DATA PROTECTION IMPACT ASSESSMENT

A Data Protection Impact Assessment (DPIA) is a process that helps an organisation identify and minimise the data protection risks of a project.

This DPIA template must be completed wherever there is a change to an existing process or service, or a new process or information asset is introduced that uses personal data.

Template Version 1.02 August 2023

Document Version Control

Version	Reason	Date	Author(s)
V1	Initial version	04/10/2024	Steve Durbin
V2	Incorporate subgroup outcomes 09/01/2025	12/02/2025	Steve Durbin
V2.1	Update changes to Heidi “free” contract	17/02/2025	Steve Durbin

Project / Work Stream Name:	**Heidi Health – AI Scribe for GP services**	
Project / Work Stream Lead: ***(Who is leading the project or implementation)***	Name:	Dr Anshumen Bhagat
	Designation:	GP Partner
	Email:	abhagat@nhs.net
Information Asset Owner: ***(Who will be responsible for the data that is stored / created)***	Name:	(No new data is produced and held by system)
	Designation:	
	Email:	
Key Stakeholder Names and Roles:		
Overview: **(Summary of the project/work stream)**	From the website: “Heidi is the ambient clinical AI that frees you from note-taking, insurance-pleading, results-finding and all the other tasks that make you hate your job.” Basically, it can create letters / referrals / summaries from multiple sources on the device including recordings, soft phones, microphones.	
Timeframe for the project / work stream:	Immediate. Some practices are already using.	

(When is it due to begin? If time limited, when does it end / need to be reviewed)	
Environmental Scan: ***Describe the consultation/checks that have been carried out regarding this initiative or, project of similar nature, whether conducted within your organisation or by other organisations.***	There are many similar products. Tortus AI is being supported by the Primary Care Team in the ICB, but practices report there is lower functionality in this product. However, Tortus has achieved DCB certification as a medical device, this product has not. NHSE are trialling AI facilities in Teams and have briefed DPOs on this. Anyone can cut-and-paste into online LLMs and use them to obtain answers – this is a risk, and Tortus/Heidi have the advantage of higher privacy levels. Hence this use must be considered as the "lesser of two evils".

Step 1: Complete the Screening Questions

Q	Category	Screening question	Check if Yes
1.1	Technology	Does the project introduce new or additional information technologies that can substantially reveal an individual's identity and has the potential to affect that person's privacy?	☒
1.2	Technology	Does the project introduce new or additional information technologies that can substantially reveal business sensitive information, specifically: have a high impact on the business, whether within a single function or across the whole business?	□
1.3	Identity	Does the project involve new identifiers, re-use or existing identifiers e.g. NHS or NI number, Local Gov. Identifier, Hospital ID no. or, will use intrusive identification or identity management processes or, electronic linkage of personal data?	□
1.4	Identity	Might the project have the effect of denying anonymity and pseudonymity, or converting transactions that could previously be conducted anonymously or pseudonymously into identified transactions?	□
1.5	Multiple organisations	Does the project involve multiple organisations, whether they are public sector agencies i.e. joined up government initiatives or private sector organisations e.g. outsourced service providers or business partners?	☒
1.6	Data	Does the project involve new process or significantly change the way in which personal data/special categories of personal data and/or business sensitive data is handled?	☒
1.7	Data	Does the project involve new or significantly changed handling of a considerable amount of personal data/special categories of personal data and/or business sensitive data about each individual in a database?	□

Q	Category	Screening question	Check if Yes
1.8	Data	Does the project involve new or significantly change handling of personal data/special categories of personal data about a large number of individuals?	□
1.9	Data	Does the project involve new or significantly changed consolidation, inter-linking, cross referencing or matching of personal data/special categories of personal data and/or business sensitive data from multiple sources?	□
1.10	Data	Will the personal data be processed out of the UK?	☒
1.11	Exemptions and Exceptions	Does the project relate to data processing which is in any way exempt from legislative privacy protections?	□
1.12	Exemptions and Exceptions	Does the project's justification include significant contributions to public security and measures?	□
1.13	Exemptions and Exceptions	Does the project involve systematic disclosure of personal data to, or access by, third parties that are not subject to comparable privacy regulation?	□

The purpose of the screening questions is to confirm that the data protection laws are being complied with, or highlights problems that need to be addressed. It also aims to prevent problems arising at a later stage which might impede the progress or success of the project.

