Data Protection Impact Assessment

NHS COVID-19 Data Store

Template Version number: 5a

Published Date:

Prepared by: Corporate Information Governance

Classification: OFFICIAL

Data Protection Impact Assessment
NHS COVID-19 Data Store

This is a technical document for those interested in the detailed arrangements for the NHS COVID-19 Data Store. Further information including FAQs are available here. If you have any queries in relation to this document please contact england.ig-corporate@nhs.net

Administrative information

NHS England	
Your name	Ming Tang
Your team and directorate	Director of Data and Analytics
Your location	Quarry House, Leeds, W. Yorkshire
Your telephone number	
Your email address	

NHS Improvement	
Your name	
Your team and directorate	
Your location	
Your telephone number	
Your email address	

Purposes - Fully describe what is the purpose of the project and how is the processing of information necessary to that work?

Context

NHSx along with NHS England and Improvement are providing a national response to the COVID-19 pandemic. We are working with companies, such as Microsoft, Faculty, McKinsey under strict contractual controls, to ensure data can be used effectively to support the national response to the COVID19 virus.

Purpose

The purpose of the processing is to bring together all data necessary to provide analysts, the government and researchers with the most comprehensive datasets related to COVID-19.

It is proposed that NHS England will utilise cloud infrastructure with secure data storage and analysis tools to ensure that the health and care service is able to maximise the usefulness of the data and future proof the database. This also aligns to the 'cloud first' policy for public sector IT introduced in 2013, endorsed by the National Information Board's Personalised Health and Care 2020 framework.

The privacy protected database known as the NHS COVID-19 Data Store will be a SQL Server database hosted under an NHS England and Improvement operated Microsoft Azure subscription. The privacy protected data hub will contain several data marts supporting the analysis to support the analytical models.

The work will build a framework that will allow for analytical models to be created with forecasting, trend analysis and data profiling to support healthcare needs as well as spread of the virus. Using national and local data sources, will facilitate a range of novel and actionable national population-level insights, using a standardised and robust approach.

Type of Data Marts

The following is a list of subsets of data within the Data Store:

1. Record level Secondary Uses Service (SUS) data extracts including data in aggregated form from emergency care, inpatient and outpatient activity data sets.
2. SITREP data collected within the Strategic Information Platform aggregated to support the COVID-19 response.
3. Patient-Level Information and Costing Systems (PLICS)
4. Costing data mart
5. UK Health Facts published data
6. North of England Commissioning Support (NECS) community and bed utilisation data mart
7. Emergency care, Outpatient, Inpatient, A&E record level data sets
8. Daily COVID SITREP v1
9. COVID SITREP v2
10. A&E sitrep
11. Ambulance SITREP
12. Record level and aggregated 111 data
13. 111 online screening data
14. 111 telephony data
15. Reference data = UK Health Dimensions, UK Health Facts, Org Hierarchy file
16. Master Patient Index (MPI) with frailty flag
17. Spec Comm (Specialised Commissioning) Data
18. PHE diagnostic COVID-19 test data
19. ESR workforce data
20. Care home bed availability – partial coverage
21. East Midlands home ventilation data
22. Mental health
23. 111 data (National Commissioning Data Repository)
24. COVID-19 Hospitalisation in England Surveillance System (CHESS)
25. Ventilator Orders
26. Oxygen gas supply and capacity

27. Deprivation data - 2019
28. Births by Clinical Commissioning Group (CCG)
29. Live bed capacity
30. Primary care data
31. Ambulance capacity
32. International data daily dump of cases (infected, recovered, fatalities)
33. Supply chain data
34. Self-isolating information
35. Other workforce data

Please note, the above list is not exhaustive as more data is being added to support the COVID-19 response. The following link contains the most up-to- date list of data: https://data.england.nhs.uk/covid-19/

Process

The data marts will be sourced from different data sources including Public Health England, NHS Digital, NHSE's National Commissioning Database Repository (NCDR), NHSE's North of England Commissioning Support Unit, published data on the Web, NHS Improvement's Strategic Information Platform. This will then be:

a) uploaded to the NHS COVID-19 Data Store
b) In its raw format the data will be stored in a secure 'blob' storage. A copy of the data will be made before it is made available. This will ensure that small number data is suppressed, and quality assurance is done.
c) Data marts will be created to meet analytical requirements while minimising the size and identifiability of the data. The reporting data marts will be pseudonymised before being made accessible to registered and approved users.
d) For the system interface there will be a requirement for automation and API updates. This data will be pulled through to the secured area before aggregation.
e) At no point during the processing will the data be made available in identifiable format to the user or the data technician.

