles within the organisation in terms of SRO’s etc).

STEP 4: Assess necessity and proportionality

5. How will you ensure data quality and data minimisation?

This would be specific to each project and the data sources being applied and the nature of the data and who it pertains to and how it is being applied.

6. What information will you give individuals?

This would be specific to each project and the data sources being applied and the nature of the data and who it pertains to and how it is being applied.

7. How will you help to support their rights?

This would be specific to each project and the data sources being applied and the nature of the data and who it pertains to and how it is being applied.

8. What measures do you take to ensure processors comply?

This would be specific to each project and the data sources being applied and the nature of the data and who it pertains to and how it is being applied, however engagement and collaboration for stakeholders is a key element of the vast majority of NHSCFA work – as the centralised body for counter fraud activity, our activity is continually conjoined with additional stakeholders and this would include the provision of MoUs and ISA's that are very clear about the expectations for compliance and cooperative activity in terms of data provision and what is/isn't permitted.

9. How do you safeguard any international transfers?

Although this would be specific to each project and the data sources being applied and the nature of the data and who it pertains to and how it is being applied, it would be expected that the transfer of data across international borders is extremely unlikely given the mandate of NHSCFA and the service it provides exclusively to England and Wales.

STEP 5: Identify and assess risks

Describe source of risk and nature of potential impact on individuals. Include associated compliance and corporate risks as necessary	Likelihood of harm Remote, Possible or Probable	Severity of harm Minimal, Significant, or Severe	Overall risk Low, Medium or High
1. There is a risk that personal data will be used for purposes other than that which are stipulated in the business case	Possible	Significant	Medium
2. There is a risk that complex processing of data by participating authorities during the analysis process may lead to information being inadvertently disclosed.	Remote	Significant	Medium
3. There is a risk that data disclosed are not required for the purposes of fraud detection, or are excessive.	Remote	Minimal	Low
4. There is a risk that incorrect information may be disclosed by participating authorities during the pilot.	Possible	Significant	Medium
5. There is a risk that data is retained for longer than it is needed.	Remote	Minimal	Low
6. There is a risk that an individual's rights under GDPR are violated.	Remote	Significant	Medium
7. There is a risk that an external attacker gains access to personal data.	Remote	Significant	Medium
8. There is a risk that information could be lost, released or shared inappropriately	Remote	Significant	Medium
9. There is a risk that processing is carried out internationally in a territory without appropriate personal data protection in place	Remote	Minimal	Low
10. There is a risk that individuals will be misidentified as a result of data processing	Possible	Significant	Medium
11. There is a risk that the quality of data will not be to a consistently high standard	Possible	Significant	Medium

STEP 6: Identify measures to reduce risk

Identify additional measures you could take to reduce or eliminate risks identified as medium or high risk in step 5.	**Effect on risk** Eliminated, Reduced, Accepted	**Residual risk** Low, Medium, High	**Measure approved by SMT Owner** Yes/No
(1) There is a risk that sensitive data will be used for purposes other than that which are stipulated in the business case	(1), (2), (3) Robust governance structure and project management techniques are used to make clear roles and responsibilities, timelines, data flows, and contingency plans. The data itself has been minimised to remove personal data from the structured data, ensure that all possible steps have been taken to prevent inclusion of personal data (beyond that included in unstructured data)	Low	Yes - TM
(2) There is a risk that complex processing of data by participating authorities may lead to information being inadvertently disclosed.		Low	Yes - TM
(3) There is a risk that incorrect information may be disclosed by participating authorities during the pilot (leading to incorrect identification of fraud)		Low	Yes - TM
(4) There is a risk that incorrect information may be disclosed to unauthorised parties	(4) Data quality review is undertaken by authorities sharing data to review individuals / organisations included and remove irrelevant ones from the matching process.	Low	Yes - TM
(5) There is a risk that data is retained for longer than it is needed	(5) Each data source is subject to a retention schedule	Low	Yes - TM
(6) There is a risk that individual's rights under GDPR are violated	(6) This risk is not considered to be increased beyond business-as-usual levels as a result of the original NHSCFA activity and is mandated and controlled by the existing procedures	Low	Yes - TM