Answering "Yes" to any of the screening questions above represents a potential Information Governance (IG) risk factor, please proceed and complete the next section.

Step 2: Identify the need for a DPIA

2.1	**Is this a new or changed use of personal data/special categories of personal data and/or business sensitive data that is already processed/shared?**	New	□
		Changed	☒

2.2 **Please specify data used. For items marked "*" please provide more detail in the box. If forms or data definitions are available, please include**

Personal Data					
Name	☒	Address	☒	Postcode	☒
Email	☒	Phone / Mobile	☒	Date of Birth	☒
NHS No. / EPR ID No.	☒	NI No.	□	Payroll number	□
Driving licence	□	Bank / Credit details	□	Tax / Benefit / Pension	□
School Records	□	Housing Records	□	Other identifiers* (e.g. mother's maiden name, passwords, logins)	☒
Special Category Personal Data					
Racial / ethnic origin	☒	Political Opinions	□	Religious or philosophical beliefs	□
Trade Union Membership	□	Physical or mental health*	☒	Sexual life*	□
Criminal offences*	□	Biometrics* (DNA profile, fingerprints, etc.)	□	Adoption records*	□
Social services records*	□	Child protection records*	□	Safeguarding records*	□

Other personally identifiable data (please list)

Further details of items marked *

Other identifiers – logins for staff; Physical/Mental health – anything that comes up in recordings which will be used to create the outputs.

Note that identifiers are restricted in part due to the web anonymisation processes, but are processed *de facto* by the system.

2.3	**Business sensitive data**	**Check if Yes**	**Details**
	Financial	□	
	Local Contract conditions	□	

	Operational data	□	
	Notes associated with patentable inventions	□	
	Procurement/tendering information	□	
	Customer/supplier information	□	
	Decisions impacting:	One or more business functions	□
		Across the organisation	□
	Description of other business data processed/shared/viewed (if any).		

Step 3: Describe the sharing/processing	
3.1	**List of organisations/partners involved in sharing or processing personal/special categories personal data.**

If more rows are needed, select the right-hand box in the last row and click "+".

Name or class of organisation	Type of Organisation (drop-down)	Completed and compliant with the Data Security and Protection (DSP) Toolkit (drop-down)
NCL GPs	Controller	Group (DSPT required)
Heidi Health UK Ltd (subsidiary of Australian company)	Processor	No
Heidi Health (Australia)	Processor	Standards Met
Google LLC (Ireland)	Subprocessor	Standards Exceeded
AWS	Subprocessor	Standards Exceeded
Kinde (Ireland)	Subprocessor	Not Required
Stripe	Subprocessor	Not Required
Intercom (Ireland)	Subprocessor	Not Required

3.2	**Is there an existing ' Data Processing Contract' or 'Data Sharing Agreement' between the Controller and the Processor?** *If no, please provide details of legal route for Article 26/28 compliance below*	**Yes**	☒	**No**	□
	A DPA is available covering the UK legal requirements; note that this is with Heidi UK – the one with the Australian company is not Article 28 compliant. As at 17/02/2025 - The "free" version on the website now recognises England and Wales law and provides for storage in UK/EEA only. This is now compliant, Heidi have stated the issue with the contract was due to a website programming error.				
3.3	**Has a data flow mapping exercise been undertaken?** *If yes, please provide a copy at Annex 2 below, if no, please provide clear reason why not in the box below.*	**Yes**	☒	**No**	□
	Details from DPA included in annex.				
3.4	**Does the project involve employing contractors external to the Organisation who would have access to personal or special categories of personal data? Please check if Yes.**	□			