Access

A limited number of individuals from NHS England's Data Services team will have access to the raw data. This is to process the data through an ETL (Extract, Transform, Load process) into a database, for continued maintenance of this database and to build the business intelligence analyses / dashboards.

External data processors Palantir using their Foundry platform, have been engaged under contract with NHS England and will have access to the data which is aggregated to required level or data which has been de-identified to mitigate the risk of identification of the individual in the data mart.

Permissions will be granted via Office365 user accounts. This is the standard methodology on NHS England's cloud infrastructure for the applications which are accessing the data through the database connections or API (Application Programming Interface). The Azure Bastion will control the access to the analytical tooling Virtual Machines.

An admin account will be temporarily created for NCC Group (Cyber Security Experts) and Pivotal, the third-party supplier contracted to deliver the COVID-19 app and Penetration Testing. They will have read only access and once the application has been delivered, this account and access will be terminated. Following this, an SQL account will be created which the application will use to read/write the database for certain task(s). This will restrict unwarranted access.

Nature of the data

Will the processing involve anonymised information[1]?	Yes
Will the processing involve pseudonymised personal data?	Yes
Will the processing involve fully identifiable personal data?	Yes – but only in an area with restricted access to a small number of NHSE/I staff for pre-processing and validation before data goes into the NHS COVID-19 Data Store

Assets

Does the proposal involve creating a new information asset?	Yes
Does the proposal involve processing data held on an existing information asset or assets?	No
Is/are the asset owner(s) aware of the proposal	

What is the timeframe for the project/programme/initiative? (Please include commencement dates and any foreseen end dates)

15th March 2020 – to agreed Control of Patient Information (COPI) Notice End Date, currently 30th September 2020

[1] anonymous information is information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable

Controllers[2]

NHS England	Yes - NHS England is the sole controller of data held within the NHS COVID-19 Data Store and for any data legally shared with them under the NHS Control of Patient Information (COPI) Notice issued by the Secretary of State.
TDA	No
Monitor	No
NHS Digital	No
Other (Please do not include any third party that we are contracting with to process personal data for us as a processor.)	

Screening questions

Does the proposal involve any of the following – drop down list to include: • NCDR • Pseudonymised by NHS Digital • Aggregate data • Anonymised data	Yes
Has processing of this nature already been captured and considered within a previous DPIA? If so, link to reference number	No
Will the processing involve a large amount of personal data (including pseudonymised personal data) and affect a large number of data subjects?	Yes
Will the project involve the use of a new technology(ies) which might be perceived as being privacy intrusive? i.e. using biometrics, facial recognition, Artificial Intelligence or tracking (such as tracking an individual's geolocation or behaviour)?	No – but it is expected that new algorithms will be created to support targeted analysis
Will the processing introduce or make use of a new platform not currently in use?	Yes
In the absence of proper controls is there the risk that the processing may	Yes but only in the absence of proper controls

[2] 'controller' means NHSE, alone or jointly with others, determines the purposes and means of the processing of personal data

