Data Privacy Impact Assessment

Heidi Health Ltd

Contents

[Reason for processing 2](#_Toc191028494)

[Description of the data 2](#_Toc191028495)

[Purpose and Benefits of sharing the data 2](#_Toc191028496)

[Types of personal data potentially being shared 3](#_Toc191028497)

[Data Provision 4](#_Toc191028498)

[Data linkage 4](#_Toc191028499)

[Data Flow 4](#_Toc191028500)

[Legal Basis for sharing 7](#_Toc191028501)

[Keeping data secure 7](#_Toc191028502)

[Measures to ensure data is only used for the identified reasons 9](#_Toc191028503)

[Data Retention 9](#_Toc191028504)

[Complying with individual rights 10](#_Toc191028505)

[Automated decision making 11](#_Toc191028506)

[Data Controllers 11](#_Toc191028507)

[Data Processor. 11](#_Toc191028508)

[Due diligence measures 12](#_Toc191028509)

[Risks and Mitigations 12](#_Toc191028510)

[Review and sign-off 15](#_Toc191028511)

# Reason for processing

Heidi works by transcribing speech into text from a healthcare encounter such as conversations between clinicians and patients or by clinicians dictating their clinical findings, impression and/or management plans before, during and after the healthcare encounter.

Heidi Health improves clinical documentation by aiding healthcare professionals in note-taking and generating consult summaries. By providing streamlined documentation across multiple access points, real-time transcription processing, and seamless integration with existing workflows, Heidi enables efficient and standardised clinical documentation. The technology enables clinicians to focus on patients during the consultation, contributing to improved patient care. It also acts as a valuable tool for medical practitioners, saving them hours of administrative time per week

# Description of the data

|  |  |
| --- | --- |
|  | Personal data |

Potentially (unless microphone turned off) all sensitive data for that patient is uploaded onto the service but the recording and transcript contains no patient identifiable data (PID) unless data synthesis allows indirect identifiability or anyone in the consulting room provides PID. Special category data including information regarding patient health, will be de-identified and pseudonymized before the transcript is processed into clinical notes or other clinical documents.

If in the consultation a third party is discussed this data will also be included.

# Purpose and Benefits of sharing the data

Using transcription software can improve the quality of the consultation and the quality of the notes.

1. By allowing real time transcription, the clinician has more time to spend time with the patient and record taking does not interrupt the flow and eye contact, potentially improving the patient experience. Allowing more free-flowing consultations also increases the chance the clinician could identify hidden concerns. Releasing time in the consultation allows time for more complex reviews and potentially more issues to be addressed.
2. By having a transcription of the notes, it is less likely important details will be missed when notes are written after the consultation. Heidi Health also gives suggested SNOMED codes which could improve data quality.
3. Having a transcription allows quicker referrals but can also allow more in-depth referrals as in specialities such as mental health, patient concerns can be easily transcribed.
4. Medicolegally, having a true reflection of the consultation can protect the patient and the clinician if concerns are later raised.
5. Heidi Health’s translation feature allows the clinician to be more certain that the translator is asking the correct questions and enables the clinician to give the patient a copy of the consultation in their language.

# Types of personal data potentially being shared

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Forename |  | Physical description and examination, for example height |  | Photograph / picture of people |
|  | Surname |  | Phone number |  | Location data   * GP practice |
|  | Address |  | Email address |  | Audio recordings |
|  | GP details |  | Date of birth |  | Video recordings |
|  | Age |  | Legal representative name (personal representative) |  | Patient information as below |
|  | Gender |  | NHS number |  | National insurance number |
|  | Third party data |  | Registered pharmacy |  |  |

|  |  |  |
| --- | --- | --- |
| **Type of data** | | **Reason required** |
|  | Information relating to an individual’s physical or mental health or condition, for example information from health and care records | As Heidi Health is recording a consultation all relevant data will be included to ensure an accurate record is kept for future care and enquiries. This could potentially include physical, mental and sexual health history, exam findings, conclusions, advice/treatment and referrals |
|  | Information relating to an individual’s sexual life or sexual orientation | The individual’s sexual life and orientation can influence clinical decision making and care recommended |
|  | Racial or ethnic origin | The individual’s race or ethnic origin can influence clinical decision making and care recommended |
|  | Religious or philosophical beliefs | The individual’s religious or philosophical beliefs can influence clinical decision making and care recommended |
|  | Information relating to criminal or suspected criminal offences | The use of non-prescription drugs can influence clinical decision making and care recommended |

**Individuals that can be potentially be identified from the data**

|  |  |
| --- | --- |
|  | Patients or service users |
|  | Carers |
|  | Staff |
|  | Wider workforce |
|  | Visitors |
|  | Members of the public |
|  | Potentially anyone could be identified in a consultation |

# Data Provision

The data will be provided by the consulting clinician and the patient. Other individuals who may input into the data include translation services and other individuals in the room including family, social support workers and other clinicians.