	If checked, provide a copy of the confidentiality agreement or contract as an annex	
3.5	**Describe in as much detail why this information is being processed/shared/viewed?** *(For example Direct Patient Care, Statistical, Financial, Public Health Analysis, Evaluation. See NHS Confidentiality Code of Practice Annex C for examples of use)*	
	Direct care	

Step 4: Assess necessity and proportionality		
4.1	**Lawfulness for Processing/sharing personal data/special categories of personal data?** ***You need to check one box in each section. For most health and care purposes, it's the first box in each section.***	
	Personally Identifiable Data	
	☒	The **DPA section 8(c)** – "the exercise of a function conferred on a person by an enactment or rule of law", specifically the NHS Act 2006 and the Health and Social Care Act 2012, allowing use of **UK GDPR Article 6(1)(e)** '…for the performance of a task carried out in the public interest or in the exercise of official authority…' **Note:** This is the most common legal basis for health and care processing
	☐	**UK GDPR Article 6(1)(a)** – "the data subject has given consent to the processing of his or her personal data for one or more specific purposes;" **Note:** This is a very rare and unusual case in health and care. Relying on this has legal consequences which need to be reviewed elsewhere in the DPIA.
	☐	**UK GDPR Article 6(1)(b)** – "…for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract;" **Note:** The DATA SUBJECT must be part of the contract; we cannot rely on this basis for contracts between health bodies.
	☐	**UK GDPR Article 6(1)(c)** – "…for compliance with a legal obligation to which the controller is subject;" **Note:** There is no need to check this one if one of the above is checked.
	☐	**UK GDPR Article 6(1)(d)** – "…in order to protect the vital interests of the data subject or of another natural person;" **Note:** There is no need to check this one if one of the above is checked.
	☐	**UK GDPR Article 6(1)(f)** – "…for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child;" **Note:** Public authorities cannot rely on legitimate interest. If you are using this, there needs to be careful review to ensure that you are not formally a public authority.
	Special Categories of Personally Identifiable Data	
	☒	The **DPA section 10(1)(c)** – health and social care via Schedule 1 Part 1 section 2 "Health or social care purposes" satisfying DPA section 10 (2) allowing use of **UK GDPR Article 9(2)(h)** '…medical diagnosis, the provision of health or social care or treatment or the management of health or social care systems…' **Note:** This is the most common legal basis for health and care processing

□	**UK GDPR Article 9(2)(a)** – “the data subject has given explicit consent to the processing of those personal data for one or more specified purposes…” **Note:** This is a very rare and unusual case in Health and care. Relying on this has legal consequences which need to be reviewed elsewhere in the DPIA.
□	**DPA Section 10(2)** meeting a criterion in Part 1 of schedule 1 allowing use of **UK GDPR Article 9(2)(b)** – “…necessary for the purposes of carrying out the obligations and exercising specific rights of the controller or of the data subject in the field of employment and social security and social protection law …” **Note:** This generally would apply staff or patients in social care situations not covered by DPA Section 10. There is no need to check this one if DPA Section 10(1)(c) above is checked.
	If you checked this section, detail below which criterion in Part 1 of Schedule 1 is met?
□	**UK GDPR Article 9(2)(c)** – “…necessary to protect the vital interests of the data subject or of another natural person where the data subject is physically or legally incapable of giving consent;” **Note:** There is no need to check this one if DPA Section 10(1)(c) above is checked.
□	**UK GDPR Article 9(2)(d)** – “…carried out in the course of its legitimate activities with appropriate safeguards by a foundation, association or any other not-for-profit body with a political, philosophical, religious or trade union aim…” **Note:** This is unlikely for health and care bodies.
□	**UK GDPR Article 9(2)(e)** – “…relates to personal data which are manifestly made public by the data subject;” **Note:** This is unusual and would only apply in rare circumstances. Note that this excludes data made public by someone *other* than the data subject.
□	**UK GDPR Article 9(2)(f)** – “…necessary for the establishment, exercise or defence of legal claims or whenever courts are acting in their judicial capacity;” **Note:** There is no need to check this one if DPA Section 10(1)(c) above is checked.
□	**DPA Section 10(1)(b)** “substantial public interest” satisfying a condition in DPA Schedule 1 Part 2, allowing use of **UK GDPR Article 9(2)(g)** “…necessary for reasons of substantial public interest…” **Note:** There is no need to check this one if DPA Section 10(1)(c) above is checked.
	If you checked this section, detail below which criterion in Part 2 of Schedule 1 is met?