give rise to discrimination, identity theft or fraud, financial loss, damage to the reputation, loss of confidentiality of personal data protected by professional secrecy (e.g. health records), unauthorised reversal of pseudonymisation, or any other significant economic or social disadvantage?	
Does the proposal introduce difficulties in ensuring that individuals are informed or able to exercise their information rights?	No
Will there be processing of genetic data, data concerning health, sex life, racial or ethnic origin, biometric data, political opinions, religion or philosophical beliefs, or trade union membership?	Yes
Will there be processing of data concerning criminal convictions and offences or related security measures?	No
Will the project involve the targeting of children or other vulnerable individuals for marketing purposes, profiling or other automated decision making?	No
Will the processing result in you making decisions or taking actions against individuals in ways which can have a significant impact on them? e.g. decisions about an individual's access to a product, service, opportunity or benefit, or recruitment aptitude test based on automated decision making (including profiling)?	No
Will there be a systematic monitoring of a publicly accessible area on a large scale (e.g. CCTV)?	No
Will the processing include any data matching e.g. the combining, comparing or linking of personal data obtained from multiple sources?	Yes
Will personal data about individuals be shared with other organisations or people who have not previously had routine access to the data?	No
Will the project/proposal use personal data about individuals for a purpose it is not currently used for or in a new way?	Yes
Will the project require you to contact individuals in ways which they may find	No

intrusive? i.e. telephoning or emailing them without their prior consent.	
Are you using a Data Processor/third party supplier or is a service/processing activity being transferred to a new supplier/organisation (or re-contracted) at the end of an existing contract?	Yes

NB. If the answer to any of the above questions is Y, please complete the rest of the form. If all of the screening questions are answered N, the local IG team must still sign off the DPIA.

Where the information will include the processing of personal data, please continue.

Personal data[3]

Why would it not be possible to do without personal data?	The current situation relating to patients affected by COVID-19 must be monitored and effectively managed. All data required for processing is assessed before it is onboarded to ensure that the minimum amount required for COVID-19 purposes is collected in line with GDPR requirements. Without the initial processing of identifiable data (which will then be de-identified) it will not be possible to create the appropriate dashboards. The COPI notice enables NHS England to collect and process any data required to support the COVID-19 response.
What are the required personal data? Please itemise them or supply a dummy sample, blank forms, screenshots from the prototype system etc.	The list of datasets is made publicly available via the following link https://data.england.nhs.uk/covid-19/
Please confirm that this is the minimum amount of personal data that is necessary.	Yes
Would it be possible for NHS England to use pseudonymised personal data for any element of the processing?	Yes, once the data has been validated against the NHS England Master Patient Index it will be pseudonymised before being linked to other datasets held within the data hub. Data marts

[3] 'personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person

	will then be created and made available as per user requirements. Two dashboards will be created within the Data Store. The COVID-19 operational dashboard will display aggregated (small number suppressed) and anonymised data to approved users. The strategic dashboard will display aggregate and record level pseudonymised data however this dashboard is further restricted to specific users with clear justification to require access to data at that level. Users are prohibited from extracting data from the dashboards unless specific approval is given by NHS England.
If Y, please specify the element(s) and describe the pseudonymisation technique(s) that we are proposing to use.	Pseudonymisation will be undertaken as per NHSE/I processes. Once the PHE data has been pseudonymised, the key will be provided to AGEM Data Services for Commissioners Regional Offices DSCRO in order that they can then de-identify any other required datasets for linkage purposes.

Scale and constituency(ies)

What is the scale of the processing (i.e. (approximately) how many people will be the subject of the processing)?	Large Scale - national
Please describe the constituency(ies).	National

Outcomes

What will be the effects of the processing (i.e. what actions/decisions will result from the processing)?	Data is a critical component of the Covid-19 response, and that the processing of this data is necessary to support all organisations involved in managing and supporting health services. With access to the output dashboards, the government and health service, will be able to effectively monitor the spread of COVID-19 and implement appropriate measures to ensure services and support is available to patients.

Purpose(s) and legal basis(es) of the processing (complete only if sole controller)

(Please tick all that apply.)	
Is the processing necessary for a task that is within NHSE's remit as a public authority? (please specify below)	Yes
(This is applicable for much of NHS England's processing of personal data using its statutory powers) GDPR Article 6(1)(e)	
Is NHSE under a legal obligation to carry out the processing? (please specify below)	Yes – The Secretary of State for Health and Social Care has issued NHS England/NHS Improvement with a notice under the Control of Patient Information Regulations Regulation 2002 4(3) of which requires the organisations to process data to support the response to COVID-19
GDPR Article 6(1)(c)	

Special categories of personal data

Will the processing involve personal data about: (Please tick all that apply.)	
• racial or ethnic origin	No
• political opinions	No
• religious or philosophical beliefs	No
• trade union membership	No
• genetic data[4]	No
• biometric data[5]	No
• data concerning health[6]	Yes

