Part 2 – Form of Annex: Specific Processing Instructions

This is an Annex to the Data Processing Agreement between the Controller and the Processor dated 22/11/2024 as updated from time to time. Capitalised terms used in this Schedule have the meaning given to them in the Agreement.

Description	Details[1]
Controller for each Category of Personal Data	NHS England is controller for all Personal Data categories
Processor(s)	Palantir Technologies UK LTD – Processor for Federated Data Platform
Subprocessors	Amazon Web Service, Inc – Cloud hosting and infrastructure, alerting and encrypted notification services
Commencement of Processing	The Ontology is a Core Capability of the national Instance of FDP and commences May 2024
Product Name	**FDP National Ontology**
Duration of the Processing	Duration of the Services Agreement or, if shorter, the Controller's use of the Ontology
Nature and purposes of the Processing	The FDP National Ontology (Ontology) is a collection of curated and transformed Pseudonymised datasets and place and metric datasets which use Aggregated Data and Operational Data, which together provide the data sources in FDP for many of the national Products. The purpose of the Ontology is to provide NHS England with a single consistent data source to support Products in the national Instance. This ensures that when different Products are developed, the data being used is consistent, quality assured and in the agreed format. The Ontology as a whole is never made available to external organisations and access to it within NHS England is limited to a small number of developers only who have responsibilities for maintaining the Ontology. When creating Products which are rolled out to the National FDP Instances, data from the Ontology may be used. Each product DPIA using data from the Ontology will detail exactly what data has been used in the development of the Product in the form of a Data Specification which detail the datasets from the Ontology and either the data items or a description of the relevant data items. All Pseudonymised Data ingested into the Platform flows through the Person Ontology before being made available to create national Products. This enables the Data to be conformed in terms of data quality standards and assigned to specific events. Processing includes using the Pseudonymised Data sourced from NHS England's NCDR/UDAL to create cleaned and enhanced data in the Person Ontology that can be further used to calculate metrics in Metrics Ontology. The FDP Processor is instructed to: • provide the core-connector as described in the Overarching DPIA for FDP, to ingest Pseudonymised Data into the FDP to be processed in the Person Ontology, following the Data Ingestion Process (as advised to the FDP Processor by NHS England from time to time); and

---

[1] Update as appropriate.

Description	Details[1]

	• Process Pseudonymised Data ingested into FDP for the purposes of enabling NHS England to create, develop and maintain the Ontology in line with the approved FDP National Ontology DPIA and the Principles below.
	• implement access controls and governance for providing access on NHS England instruction to the limited internal users who may access the Pseudonymised Data in the Person Ontology for the purposes of managing the Ontology in line with the NHS England Role Based and Purposed Based Access Controls.
	
	Principles of the Ontology:
	
	• Patient level Pseudonymised Data is organised as entities and events
	• A sub folder to be created as per the names of the entities and events
	• Patient pseudo id serves as the link between the pseudonymised datasets
	• Data will be organised as bronze, silver and gold. The Classification is an approach to organise the data within FDP. These are data quality categories.
	
	    ○ **Bronze Data** quality means that the datasets are mostly raw and are not enhanced or transformed in any way. This is mainly used in combination with the Person Ontology to produce the silver and gold data sets. The Bronze data set may also be used when a Product requires raw data as it is ingest. The use of this dataset is minimized to where it is absolutely necessary.
	
	    ○ **Silver Data** quality ensures that the data set have following standards:
	
	        ▪ IG restrictions and requirements are documented on the dataset
	        ▪ Metadata of columns will be documented on the dataset
	        ▪ Dataset has Health checks and monitoring to ensure that they are timely refreshed and data quality checks are applied to ensure row counts are as expected.
	        ▪ The documentation should capture the process on versioning datasets. e.g. deprecation of older versions of datasets.
	        ▪ Use reusable functions for standard transformations / business logic.
	
	    ○ **Gold Data** quality datasets ensure that where there are multiple sources of data, data is combined to have the most reliable and timely data . For example for Inpatients dataset (which land in Person Ontology and is then aggregated in Metrics Ontology and are used to create Products such as ERD Dashboard and POD), where we have data from SUS and SUS sources, SUS Data is more reliable and Faster SUS is more timely and is used to create national Products such as ERD Dashboard and POD. Hence they are combined to fill in the latest data from Faster SUS and older data from SUS. All silver level standards are also applicable to Gold data category.
	
	• The IG Restrictions are documented in Ingestion forms. IG Restrictions will be implemented using the user access model in FDP.

Description	Details[1]
	The datasets in the Ontology are not linked together however they may be linked at the point they are required for a particular Product. This linkage will be detailed in the relevant Product DPIA.
Type of Personal Data	All or the data categories below are ingested into FDP in Pseudonymised form. • Postcode • Age • Sex • Gender • Living Habits • General Identifier • Physical/Mental health or condition, diagnosis/ treatment • Racial/ ethnic origin The Data Schema in the FDP National Ontology DPIA sets out the exact data items and datasets to be ingested into the FDP.
Categories of Data Subject	Patient
Plan for return and destruction of the data once the Processing is complete **UNLESS requirement under Union or Member State law to preserve that type of data**	Processor's - Palantir as the FDP provider is configured only to retain data for defined periods that can be configured by Controller. Processor will delete or provide access for Controller to remove data from the Processor's services at the end of duration of processing.
Transfers of data outside the UK	All personal data is stored in the UK and is not to be accessible or processed from outside the UK.
Issued on behalf of the Controller by	**Rebecca Llewellyn** Federated Data Platform Programme Delivery Director **Data Services** Data and Analytics Transformation Directorate NHS England
Date of Issue	16/05/2024