	☐	**DPA Section 10(1)(d)** "public health", satisfying Section 3 of DPA Schedule 1 Part 1, allowing use of **UK GDPR Article 9(2)(i)** – "…necessary for reasons of public interest in the area of public health…" **Note:** This is very distinct from the health and care purpose 10(1)(c). It is possible, but uncommon, to have processing that is in both areas.
	☐	**DPA Section 10(1)(e)** *"archiving, research and statistics"*, satisfying Section 4 of DPA Schedule 1 Part 1, allowing use of **UK GDPR Article 9(2)(j)** – "necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes …" **Note:** Most statistics are covered by health and care processing, this is most commonly used for approved medical research, the National Archive or non-medical statistics involving personal data.
		If you checked this section, detail below how the requirements of DPA Section 19 and Article 89 UK GDPR are met?

4.2	**If the use case is NOT direct care, how are the requirements of Section 251 of the NHS Act 2006 being met, in order to set aside the Common Law Duty of Confidentiality?** *Please describe in detail below what the CAG approval is, if obtained, or otherwise how this is achieved.*				Confidentiality Advisory Group Approval	☐
					Other Legal route	☐
	If CAG approval, please provide App No and Opinion Reference *These can be obtained from the HRA website registers*	App		Ref		
	N/A – direct care					
4.3	**Will the information be processed/shared electronically, on paper or both?**				Electronic	☒
					Paper	☐
4.4	**How will you ensure data quality and data minimisation?** *Note that if existing processes cover the rights, please answer "Existing Processes". The reviewer may confirm these processes.*					
	Existing processes. Note that the system does NOT store identifiable data in the cloud – it is anonymised on entry and re-identified on exit.					
4.5	**Have individuals been informed about the proposed use of their personal or special categories of personal data? If No, how is the transparency principle met?** *For example, do the organisations/partners listed in section 3.1 have updated Fair Processing Notice available to patients on their websites?*				Yes	☒
					No	☐
	Added to GP practice privacy notice.					

4.5	**How will you recognise and respond to rights requests from individuals?** *Note that if existing processes cover the rights, please answer "Existing Processes". The reviewer may confirm these processes. Note that ALL rights must be covered, not just subject access.*		
	Existing processes		
4.7	**How will you recognise and respond to data breaches?** *Note that if existing processes cover the right, please answer "Existing Processes". The reviewer may confirm these processes.*		
	Existing processes		
4.8	**Will the processing of data include automated individual decision-making, including profiling?** *If yes, please outline the profiling processes, the legal basis underpinning the process, and the rights of the data subject*	Yes	☒
		No	☐
	The summarisation and production is fully automated – we therefore need to ensure that a human review takes place i.e. the practitioner must review each output before use.		
4.9	**If Consent is used as the legal basis above, how will consent be obtained and recorded?** *If not using consent, respond N/A*		
	Consent is not legal basis – however, as the system records we MUST inform the patient if it is used during remote patient consultation and they have a right to object to recording under UK Law. In person-to-person conversations, there is a reasonable expectation of privacy and therefore recording must be informed to the patient and we recommend patient is given opportunity to object. Note that Heidi Health have stated that a notice in the waiting room is sufficient – the GP DPO strongly disagrees with that view.		
4.10	**As part of this work is the use of Cloud technology being considered either by your own organisation or a 3rd party supplier? If so please complete the cloud security questionnaire and add as an annex or state below why it is not required.** *Note that the questionnaire is not required for systems that already have a separate DPIA.*	Yes	☒
		No	☐
	AWS UK cloud		
4.11	**Where will the data will be stored?** *Please state countries of storage, locations and types used. Examples of Storage include bespoke system, e.g. EPR, Emis & other clinical systems, SharePoint, data repository, Network Drives, Filing cabinet, storage area/filing room and location, etc.*		
	Anonymised data in UK/EU cloud		
4.12	**Data Retention Period** *How long will the data be kept?*		
	Until end of contract plus 10 days.		
4.13	**Will this information being shared/processed outside the organisations listed above in question 3?** *If yes, describe who and why:*	Yes	☐
		No	☒