# Data linkage

|  |  |
| --- | --- |
|  | Yes |
|  | No |

# Data Flow

**Personal Information Flow**

Heidi transcribes speech into text from a healthcare encounter. The clinician can also add additional contextual notes about the healthcare encounter which they may not wish to verbalise during the healthcare encounter. The clinician is also able to set and modify various settings within the Heidi platform in order to customise their Heidi experience as well as how their clinical documentation is structured and written. To generate the requested clinical documentation, the transcribed text and contextual notes along with the various user controlled settings are then through an artificial intelligence model which then generates the requested clinical documentation based on the data that has been given to the AI model.

The comprehensive clinical documentation generated by Heidi can then be copied or integrated into an electronic medical record system or used with other word processing or communication tools to provide other clinicians and/or the patient with relevant information related to the healthcare encounter and the patient's care. **Clinicians have full autonomy to delete the sessions on Heidi, either manually or via an automatic deletion schedule that will be set for 24 hours. Once a session is deleted, it is permanently removed from Heidi's servers and is unrecoverable.** This ensures that clinicians have complete control over the data they generate and maintain strict privacy and data management standards. If a clinician chooses not to delete their sessions from Heidi, then the deidentified transcripts will remain on Heidi’s UK servers until deletion occurs.

Data will flow every time a clinician performs a session on Heidi - in other words, every time they click 'Start Transcribing', ‘Start Dictating’, or upload audio to Heidi. This data will cover both data from the patient themselves (though it is de-identified) and the clinician too - such as their templates, note-taking style, clinician type, email address etc

A white rectangular object with black text

AI-generated content may be incorrect.

**De-identification Flow**

A diagram of a flowchart

AI-generated content may be incorrect.

Heidi uses a cloud-based artificial intelligence medical scribe platform. It processes clinical conversations through multiple secure access methods:

1. **Web Browser Interface**
   * Accessible via desktop and mobile browsers
   * Direct audio streaming from microphone input
   * Real-time processing of clinical conversations
   * More detailed technical diagrams can be found in [Appendix A](#appendixa)
2. **Mobile Application**
   * Accessible on both iOS and Android devices
   * Network resilience features for connectivity issues
   * Secure temporary storage and processing of audio data in sandboxed environment
   * More detailed explanation on the security features of the mobile application as well as technical diagrams can be found in [Appendix B](#_heading=h.rjd1143yspkm).
3. **Desktop Application**
   * Accessible on both MacOS and Windows devices
   * System audio integration capability for telehealth consultations
   * Secure temporary storage and processing of audio data on users’ device
   * More detailed explanation on the security features of the desktop application as well as technical diagrams can be found in [Appendix B](#_heading=h.rjd1143yspkm).
4. **EHR Integration (Widget/SDK)**
   * Integration with electronic health record systems
   * Bi-directional data flow for patient demographics and clinical notes
   * Template synchronisation with EHR systems
   * Context-aware processing capabilities
   * Currently available on select EHRs only
5. **Chrome Extension**
   * Operates only on HTTPS-secured websites
   * Facilitates direct insertion of processed notes into text fields
   * Includes specialised processes for contextual note placement

All data is hosted locally within the UK for all UK users, with servers maintained under strict security protocols. Heidi employs end-to-end encryption for data transmission and secure storage practices for any temporarily retained data.

All data processing activities for UK customers are conducted exclusively within the UK. The primary data centres of their sub-processors are located in London, UK, Dublin, Ireland and Frankfurt, Germany, with redundant backup facilities in each respective region to ensure both data sovereignty and service reliability.

# Legal Basis for sharing

Although the technology has been designed to not identify patients, the clinician cannot be certain what the patient will disclose in the consultation which may make the patient identifiable. Therefore, patient or carer consent is obtained before using this tool. Who consented will be recorded in the patient medical record. Patients will be advised that is no requirement to use this tool and they can easily opt out at the start or during the consultation.

|  |  |
| --- | --- |
|  | (a) We have consent GDPR Article 6 (1) a |
|  | (b) We have a contractual obligation |
|  | (c) We have a legal obligation |
|  | (e) We need it to perform a public |
|  | (f) We have a legitimate |

Special category data is required routinely in consultations and although consent will be obtained from patients, the legal basis is for the provision of health care ie Article 9 (2) h, DPA2018 Schedule 1 Part 1. Article 2 (1) 2 (d). Heidi Health ensure their systems fulfil article 5 of the GDPR.