[4] 'genetic data' means personal data relating to the inherited or acquired genetic characteristics of a natural person which give unique information about the physiology or the health of that natural person and which result, in particular, from an analysis of a biological sample from the natural person in question

[5] 'biometric data' means personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data

[6] 'data concerning health' means personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status

- data concerning the sex life or sexual orientation of the data subjects	No

If there are no special categories of data processed, please skip the following section and proceed to the 'Common law duty of confidentiality' section…

Legal basis(es) for special category personal data

Legal basis	Personal data to which this legal basis relates:
- necessary for health or social care purposes	GDPR Article 9(2)(h)
- necessary for public health	GDPR Article 9(2)(i)

Common law duty of confidentiality

Are any of the data subject to a duty of confidentiality (e.g. clinical records, OH details, payroll information)? If so, please specify them.	Yes	
Where it is planned to disclose such data, what are the grounds for doing so?	- consent - safeguarding - other overriding public interest - please specify - legal duty or permissive power e.g. s251 support – please specify (e.g. court order)	The Secretary of State for Health and Social Care has issued a notice under the COPI Regulations 2002 which require NHS England to process Confidential Patient Information for COVID-19 purposes.
If the processing is of data concerning health or social care, is it for a purpose other than direct care[7]?	Y	

[7] direct care: a clinical, social or public health activity concerned with the prevention, investigation and treatment of illness and the alleviation of suffering of individuals. It includes supporting individuals' ability to

Consultation

Would it be appropriate to seek the views of data subjects or their representatives on the proposed processing?	No
If Y, how will this be done?	
If N, why is this the case?	Various reasons, including the fact that this is large scale processing which needs to be facilitated very quickly to support the emergency response for COVID-19
Would it be helpful to seek advice from independent experts (clinicians, security experts, ethicists etc.) where their specialist knowledge would be useful in understanding and managing privacy risks?	Subject matter experts are involved in ensuring that the processing meets safe, efficient and effective standards.
If Y, how will this be done?	
Will any other stakeholder(s) (whether internal or external) need to be consulted about the proposed processing (e.g. NHSE Central team, Public Health England, NHS Digital, the Office for National Statistics)?	Yes – NHSx are leading on the management of the NHS COVID-19 Data Store and will implement a triage process which will continually monitor what data is being processed, the access controls and controls to mitigate the risk of re-identification.
What was/were the outcomes(s) of such consultation?	

Datasets and access

The datasets available are listed in the following link

https://data.england.nhs.uk/covid-19/.

Access to all data or the dashboards is managed by the front door triage process attended by representatives from NHSE/I, PHE, NHSx, NHSD, DHSC.

Purpose / process	**Required data items**	**Accessed by (Roles)**	**Storage location**

function and improve their participation in life and society. It includes the assurance of safe and high quality care and treatment through local audit, the management of untoward or adverse incidents, person satisfaction including measurement of outcomes undertaken by one or more registered and regulated health or social care professionals and their team with whom the individual has a legitimate relationship for their care.

Data processor[8]

Will the processing be wholly or partly performed on our behalf by a data processor(s)?	Yes
If Y please give details	Palantir
Where is the data to be processed by the data processor?	• In UK • Outside UK but in EEA

If the processing is not completed by a data processor, please ignore the following questions and proceed to the 'Collection of personal data' section …

What assurance has been/will be sought about the/each processor's compliance with the GDPR?	Contractual arrangements confirm GDPR assurance statement and provides necessary safeguards
Will the contract use NHS England's standard data processing agreement template?	No – A G-cloud contract has been put in place
Will the contract contain standard clauses to require compliance with the GDPR?	Yes
Will the contract contain clauses to address the secure transfer of the personal data to a successor data processor should this become necessary or upon the expiry of the term?	Yes

Collection of personal data

Will personal data be collected from the data subject?	No
Will personal data be obtained from sources other than the subject?	Yes
Will personal data be collected from a third party(ies)?	Yes