4.14	**Is there linking of information between different data sets?** *If yes, describe how this is achieved and the legal basis for doing so. Note that linking data where patients are not clients of both organisations needs particular review.*	Yes	□
		No	☒

Step 5: Information Security Process			
5.1	**Is there an ability to audit access to the information?** *If no, please provide a reason why this is not required. If yes, please describe auditing.*	Yes	☒
		No	□
	Audit of the cloud system is given in Heidi's DPIA.		
5.2	**How will access to information be controlled?**		
	Username / password and MFA		
5.3	**What roles will have access to the information?** (list individuals or staff groups)		
	Practitioners and support persons working for Heidi Health where requested by clinicians. Heidi do not access data without practitioner agreement.		

5.4	**What security and audit measures have been implemented to secure access to and limit use of personal data/special categories of personal data and/or business sensitive data?**					
	Username and password	☒	Smartcard	□	key to locked filing cabinet/room	□
	Secure 1x Token Access including authenticator apps, SMS	☒	Restricted access to Network Files			□
	Other: *Provide a Description Below*:					

5.5	**Is there a documented System Level Security Policy (SLSP) for this project? If yes, please add a copy as an annex.** SLSP is required for new systems. *SLSP refers to the architecture, policy and processes that ensure data and system security on individual computer systems. It facilitates the security of standalone and/or network computer systems/servers from events and processes that can exploit or violate its security or stature.*	Yes	□
		No	☒
		Not a new system	□
5.6	**Are there Business Continuity Plans (BCP) and Disaster Recovery Protocol for the proposed/existing system or process?** *Please explain and give reference to such plan and protocol*	Yes	□
		No	☒

		Not a new system □
	No SLSP; fallback is to perform manually.	

5.7	**Is Mandatory Staff Training in place for the following?**	**Yes**	**No**	**N/A**
	Data Collection, where there is new data being collected?	□	☒	□
	Use of the System or Service, for new systems?	□	☒	□
	Information Governance, where organisations involved are not toolkit certified?	□	☒	□
	If you answered any of the above as “Yes” or “No” please detail training *A response is not needed if all answers were N/A*			
	No training is provided as far as we can determine.			

5.8	**Are there any new or additional reporting requirements for this project?** *If no, skip to 5.9. If yes, provide details below.*				Yes	□
					No	☒
	What roles will be able to run reports?					
	N/A					
	What roles will receive the report or where will it be published?					
	N/A					
	Will the reports be in person-identifiable, pseudonymised or anonymised/aggregate format?					
	Identifiable	□	Pseudonymised	□	Anonymised/Aggregate	□
	Will the reports be in business sensitive or redacted format (removing anything which is sensitive)?					
	Business Sensitive	□	Business Redacted	□	No Business Data	□

Step 6: Identify and Assess Risks		
Note that this section should only cover risks of the NEW or CHANGED process. If existing systems are used, their protective measures are covered in the DPIA / SLSP for the system and should not be repeated here.		
Risk Scoring		
Score	Likelihood of Harm (L)	Severity of Harm (S)
1	Rare – can't believe this will ever happen	Insignificant – no injury or adverse outcome; no risk to persons or organisation. Unlikely to cause complaint. Litigation risk remote
2	Unlikely – do not expect to happen, but possible	Minor – short term injury / damage; minimal risk to persons or organisation. Complaint possible, litigation possible.
3	Occasionally - May occur	Moderate – semi-permanent injury / damage risk to persons or organisation. Complaint likely; litigation possible, but not certain
4	Likely – will probably occur but it is not a persistent issue	Major – risks of severe injury / damage to persons or organisation. Litigation expected/certain. Local adverse publicity.
5	Almost certain – likely to occur on many occasions, a persistent issue	Catastrophic – risk of death or inability of organisation to continue. Formal investigation likely, litigation certain, national adverse publicity.