# Keeping data secure

Data is being collected, transferred and transcribed in real time as the patient and clinician plus any others involved e.g. translator, carer, family member discuss concerns in the consultation and any additions are added by the clinician afterwards. The transcribed data is then stored on the cloud in the UK before being transferred to the clinical record within 24 hours. The records should be updated in real time but it is accepted that, by exception, the clinician may have insufficient time to enter the details into the record. For this reason, the recorded data is held on the cloud for 24 hours.

|  |  |  |
| --- | --- | --- |
| **Storage location** | | **Details** |
|  | Physical storage, |  |
|  | Local organisation servers |  |
|  | Cloud storage | [Trust Center - Heidi Health](https://trust.heidihealth.com/subprocessors) |
|  | Other |  |

|  |  |  |
| --- | --- | --- |
| **Security measure** | | **Details (leave blank if not applicable)** |
|  | Encryption | End-to-end (TLS 1.2) and at rest encryption (AES-256) |
|  | Password protection |  |
|  | Role based access controls (RBAC) | Heidi determines the type and level of access granted to individual users based on the "principle of least privilege." This principle states that users are only granted the level of access absolutely required to perform their job functions. Permissions and access rights not expressly granted shall be, by default, prohibited. Heidi’s primary method of assigning and maintaining consistent access controls and access rights shall be through the implementation of Role-Based Access Control |
|  | Restricted physical access |  |
|  | Business continuity plans |  |
|  | Security policies | To ensure continuous compliance with GDPR, Heidi utilise Vanta, a third-party continuous auditing platform. Vanta provides real-time monitoring and assessment of Heidi’s compliance frameworks. The continuous nature of this monitoring ensures that we remain vigilant and responsive to any changes in compliance requirements, thereby safeguarding patient and clinician data with the highest standards of privacy and security. |
|  | Other | Heidi’s Personally Identifiable Information (PII) detection model is evaluated quarterly to assess the accuracy and effectiveness of de-identification methods applied within Heidi's processing of medical transcripts. Heidi utilises a fine-tuned Bidirectional Encoder Representations from Transformers (BERT) model specialised for PII detection. Their systems show high levels of precision and recall, crucial for maintaining privacy standards. The latest performance metrics (September 2024) demonstrate an overall precision of 94.38%, a recall of 96.63%, and an F1 score of 95.49%  Pseudonymization is done by first identifying sensitive data types, including transcripts, patient information, clinician notes, and generated notes. Sensitive data is encrypted both while in transit, and at rest, and all keys are managed securely. Machine learning (ML) techniques are used to de-identify transcript data, targeting entities like names, genders, addresses, emails, and phone numbers This is done by replacing identifying fields within a data record with artificial identifiers, or pseudonyms. For example, instead of storing a person's name, the data might store a unique code that only authorized personnel can trace back to the original individual. The key to re-identify the data is kept separately and securely, ensuring that even if the pseudonymized data is accessed, the privacy of individuals remains protected. This technique is crucial for maintaining privacy while still allowing data to be used for analysis and improving ML accuracy |

# Measures to ensure data is only used for the identified reasons

|  |  |  |
| --- | --- | --- |
| **Security measure** | | **Details (leave blank if not applicable)** |
|  | Contract | Between Heidi and data controller |
|  | Data processing agreement | Heidi enter into strict data processing agreements with all third parties involved in the handling of data (e.g cloud hosting provider for storing transcripts and clinical notes). These agreements are designed to ensure that no user data can be accessed, used, or stored by third parties beyond what is necessary for the specific purpose for which it was shared. They enforce zero retention policies with all third-party service providers. This means that after the necessary data processing tasks are completed, no data is retained, ensuring that information cannot be reused or accessed for any other purpose. |
|  | Data sharing agreement |  |
|  | [Data sharing and processing agreement (DSPA)](https://transform.england.nhs.uk/information-governance/guidance/universal-ig-templates-faqs/) |  |
|  | Audit |  |
|  | Staff training | All Heidi Health employees have full IG training.  If a Heidi Health employee, contractor, user, or customer becomes aware of an information security event or incident, possible incident, imminent incident, unauthorised access, policy violation, security weakness, or suspicious activity, then they shall immediately report the information to the Compliance Lead or CTO. All relevant authorities, and stakeholders will be notified of the privacy breach as soon as possible.  Where applicable, in the event of a personal data breach that meets the criteria for mandatory notification under the UK General Data Protection Regulation (UK GDPR), Heidi Health will report the incident to the Information Commissioners Office (ICO) within 72 hours of becoming aware of the breach. |
|  | Other |  |