(7) There is a risk that an external attacker gains access to personal data	(7) NHS CFA have approved and assured methods of managing data and for information security and are ISO27000 compliant. The data analytical platform itself is password protected and secured on the Cloud.	Low	Yes - TM
(8) There is a risk that information could be lost, released or shared inappropriately	(8) This risk can be mitigated with robust governance structures, as well as security accreditation and adherence to a common set of data standards, set out in the security statement and information sharing agreement.	Low	Yes - TM
9. There is a risk that processing is carried out internationally in a territory without appropriate personal data protection in place	(9) Data is stored on UK based cloud storage and does not leave national borders as any part of the storage or data processing	Low	Yes - TM
10. There is a risk that individuals will be misidentified as a result of data processing	(10) Any decisions made using matched data will be informed by the strength of the match and resulting outlier. The strength of all matches will be assessed to manage false positives and triage follow up investigations.	Low	Yes - TM
(11) There is a risk that the quality of data will not be to a consistently high standard	(11) Data quality reviews are routinely undertaken as part of the analytical process and the impact of these findings will impact on the outputs produced. In terms of outlier detection, low data quality can sometimes be a useful factor in determining fraud risks	Low	Yes - TM

STEP 7: Sign off and record outcomes

Item	Name/date	Notes
Measures approved by SMT Owner:	Tricia Morrison , 5th July 2022	Confirmed happy to authorise
Residual risks Approved by SMT Owner:	Tricia Morrison , 5th July 2022	Confirmed happy to authorise
DPO advice provided		DPO should advise on compliance, step 6 measures and whether processing can proceed.
Summary of DPO advice: Having reviewed the DPIA I am satisfied that a comprehensive review and assessment of the Data Analytic Platform has been undertaken. It focuses primarily on how it stores, processes, analyses and outputs from NHSCFA data sources as opposed to the platform itself. Individual DPIAs exists for the specific data sources themselves and therefore this DPIA is completed to consider the system itself and the manner in which it processes data generally. The software platform is primarily accessed by approximately 12 members of staff which includes database administrators and is fully auditable. Outputs and wider data disseminations where required will be specific to individual projects and their designated outcomes. While the platform has direct interconnections with other NHSCFA systems and applications, in terms of data drawn in from other sources; each of these have their own access and control measures in place to mitigate any risk of unauthorised access. I am therefore satisfied that any data stored and processed within the analytic platform will be in accordance with current legislative requirements and handled pursuant to organisational best practice, its retention policy and the security measures employed.		
DPO advice accepted or overruled by:	Trevor Duplessis - 14th April 2023	If overruled, you must explain your reasons
Comments:		
Consultation responses reviewed by:		If your decision departs from individuals' view, you must explain your reasons
Comments:		
This DPIA will be kept under review by the Information and Records Management Officer:		The DPO should also review ongoing compliance with DPIA

Ownership

The following table describes the roles and responsibilities

Table 1 - Roles and Responsibilities

Role	Responsibility
Information Asset Owner (IAO)	Performance, Analytics & Improvement Manager.
Senior Information Risk Owner (SIRO)	Head of Intelligence and Fraud Prevention
Application/Database Owner	Data Analytics Lead
Data Protection Officer	Trevor Duplessis Information Governance and Risk Management Lead

DPIA Report

Section 1: Data Maintenance and Protection Overview

1. The impact level of the NHSCFA Data Analytics Platform was assessed as OFFICIAL.
2. The following measures briefly describe what controls have been implemented to protect the Data Analytics Platform and the personal data recorded:
    a. The database/system is primarily accessed by approximately 12 members of staff from NHSCFA, which includes the database administrators. However, outputs and wider data disseminations are specific to individual projects and their designated outcomes

    b. The database/system has direct interconnections with other NHSCFA systems and applications, particularly in terms of data drawn in from other NHSCFA data sources. Each of these have their own access control measures and controls in place to mitigate any risk of unauthorised access. The nature of each would be specific to each project and their intended data usage