Describe source of risk and nature of potential impact on individuals. Include associated compliance and corporate risks as necessary. ***If you need more rows, click in the last box on the right of the last row, then click the "+".***	**Likelihood of harm (L)**	**Severity of harm (S)**	**Overall risk** **(L x S)**
Risk that the data is incorrectly depersonalised, resulting in processing of personal data which is not necessary	3 - Occasionally	2 - Minor	6 to 10 - Medium
Risk that data interpretation is incorrect, resulting in invalid entries on medical records or letters	3 - Occasionally	4 - Major	11 to 25 - High
Risk that the device is not certified medically and this is adjudged a "medical use" resulting in potential legal risks.	2 - Unlikely	4 - Major	6 to 10 - Medium
Patients nor correctly informed of use of data by practitioner. In particular, in ways that are intelligible to them e.g. correct level and language.	4 - Likely	4 - Major	11 to 25 - High

Describe source of risk and nature of potential impact on individuals. Include associated compliance and corporate risks as necessary. ***If you need more rows, click in the last box on the right of the last row, then click the "+".***	**Likelihood of harm (L)**	**Severity of harm (S)**	**Overall risk** **(L x S)**
The system does not appear to have compensation for accents, dialects, and intonations leading to possible misinterpretation.	3 - Occasionally	4 - Major	11 to 25 - High

Step 7: Identify Measures to reduce risk

Identify additional measures you could take to reduce or eliminate risks identified as medium or high risk in step 6

If you need more rows, click in the last box on the right of the last row, then click the "+".

Risk	**Options to reduce risk likelihood or severity**	**Effect on risk**	**Residual risk (see above for scores)**			**Measure approved**
			L	**S**	**Score**	
Risk that the data is incorrectly depersonalised, resulting in processing of personal data which is not necessary	Heidi have measures for the purpose.	Mitigated	2	2	4	Choose an item.
Risk that data interpretation is incorrect, resulting in invalid entries on medical records or letters	Practitioners MUST review all outputs as they are confirming them as their medical opinion; they retain liability. Note that this is a legal requirement to avoid Article 22 (automated processing) being engaged.	Mitigated	2	4	8	Choose an item.
Risk that the device is not certified medically and this is adjudged a "medical use" resulting in potential legal risks.	Heidi are pursuing certification, but it has not yet been obtained.	Accepted	2	4	8	Choose an item.

Risk	Options to reduce risk likelihood or severity	Effect on risk	Residual risk (see above for scores)			Measure approved
			L	S	Score	
Patients nor correctly informed of use of data by practitioner. In particular, in ways that are intelligible to them e.g. correct level and language.	Ensure that privacy notice is fully updated. Ensure that patients are informed of recording of both conversations and telephone calls at ALL times and that AI is being used to analyse. Ensure that such notifications are presented in a language and manner the patient can understand.	Reduced likelihood	2	4	8	Choose an item.
The system does not appear to have compensation for accents, dialects, and intonations leading to possible misinterpretation.	Ensure review as per previous risks. This is separated because there is risk of the patient as well as the practitioner being misunderstood.	Reduced likelihood	2	4	8	Choose an item.