# Data Retention

No audio is ever stored during clinical encounters (patient-clinician conversations); the audio stream is processed in real-time and immediately discarded. For dictation recordings (clinician audio only, captured outside of the clinical encounter), clinicians can opt-in to save their dictated audio for later review. It is important to note that the clinician is responsible for ensuring that no patient audio is included in any dictations. Temporary audio storage varies by access method:

* Mobile app: Secured in sandboxed environment until processed
* Desktop app: Protected in secure storage until processed
* Web browser / Chrome extension / EHR Integration: Direct streaming without storage

Clinicians retain control of all transcripts, clinical notes, and clinical documents and can decide how long this data is stored in the cloud up to a maximum of 28 days. Additionally, the patient information contained in these transcripts and clinical notes/documents will only be accessed externally for the purpose of troubleshooting with the express permission of clinicians.

**Retention Policy**

|  |  |  |
| --- | --- | --- |
| **Action** | | **Details** |
|  | Secure destruction (for example by shredding paper records or wiping hard drives with evidence of a certificate of destruction) | The system will automatically delete all data within the time frame specified by the data controller noting this cannot exceed 28 days, usually within 24 hours. |
|  | Permanent preservation by transferring the data to a Place of Deposit run by the National Archives |  |
|  | Transfer to another organisation |  |
|  | Extension to retention period |  |
|  | It will be anonymised and kept |  |
|  | The controller(s) will manage as it is held by them |  |
|  | Other |  |

# Complying with individual rights

|  |  |  |
| --- | --- | --- |
| **Individual right** | **How you will comply** | |
| **The right to be informed**  The right to be informed about the collection and use of personal data. |  | Patients will be informed about the use of data for Heidi Health via the below methods: |
|  | Practice privacy notices |
|  | Patients will be able to ask any questions when consent is obtained prior to the commencement of consultation.recording. |
| **The right of access**  The right to access details of data use and receive a copy of their personal information - this is commonly referred to as a subject access request. | Not applicable | |
| **The right to rectification**  The right to have inaccurate personal data rectified or completed if it is incomplete. | Before being inserted into the patient record the clinician is required to check the data for accuracy. Any errors the are identified by individuals will be managed through the usual practice policies on record rectification. | |
| **The right to erasure**  The right to have personal data erased, if applicable. | Not applicable | |
| **The right to restrict processing**  The right to limit how their data is used, if applicable. | To prevent function creep, Heidi’s data processing activities are strictly governed by clearly defined use cases and regular reviews. Any expansion of data usage or new features undergoes a rigorous evaluation process to ensure alignment with the original purpose and compliance with regulatory requirements. | |
| **The right to data portability**  The right to obtain and re-use their personal data, if applicable. | Patients can make SARs as per usual practice policies. | |
| **The right to object**  The right to object to the use and sharing of personal data, if applicable | Patients will give consent before clinicians use Heidi Health. They can refuse at this time and their care will be un-affected by their decision. | |

**National Data Opt Out**

|  |  |
| --- | --- |
|  | Yes |
|  | No as patient can refuse recording if they wish. The data obtained is subject to the national data opt out only once entered into the patient record |
|  | Unsure |

# Automated decision making

|  |  |
| --- | --- |
|  | Yes |
|  | No. The data is transcribed via automated software but this is checked by a human before being entered into the patient record. |

# Data Controllers

The data controller is the individual practitioner or if on behalf of a practice who has signed this document.

# Data Processor.

Heidi Health Ltd is the Data processor.

Heidi Health works with carefully selected third-party service providers (sub-processors) to provide specific services necessary for their operations. They enter into strict data processing agreements (DPA) with these sub-processors to ensure that they process data only as instructed by Heidi Health and in accordance with their privacy and security requirements. These DPAs stipulate that their sub-processors do not retain data after the necessary processing has occurred, and do not use your data for any secondary purposes. A complete and up-to-date list of Heidi’s sub-processors and their activities can be found [here](http://trust.heidihealth.com/subprocessors).