    c. The Data Custodian must comply with the data protection requirements. Examples include: regularly reviewing the business requirement to record the personal data; ensuring that the data is not excessive; it is being used for the purpose intended; that there is a deletion and disposal policy; that the application is registered on the NHSCFA register and the NHSCFA DPO is aware of its existence.
3. It is assessed that there are no residual privacy risks to the personal data used by the Data Analytics Platform which cannot be addressed through the management and oversight related to individual projects
4. This DPIA must be reviewed if any changes are made to the personal information if used by the database or any other changes are made that affect the privacy of an individual.

Section 2: Uses of the Application and the Data

5. Administration of the Data Analytics Platform is managed by the Database Administrators, currently within the Information Systems Team. Requests for changes / amendments to the system are managed by the NHSCFA Service Desk who are the first point of contact and record and manage requests.

6. Information in the Data Analytics Platform would be project and purpose driven - this covers a wide range of data sources too numerous to name. However, there are a number of key data systems and data owners that are internal and external to the NHCFA and would be considered fundamental to expected / business as usual activities for the software. These are

Internal:

- NHSCFA Case Management System (CLUE), as well as legacy systems (FIRST)
- NHSCFA Intelligence Database (IBase)
- Performance Reporting Database (Verto) as well as legacy systems (MRT)
- The LCFS/DoF nominations database (CPOD)

External

- The NHS provider and commissioning sectors via the NHSCFA Data Capture System (DCS) as calibrated for specific data capture exercises
- The NHSCFA Fraud and Corruption Reporting Tool (FCROL)
- NHSBSA Dental and Pharmaceutical services (project specific)
- HM Cabinet Office via National Fraud Initiative (NFI) data (project specific)

7. Some of the above data sources capture sensitive data. The individual DPIA's for each would be more specific in outlining what these are. It is not necessarily the case that the information captured on these systems will be transferred or otherwise processed by the Data Analytics Platform, although this may be the case for individual

8. The IAO must comply with data protection requirements. Examples include: regularly reviewing the business requirement to use the personal data; ensuring that the data is not excessive, it is being used for the purpose intended; that there is a deleting and disposal policy; that the application is registered and the DPO is aware of its existence.

Section 3: Data Retention

9. The Data Analytics Platform is subject to NHSCFA Data Handling and Storage Policy by virtue as to how this is applied to the individual data sources that it processes. All records are electronic and there are no paper based records produced by this system.

10. The IAO is required to review the retention period and any requirement to change must be submitted to the Senior Information Risk Owner.

Section 4: Internal Sharing and Disclosure of Data

11. Because of the nature of the Data Analytics Platform in producing and disseminating data, the Database/System is theoretically accessible by all members of staff from NHSCFA, however this is not to say that all the applications of outputs will be utilised in this widespread manner and more targeted controls (and alternative mediums) are available to produce and disseminate data that requires restricted audiences.

Section 5: External Sharing and Disclosure of Data

12. The only information shared directly with external organisations would be a) through controlled sharing of findings via formal reports and b) via the dissemination element of the analytical platform which would designed specifically to produce reports for external users.

Ultimately, individual projects which MAY have this element would require their own considerations, as would the data sources that capture these data in the first place. The nature of this dissemination is specific to each project, and it is extremely unlikely any would include personal data. However, the disclosure of data via electronic reporting can be controlled through rigorous security settings, linked to the NHSCFA nomination process and database.

Individual outliers would be shared in specific circumstances, if necessary for the administration of justice and/or in accordance with appropriate Information Sharing Agreement/Memorandum of Understanding.

Section 6: Notice/Signage

13. NHSCFA's privacy policy on its website hosts separate sections in relation to data collection, retention and storage. This broadly covers all elements of the NHSCFA usage of data, in a nonspecific manner.

14. The use of signage or other notifications to notify the public of the gathering and use of personal data is not relevant to this Database/System and therefore outside the scope of this DPIA.