Step 8: Sign off and record outcomes

Item	Details		Notes
IG Subgroup Reviewed	Date:	12/10/2024	Subgroup asked that all AI systems have risks regarding interpretation added 06/01/2025
Risk Measures approved by:	Name:		
	Role:		
	Date:		
Residual risks accepted by:	Name:		Must be a SIRO, IAO or Caldicott Guardian
	Role:		
	Date:		

DPO Advice section – Each Controller to Complete, duplicate this section as needed.		
Controller:	NCL GPs, Federations and PCNs	
DPO advice provided by:	Name:	Steve Durbin
	Date:	12/02/2025
Summary of DPO advice:	All DPO advice has been incorporated into this document. The lack of completed clinical certification remains a concern – practitioners are reminded that the final entry is THEIR medical opinion and cannot be offloaded to the product. Heidi inform us that this is to be completed soon. There are two processing agreements being shared by Heidi Health – controllers must ensure the UK compliant one is signed and used otherwise the processing fails the requirements of the UK GDPR. As of 17/02/2025, the free version on the website is compliant – please check that the jurisdiction paragraph states “England and Wales” (currently para 13) As long as human review is undertaken for **all** entries, and patients are always informed of being recorded, the risks are generally well understood and controlled.	

	Note that informing patients cannot be limited to notices – there must be notification at time of recording, in a manner that is intelligible to the patients (e.g. correct languages and level of language		
DPO advice accepted or overruled by:	Name:		If overruled, you must explain your reasons in explanation, below.
	Role:		
	Date:		
Explanation:			
End of DPO Advice Section			

Item	Details		Notes
Consultation responses reviewed by:	Name:		If your decision departs from individuals' views, you must explain your reasons below.
	Role:		
	Date:		
Explanation:			
The DPIA will be reviewed by the respective DPOs / IAOs of each organisation when required.			

ANNEX 1. GLOSSARY OF TERMS

Anonymised Data - means data in a form where the identity of the individual cannot be recognised i.e. when: Reference to any data item that could lead to an individual being identified has been removed; The data cannot be combined with any other data sources to produce personal identifiable data.

Controller means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data.

Data Flow Mapping (DFM) means the process of documenting the flows/transfers of Personal Data, Sensitive Personal Data (known as special categories personal data under UK GDPR) and Commercially Confidential Information from one location to another and the method by which they flow.

Data Subject – an individual who is the subject of personal information or can be identified from it.

Direct Care - means 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 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. (National Data Guardian definition)

EPR means electronic patient record system; these are the primary systems which provide access to patient

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.

Processing means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.

Processor means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller.

Pseudonymisation means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.

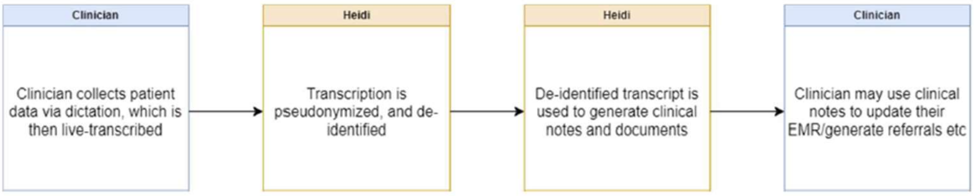
Special Categories of Personal Data mean data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation.

