APPENDIX B

A-D

Name	Definition	Guidelines	Example	
ADP	Application Development Projects. This is the HIC team that carries out software development to facilitate secure data collection, consisting entirely of HIC Developers and working mainly with consented data. This team has since been renamed: Software.	• •	D..	
			D..	
Aggregate(d) Data	See Aggregate-Level Data			
Aggregate-Level Data	Summary data which is acquired by combining individual-level data and may be collected from multiple sources and/or on multiple measures, variables, or individuals. It is synonymous with 'aggregate(d) data' or 'statistical results'.			
Algorithm	A set of instructions to execute a task. This is a very general definition; algorithms may be deterministic (always giving the same answer when presented with the same input) or stochastic (giving different answers with various probabilities). Algorithms are not necessarily run by a computer; humans also use algorithms implicitly when making decisions. We will usually use the term to mean a set of instructions which only refer to data generically, rather than a specific dataset. An example of an algorithm is the ordinary least squares (OLS) method for fitting linear models.			
Anonymised Data Return to top	Any and all data that could allow individuals to be identified has been removed. These include CHI, name, date of birth, address, full postcode, GP code, General Medical Council registration number, GP Practice code. Any request for data containing any of this information will be treated as a request for identifiable data, which will require explicit Caldicott approval from the NHS Board(s) of residence of the patient(s). HIC Services recognises that while the data is anonymised it is potentially disclosive, so is treated by HIC Services as potentially personal data.		D..	
			D..	

Application Return to top	The implementation of a service (web application, console application, etc.)			
Approved Data User Return to top	An Approved Data User is an approved DLS Project Principal Investigator (PI) as named on the PM System, or a person who is authorised by the PI to also have access to the DLS Project Dataset. The Approved Data User, to whom data has been made available, will be recorded on the PM system. • The Approved Data User must complete approved Information Governance training and provide the certificate to HIC. • Employees of NHS Tayside, NHS Fife, the University of Dundee, or the University of St Andrews need to read, sign, and follow the terms of the HIC Data User Agreement. • Where an Approved Data User is not such an employee the HIC Data User Agreement must also be signed by a senior representative of the Approved Data User's organisation.			
Approved Project Return to top	An approved project is a project that is logged into the Project Management System and has Ethics, Caldicott and NHS R&D governance approval, as required.			
Attack Types - Black Box Return to top	A type of model attack where the attacker has only query access to the model. That is, they can present input data to the model and observe the predictive outputs that the model makes. For example, in a model that detects the presence/absence of a tumour in an x-ray image, the attacker can present an image to the model and will receive the probabilities that a tumour is present or not. Black box attacks do not have access to the interior of the model.			
Attack Types - White Box Return to top	A type of model attack where the attacker knows some information about the training data, the target model classifier, architecture and learned parameters of the target model. For example, this might include knowing the weights of a neural network, or the decision thresholds in a rule or tree-based model.			
Attacker or Adversary	A person or group of persons who attempts to extract, from the trained ML model, some or all of the personal data that			

Return to top	was used to train it.			
Attribute Inference Return to top	The risk that an attacker, given partial information about a person, can retrieve values for missing attributes in a way that gives them more information than they could derive just from descriptions of the overall distribution of values in the dataset.			
Attribute Inference Attacks (AIA) Return to top	A type of attack where the adversary is capable of discovering a few characteristics of the training data. Individual Disclosure Occurs when outputs from an analysis segment a participant with a specific condition, e.g. rare genetic disease, or a unique combination of conditions that might put the data of this individual at high risk of being identified or disclosed.			
Automated Build Environment Return to top	Software system used by HIC to manage consistent, tested, and automated building and release of applications.			

E E-K

Name	Definition	Guidelines	Example

L L-O

Name	Definition	Guidelines	Example

P P-S

Name	Definition	Guidelines	Example

T T-Z

Name	Definition	Guidelines	Example