Section 7: Rights of Individuals to Access, Redress and Correct Data

15. Individuals subject to certain exemptions, have the right to gain access to their own personal data. In the event an access request is directly or indirectly received by NHSCFA, we are required to provide the individual who has made the request with details of the personal data recorded about them.

16. It is unlikely that many access requests will be received as the personal data recorded is all in relation to individual data sources and/or their outputs, as opposed to the platform itself which acts as intermediary.

17. In the unlikely event that information is identified as being incorrect, NHSCFA staff will take appropriate steps to correct the record where permissible.

18. All NHS employees and member of the public have the right to access, redress and correct personal data recorded about them.

Section 8: Technical Access and Security

19. The security and technical access architecture of the Database/System is as explained in this DPIA. The application and the hosting infrastructure was assessed at **Official-Sensitive** and the hosting infrastructure is subject to the ISO27000 and ISO27001

20. Access to the system itself, with the exception of disseminated products, is restricted to internal staff only.

21. The technical controls to protect the database include:

a. Anti-virus protection;

b. Permission based access controls to shared drive.

c. Logging, audit and monitoring controls.

d. Vulnerability Patching Policy for the underlying infrastructure.

Section 9: Technology

22. The Database/System holds personal information obtained electronically and is located in the NHS Counter Fraud Authority cloud infrastructure, subject to the principles of use outlined in the following NHSCFA policies

- Information Governance Policy
- Information breach Reporting Policy
- Data Quality Policy
- GDPR – Data Protection Policy
- Information Security Policy
- NHSCFA Acceptable use Policy

DPA 2018 Compliance Check

1. The DPO must ensure that the System, and the personal data that it records, and its business activities, are compliant and maintain compliance with:

a. The GDPR and the Data Protection Act in general;

b. The Data Protection Principles;

c. The interpretations of the Principles.

2. **This is not a recommendation but a requirement of law.**

3. The roles and responsibilities for the protection of personal data are described in the NHSCFA security policy.

4. The Database/System processes sensitive personal data so a Data Protection Compliance Check Sheet has been completed (see Annex B) describing how the requirements of GDPR and the Data Protection Act 2018 have been complied with.

The Privacy and Electronic Communications Regulations

5. The Privacy and Electronic Communications Regulations is not applicable as personal data is not exchanged with external organisations for commercial purposes.

The Human Rights Act

6. The decisions and activities of the organisation are undertaken in compliance with the Human Rights Act, having due regard to appropriateness and proportionality to ensure compatibility with Convention rights.

The Freedom of Information Act

7. As a public authority we are compliant with the provisions of the Freedom of Information Act, in proactively publishing and making available upon request, certain recorded information held by the organisation subject to any relevant exemption(s). However, no personal information would be disclosed under the Freedom of Information Act as this would breach the data protection principles.

Conclusion

8. There are no residual privacy risks to the personal data recorded in the Database/System that are not inherent to the individual projects or data sources that make use of them. The controls described in this DPIA explain in detail how the data is protected and managed in accordance with the GDPR and Data Protection Act 2018. The DPO is responsible for ensuring that the controls are implemented through the life cycle of the system.

Annex A - Definition of Protected Personal Data

Personal data includes all data falling into Categories A, B or C below:-

A. Information that can be used to identify a living person, including:

Name;
Address;
Date of birth;
Telephone number;
Photograph, etc.

Note: this is not an exhaustive list.

B. Information which, if subject to unauthorised release, could cause harm or distress to an individual, including:

Financial details e.g. bank account or credit card details;
National Insurance number;
Passport number;
Tax, benefit or pension records;
DNA or fingerprints;
Travel details (for example, at immigration control or oyster records);
Place of work;
School attendance/records;
Material related to social services (including child protection) or housing casework.

Note: this is not an exhaustive list.

C. Sensitive personal data relating to an identifiable living individual, consisting of:

Racial or ethnic origin;
Political opinions;
Religious or other beliefs;
Trade union membership;
Physical or mental health or condition;
Sexual life
Commission or alleged commission of offences;
Proceedings relating to an actual or alleged offence.