ANNEX 2. DATA FLOW DIAGRAM

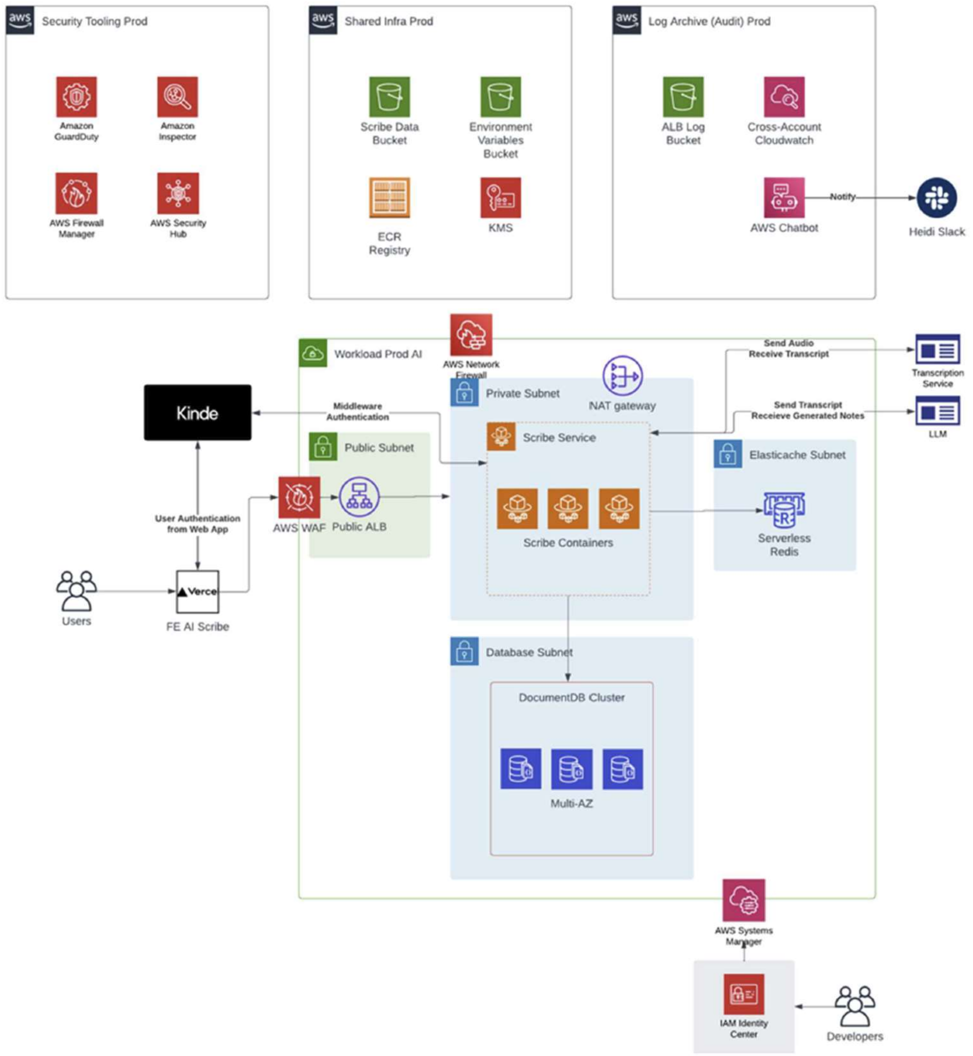
Personal Information Flow Table

Description	**Type**
Clinician obtains consent to record consult, which is then live-transcribed	Collection
Processing transcript through pseudonymization[1] and de-identification models	Use
De-identified transcript processed to generate clinical notes and documents	Use
Clinician may use clinical notes and documents to update their EMR/generate referrals etc	Use

Personal Information Flow Diagram


Clinician
Clinician collects patient data via dictation, which is then live-transcribed
Heidi
Transcription is pseudonymized, and de-identified
Heidi
De-identified transcript is used to generate clinical notes and documents
Clinician
Clinician may use clinical notes to update their EMR/generate referrals etc

Software Architecture & Data Flows


Security Tooling Prod
Amazon GuardDuty
Amazon Inspector
AWS Firewall Manager
AWS Security Hub
Shared Infra Prod
Scribe Data Bucket
Environment Variables Bucket
ECR Registry
KMS
Log Archive (Audit) Prod
ALB Log Bucket
Cross-Account Cloudwatch
AWS Chatbot
Notify
Heidi Slack
Workload Prod AI
AWS Network Firewall
Send Audio
Receive Transcript
Transcription Service
Send Transcript
Receieve Generated Notes
LLM
Kinde
Middleware Authentication
Private Subnet
NAT gateway
Scribe Service
Public Subnet
Elasticache Subnet
User Authentication from Web App
AWS WAF
Public ALB
Scribe Containers
Serverless Redis
Users
Verce
FE AI Scribe
Database Subnet
DocumentDB Cluster
Multi-AZ
AWS Systems Manager
IAM Identity Center
Developers