# Due diligence measures

|  |  |  |
| --- | --- | --- |
| **Due diligence measures** | | **Details (leave blank if not applicable)** |
|  | Data Security and Protection Toolkit (DSPT) compliance | HHA001 11/4/2024 |
|  | Registered with the Information Commissioner’s Office (ICO) | **Registration reference:**  ZB671518  **Date registered:**  18 March 2024  **Registration expires:**  17 March 2026 |
|  | Digital Technology Assessment Criteria (DTAC) assessment |  |
|  | Stated accreditations | DCB0129 |
|  | Cyber Essentials or any other cyber security certification | Date of certification2024-04-08  Issued on2024-04-08  Certificate ID  0fd4036c-35bd-4493-b704-95adc158e165 |
|  | Other checks |  |

# Risks and Mitigations

|  |  |  |  |
| --- | --- | --- | --- |
| **Source of Risk** | **Likelihood of harm** | **Impact of harm** | **Overall risk** |
| 1: Heidi Health systems data is breached by unauthorised persons | 3 | 3 | 9 |
| 2: Heidi Health data is breached due to human error and/or misunderstanding of company requirements | 3 | 4 | 12 |
| 3: Heidi Health staff are not aware of security threats and best practices resulting in a compromise of company systems and data | 3 | 4 | 12 |
| 4: Heidi Health systems and data are breached in transit due to improper encryption | 2 | 4 | 8 |
| 5: Heidi Health systems and data are breached in a non-production environment | 2 | 4 | 8 |
| 6: Heidi Health is unavailable during critical times | 3 | 2 | 6 |
| 7: Speech-to-text model fails to accurately transcribe medical terminology, drug names etc | 3 | 3 | 9 |
| 8: Public facing Privacy Notice is inaccurate, omits critical information, and/or is out of date. | 3 | 3 | 9 |
| 9: Consent for processing of PID is not captured and can't be demonstrated when needed. | 2 | 4 | 8 |
| 10: Temporary audio storage in mobile/desktop apps is compromised before processing completion | 2 | 4 | 8 |
| 11: Cross-platform data synchronization leads to data inconsistency or loss | 2 | 3 | 6 |
| 12: Mobile app network resilience features result in extended local storage of sensitive data | 2 | 4 | 8 |
| 13: Chrome extension enables unauthorized access to clinical data through browser vulnerabilities | 2 | 5 | 10 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Risk** | **Options to reduce or eliminate risk** | **Likelihood of harm** | **Residual risk** | **Measure approved by DPO** |
| 1 | Strict physical access controls and multi factor authentication | 1 | 3 |  |
| 2 | Regular policy and security training | 1 | 4 |  |
| 3 | Regular policy and security training | 1 | 4 |  |
| 4 | Adequate encryption process in place | 1 | 4 |  |
| 5 | Strict data access controls and admin roles | 1 | 4 |  |
| 6 | Redundant systems | 2 | 4 |  |
| 7 | Using speech to text model trained on medical data | 2 | 6 |  |
| 8 | Internal systems to ensure public facing Privacy Policy is regularly updated | 1 | 3 |  |
| 9 | Consent pop-up is enabled | 1 | 4 |  |
| 10 | Implement secure enclave storage and automatic purging protocols | 1 | 4 |  |
| 11 | Real-time synchronization monitoring and verification | 1 | 3 |  |
| 12 | Implement strict timeout policies and storage limits | 1 | 4 |  |
| 13 | Conduct periodic audits to verify that consent documentation and data deletion protocols are strictly followed. | 1 | 5 |  |

**\*Risk scoring table**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | **Impact (I)** | | | | |
| **Negligible (1)** | **Low**  **(2)** | **Moderate (3)** | **Significant (4)** | **Catastrophic (5)** |
| **Likelihood** | **Rare (1)** | **1** | **2** | **3** | **4** | **5** |
| **Unlikely (2)** | **2** | **4** | **6** | **8** | **10** |
| **Possible (3)** | **3** | **6** | **9** | **12** | **15** |
| **Likely (4)** | **4** | **8** | **12** | **16** | **20** |
| **Almost certain (5)** | **5** | **10** | **15** | **20** | **25** |

Review and sign-off

|  |  |  |
| --- | --- | --- |
| **DPO Approver sign-off** | | |
| Reviewer name: | Mr Paul Couldrey | |
| Reviewer job title: | Data Protection Officer | |
| Reviewer contact details: | [**paul.couldrey@nhs.net**](mailto:paul.couldrey@nhs.net) | |
| Date of review: |  | |
| Comments: |  | |
| Date for next review: |  | |
| **Data Controller** | | |
| Name: | Dr Paul Rwezaura | |
| Job title: | GP Partner | |
| Complete if signing on behalf of an organisation : | Name of Organisation | Role in Organisation |
| Date of acceptance of risks/responsibilities: |  | |
| Comments: |  | |