[8] 'processor' means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller

If Y, please identify the third party(ies)?	All sources of data are listed in the data dictionary
Is the provision of personal data obligatory or voluntary?	Obligatory
If obligatory, why/how is that the case?	The COPI notice issued by the Secretary of State for Health and Social Care requires that data is NHS England/NHS Improvement processes data to support the response to COVID-19 – this is a National pandemic – processing is being undertaken in the public interest to protect citizens
What are the possible consequences for a data subject if there is a failure to provide the requested personal data?	None

Privacy information

How will the data subjects be informed of the processing of personal data about them?	NHS England/NHS Improvement will update their privacy notices to inform patients that this processing is taking place. NHSx will also publish in depth information about the NHS COVID-19 Data Store with FAQs

Accuracy of personal data

How will we ensure the accuracy of the personal data (including their rectification or erasure where necessary)?	The Data Store is being operated by NHSE/I teams who already manage NHSE/I warehouses, which have existing very mature data validation and cleansing scripts, built up through experience with the data and supported by audits by NHS Digital. Recent experience has suggested that the quality of a national dataset is of higher quality than the sum of the quality of regional datasets, due to duplicate and unrecognised missing record.
How will we monitor the quality of the personal data?	Through data validation and data quality reports, automatically generated every time data is refreshed

Subject access and data subjects' rights

How will it be possible to provide a copy of the personal data processed about a particular individual to them (redacted as necessary) should they request access to this information? (If you are purchasing an information management system, you should consider including requirements in the specification about searching and subject access requests.)	Citizens can make a Subject Rights Request in line with NHS England/Improvement policy – england.dpo@nhs.net
What processes will be put in place to ensure that other data subjects rights can be appropriately applied to the personal data if necessary?	NHS England/NHS Improvement policy will be followed.

Data sharing (other than between NHSE and NHSI)

Will some or all of the personal data be shared with a third party (other than NHSE / NHSI)	No confidential personal data will be shared unless there is a specific supporting legal basis to do so i.e. where another organisation in receipt of a Control of Patient Information Notice issued by the Secretary of State. NHSE/I can also share data using their powers under the Health and Social Care Act 2012 s13Z3.
If Y, will the personal data be disclosed to a recipient(s) in a country outside the EEA or an international organisation?	

Risks

What are the identified risks of the processing?

Risk Assessment for Data Store.xlsx

Incident reporting

What plans are in place in relation to the internal reporting of a personal data breach? (NB Unless the personal data breach is unlikely to result in a risk to the rights and freedoms of the individual(s), it will normally need to be reported to the ICO within 72 hours.)	NHSE / NHSI incident reporting procedure NHS England have a standard process and are working with analysts to set up a process whereby they will log incidents of inadvertent re-identification to ensure lessons are learned. Given the current controls the risk of re-identification remains unlikely without unreasonable effort.
What plans are in place in relation to the notification of data subjects should there be a personal data breach? (NB Where a personal data breach is likely to result in a high risk to the rights and freedoms of the individual(s), they should be notified as soon as reasonably feasible and provided with any recommendations to mitigate potential adverse effects.)	NHSE / NHSI incident reporting procedure

Business continuity planning

How will the personal data be restored in a timely manner in the event of a physical or technical incident?	Network backup and recovery in line with organisational BCP plans.

Records Management

Will corporate records be created and / or managed as a result of this processing?	No
Where will these records be stored?	
Is there a trained Records and Information Management Coordinator (RIMC) responsible for these records?	

Retention of personal data

What is/are the retention period(s) for the personal data?	Retention period will be in line with COPI notice issued by the Secretary of State for Health and Social Care which currently ends on the 20th September 2020.
What is the basis for this retention period? (Please indicate applicable guidance or rationale)	Where there is a requirement to retain records for longer, the data will be retained in line with NHSE/I business requirements and the NHS Records Management Code of Practice.
Where personal data are processed outside of NHSE's premises or systems, how will they be securely returned to NHSE for the remainder of the retention period(s) as and when this becomes necessary (e.g. following the closure of the project)?	A G-cloud Data Processing Agreement (DPA) has been implemented between NHS England (AGEM CSU) and Palantir. This DPA includes GDPR compliant clauses and a requirement that data is either destroyed or returned to NHS England/NHS Improvement at the end of the contract. The risks assessed are limited to the four changes to the way data is managed in a cloud environment: • Access Authentication – this manages user-ids and passwords across multiple network domains. User information is shared in the cloud – this supports the strategy of Single Sign-On. • Access Authorisation – this manages permissions to access resources accessible via the cloud environment (e.g. access to database objects). Authentication will take place in the cloud, meaning user information is shared in the cloud. • Data-in-transit – this relates to the movement of patient-level data from on-premises into a cloud environment