Any data set containing this information must be processed in accordance with the GDPR and the Data Protection Act 2018 (DPA 2018).

Particular care must be taken with data in Category B and with any large data set. Information on smaller numbers of individuals may justify additional protection because of the nature of the individuals, source of the information, or extent of information.

There are additional, specific constraints within the provisions of GDPR and the DPA (2018) on the processing of data in Category C.

Annex B - Data Protection Compliance Check Sheet

PART 1: BASIC INFORMATION - New or existing Project, System, Technology or Legislation

1. Organisation and project.

Organisation	NHSCFA
Branch / Division	Information Analytics
Project	Data Analytics Platform

2. Contact position and/or name.
(This should be the name of the individual most qualified to respond to questions regarding the DPIA)

Name, Title	Trevor Duplessis
Branch / Division	Finance and Corporate Governance, NHSCFA

3. Description of the programme / system / technology / legislation (initiative) being assessed.
(Please note here if the initiative does not collect, use or disclose personal data*). If this is a change to an existing project, system, technology or legislation, describe the current system or program and the proposed changes.

The purpose of the system is for collecting, collating, managing, analysing and disseminating information in support of all internal business units and externally to the NHS and DH. It is also used to manage the data and the analysis tools used to identify patterns and anomalies that may indicate fraudulent behaviour.

The resulting analyses can be used in a number of project specific ways. Detected outliers may be used to support on-going investigations as well as support the creation of intelligence that can prompt new investigations. The findings can also drive changes to policy and guidelines in order to address fraud risks. Data can additionally be used to provide electronic Management-style information reports, in some cases linking to other data sources or being adapted to new formats (i.e. graphical representation via graphs or adapted to show weighting or ratios for benchmarking purposes)

Because the scope for the types of data that the system can process is universal, and the wide-ranging nature of data types that can be processed by it for individual projects and data exercises (which may include personal data), it has been determined that a specific DPIA should be produced to cover this processing specifically, as a separate review to the original datasets and/or projects that this activity supports

4. Purpose / objectives of the initiative (if statutory, provide citation).

NHSCFA leads on a wide range of work to protect NHS staff from economic crime and the Data Analytics Platform facilitate this work through the provision of data services. This can take a variety of shapes and purposes (as described above), which in turn is defined by the project itself and its objectives. The system itself is universal in acting as the conduit for the data source(s) that need to be utilised – this can include any data source owned or accessed by the NHSCFA for their counter fraud purposes .

Access of the main system is restricted to approximately 12 members of staff within NHSCFA, including the database administrators. The wider data products are available organisationally

5. What are the potential privacy impacts of this proposal?

Data Protection Impact Assessments (DPIA) have been considered in the light of the potential for any personal data that might be gathered and utilised by the system. In most cases, specific DPIA considerations relate to the data source itself and/or the project utilising it. However, this DPIA considers specifically the role of the Data Analytics Platform in applying it.

As such, the data in the Data Analytics Platform has been gathered for a specific, justifiable and proportional purpose (as outlined by the data source and/or the data project) and found to be mitigated by the steps put in place to minimise the possibility of unauthorised access or use (see Compliance Checks in section 3 of this document)

6. Provide details of any previous DPIA or other form of personal data* assessment done on this initiative (in whole or in part).

This is the first DPIA carried out on the system.

IF THERE IS NO PERSONAL DATA INVOLVED, GO TO PART 3: DPA COMPLIANCE - CONCLUSIONS

*IMPORTANT NOTE:
‘Personal data’ means data which relate to a living individual who can be identified:
(a) from those data, or

(b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller,

And includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual.

Page left intentionally blank.

NHSCFA offices

Coventry

9th Floor
Earlsdon Park
55 Butts Road
Coventry
West Midlands
CV1 3BH

London

HM Government Hub
10 South Colonnade
Canary Wharf
London
E14 4PU

Newcastle

1st Floor
Citygate
Gallowgate
Newcastle upon Tyne
NE1 4WH