	• Data-at-rest – this relates to the storage of patient level data in a cloud environment In all other respects, the controls already in place still apply and are relevant.

Direct marketing[9]

Will any personal data be processed for direct marketing purposes?	No
If Y, please describe how the proposed direct marketing will take place:	

Data portability

Where the processing is based on consent or due to a contract, it is carried out by automated means and the data subject has provided the personal data to us, will it be possible to provide them or a different controller with the personal data in a structured, commonly used and machine-readable format? (NB This does not apply to processing necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller – GDPR Article 6(1)(e).)	N/A

Automated processing

Will the processing result in a decision being made about the data subject solely on the basis of automated processing[10] (including profiling[11])?	No
If Y, is the decision:	

[9] direct marketing is "the communication (by whatever means) of any advertising or marketing material which is directed to particular individuals" - all promotional material falls within this definition, including material promoting the aims of not-for-profit organisations

[10] examples include the automatic refusal of an online credit application and e-recruiting practices without any human intervention

[11] 'profiling' means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements

• necessary for entering into, or performance of, a contract between the data subject and a data controller • authorised by law • based on the data subject’s explicit consent?	
Please describe the logic involved in any automated decision-making.	
Please outline the significance and the envisaged consequences of such processing for the data subject.	

ICT

Will we, or the data processor(s), be using a new system to process the personal data?	No If Y, please describe:

If Y to the above question around new systems, please ensure that a System Level Security Policy is completed, and risk assessed by ICT before proceeding to the sign off stage below.

Sign Off of DPIA and Processor Checklist (as appropriate)

IG Lead’s assessment of the level of risk

All risks identified have mitigations in place to minimise any impact. AGEM CSU are have excellent operating standards which ensure that data is securely processed and de-identified to agreed NHS standards. However, all processed should be reviewed and where required updated on a regular basis to take into account any incidents/issues raised and any change in requirements.	
Name	Wendy Harrison
Signature	
Date	April 2020

ICT assessment of the level of risk

(Where SLSP provided)

Post	

Name	
Signature	
Date	

Data Protection Officer

Advice of the DPO:	
Name	Carol Mitchell, Head of Information Governance
Signature	C. Mitchell
Date	April 2020

Information Commissioner's Office

(Where DPIA submitted for review)

Advice of the ICO:

Caldicott Guardian

(If any or all of the information as part of the initiative, project etc. is subject to the common law duty of confidence – e.g. patient identifiable data)

Advice of the CG: None	
Approved as flows supported under the COPI notice issued by the Secretary of State for Health in March 2020. Reviewed retrospectively due to urgency to ensure the NHS could meet the requirements necessary to support the response to COVID-19.	
Post	NHSE/I Medical Director
Name	Professor Stephen Powis
Signature	
Date	11 May 2020

Senior Information Risk Owner

(In all cases)

Decision of the SIRO: DPIA approved subject to regular review given the addition of new datasets, wider access to the dashboards and some agreed disseminations.	
Post	Director of Corporate Operations
Name	Mark Blakeman
Signature	
Date	April 2020

Processor Checklist

SUPPLIER INFORMATION

Supplier name: Palantir Technologies UK, Ltd.

Address:

New Penderel House 4th Floor 283 - 288 High Holborn London WC1V 7HP

Telephone number: +44 (0) 203 856 8404

Name of key contact: Palantir Technologies UK, Ltd.

Telephone number and email address of key contact:

info@palantir.com

+44 (0) 203 856 8404

If your organisation is registered with the ICO please provide Data Protection Notification [ico.org.uk] number: ZA146218

Is your organisation compliant with the Data Security and Protection Toolkit: YES

SERVICE TO BE SUPPLIED

Please describe the service which is to be provided:

Provisions of license to Palantir Foundry together with associated services in accordance with the G-Cloud 11 Call Off Contract between Palantir Technologies UK, Ltd. and NHS England (Arden & GEM CSU).

Palantir licenses software and provides related services to enable customers to integrate, manage, analyse, and interact with data. The Palantir Foundry Platform integrates common data sources (including but not limited to HDFS, JDBC, SQL databases, flat files, etc.) out of the box and can be configured to support other source systems or legacy technologies. It is designed to be modular, configurable, and scalable in order to meet the needs of a wide range of organizations. The Palantir Foundry Platform includes tooling for data integration, pipeline management, security, search, analysis, reporting, and collaboration.

The NHS secure instance of Palantir Foundry is to be named NHS Foundry.

CHECKLIST

1. What data will you be processing on our behalf?	The data catalogue is available at: https://data.england.nhs.uk/covid-19/
2. Will you be processing at an NHS site? If so, where?	No. Palantir Foundry will be provided for use on Amazon Web Services London environment and data will be processed in this environment.
3. Will your staff have remote access to NHS England data? If yes, please explain.	Yes, the environment will be hosted on AWS.
4. Will you be storing any NHS England data in paper format for any length of time? If yes, how will the data be stored?	No
5. Will you be storing any NHS England data in electronic format?	Yes
6. If yes to Q5: - Do all users of your systems have their own log-in and password?	Yes, Palantir requires all users to have a unique ID and password. Additionally, Multi-Factor Authentication is required to log in.
- How are access rights to systems controlled?	NHSE will utilise their NHS Improvement SAML 2.0 Single Sign-on Solution to manage, review, and authenticate users. NHSE access control policies will determine who should access the data integrated in NHS Foundry and for what purpose. Only members of the Palantir team authorised and onboarded by NHSE to NHS Foundry may have access to NHSE data. Palantir employs controls which prevent Infrastructure Operators accessing data integrated in NHS Foundry. Palantir employee access is approved by designated Palantir team leads following Palantir's annually updated and management approved Access Control policy and processes. They include requirements for the verification of identity, regular verification of users and access, and procedures for new user access requests, changing access, and updating and deleting users upon termination or when responsibilities change.

- Are back–ups encrypted?	Yes
- What protection do you have against malicious code?	Palantir relies on a broad spectrum of technical and operational intrusion protection and detection controls. These controls include endpoint malicious code (virus) protection, operating system hardening including mandatory access controls, mandatory operating system and application patching, automated vulnerability scans, tightly controlled network ingress/egress controls, encryption of data at rest and in transit, host and network intrusion detection systems, monitoring of logs and system state, and alerting strategies designed to detect and respond to deviations from expected state and other specific indicators of malicious activity to reduce the exposure and potential impact of malicious software.
- How often do you apply security patches?	Patches, configuration changes, upgrades, and updates are delivered continuously after undergoing Palantir's pre-release and post-release review processes. Security patches are applied as quickly as possible and are prioritized based on criticality, severity, impact, and mitigating controls.
- How often do you risk assess your security controls?	Palantir regularly, at least annually, performs both a Risk Assessment and an Effectiveness Review of the Palantir information security program. In addition, Palantir regularly, at least annually, reviews its Information Security policies, conducts internal audits, and is externally audited by independent third parties.
- What business continuity/disaster recovery plans do you have in place?	Palantir maintains a formal business continuity program for its critical corporate systems and for the NHS Foundry hosting infrastructure following guidance provided by ISO/IEC 27001, 22301, and 22313 as well as Business Continuity Institute's (BCI) Good Practice Guidelines (GPG) 2018.
- Are USB ports/CD writers on staff equipment disabled?	Yes. Usage of USB devices, including CD writers, are only permitted with a documented exception granted by the Palantir information security team. By default, users are not permitted to copy data to any unauthorised removable media device.
- Will you be storing data outside the UK? If so, where? What information governance considerations have been taken into account?	All customer data will be stored in NHS Foundry secure environment hosted on AWS in the UK region.
- Will the data be linked with any other data collections? If so, how will the linkage be achieved?	Yes, Palantir Foundry is a data integration platform. Palantir Foundry enables NHSE to integrate common data sources (including but not limited to HDFS, databases, flat files, etc.) into their owned and controlled Foundry instance, NHS Foundry. Data ingestion requests go through a robust NHSE governance process prior to being made available to the

	approved users. More details about the datasets are found on: https://data.england.nhs.uk/covid-19/
7. What security controls do you have in place for your office premises?	Palantir has a documented physical and environmental security program to protect Palantir facilities. Palantir's physical security controls include security guards, CCTV surveillance, badge readers, restricted access to sensitive spaces, and a separated reception area for visitors.
8. What controls do you have in place for the security of your equipment?	Palantir utilises full disk encryption, host-based firewalls, enforced inactivity timers, and industry standard software to protect corporate workstations. The software is designed to prevent, detect, and remediate against malicious software. In addition, Palantir workstations are managed by one or more endpoint management systems. These systems log details that include date of most recent reconfiguration and the current versions of operating systems, security tools, patching, etc. Endpoint management systems also authenticate, authorise, log, and patch all workstations to meet Palantir's security standards. Endpoints that fall out of compliance are not allowed to connect to Palantir's systems.
9. Have you had any security incidents relating to data in the last three years? If so, please explain.	No. Palantir has not suffered a substantive breech resulting in a financial, operational, or reputational loss within the last three years.
10. What policies do you have in relation to information security, data protection and incident reporting? Please provide copies.	The following documents had been provided for NHS's review: 1. Palantir Information Security Policies and Procedures Intro 2. Palantir Information Security Policies Table of Contents 3. Palantir Data Handling Policy 4. Palantir Data Incident Notification Policy
11. Describe potential disciplinary actions for breach of policy.	Palantir's legal and human resources teams follow an internal investigation and disciplinary process for instances of employees and contractors suspected of violating company privacy and security policies.
12. What steps do you take to ensure that the people your recruit have the honesty and integrity to handle person identifiable data?	Palantir verifies the identity of and runs background checks on all employees and contractors prior to granting access to NHS Foundry - these checks are compliant with BS7858 and include verification of academic and professional qualifications, exclusion from sanction lists, work eligibility, and criminal background checks - subject to the laws and regulations of applicable jurisdictions.
13. How do you ensure that your staff understand the importance of data security and how to keep person	All employees receive data handling and security training when hired and annually thereafter; employees receive additional training specific to the customer they're supporting when they're onboarded to a new team, and Palantir annually sends communications from

identifiable data secure?	the Compliance team & Management regarding policies, data handling, and security obligations.
14. How frequently do you provide your staff with any training on data security and confidentiality and is their learning tested?	See above for details on training. Understanding is verified through tests.
15. Will you ever transfer NHS England data electronically? If so, how will it be transferred?	Yes, NHSE information systems will connect to NHS Foundry through a NHSE-hosted Data Connector which securely captures data and metadata from source systems and imports them into the Foundry platform via encrypted, unidirectional, outbound requests. Data will be transferred using an Azure VM which is located inside the NHSE Azure environment. The Azure VM runs the Data Connector and transfers the data to NHS Foundry. The Data Connector is restricted to applicable data through scoped credentials and host and network level firewalls.
16. Will you ever transport any NHS England hard copy data? If so, what security controls will be in place?	There will be no hard copy data transferred.
17. Will you ever destroy any NHS England data? If so, how will this be done and what evidence of the destruction will you provide?	Palantir will return or destroy any customer confidential information per Palantir policy, regulatory, and contractual requirements. Evidence of destruction can be provided through certification. Data in the Palantir Platform is sanitised through the destruction of the encryption keys. Palantir's cloud service providers decommission media using techniques detailed in NIST 800-88. Details can be found on their website.
18. Will you ever sub-contract work in relation to NHS England data? If so, in what circumstances?	There are no plans to sub-contract work.
Date completed	Last revised on 2 May 2020
Completed by	Palantir Technology Compliance Team
Telephone number and email address	+44 (0) 203 856 8404