CURISTICA

HANLEY CONSULTING

Surgery Assist V5 Data Protection Impact Assessment (DPIA)

Version: 1.0

Date: 12/06/2025

Status: PUBLISHED

Director: Sharon Hanley

Document Management

Revision History

Version	Date	Summary of Changes
0.1	10/06/2025	First Draft
0.2	11/06/2025	Second Draft - AI section added
1.0	12/06/2025	First release

Authors

Name	Title/Responsibility	Date	Version
Curistica Ltd	Data Protection Officer	12/06/2025	1.0

Reviewers

This document must be reviewed by the following people:

Name	Title/Responsibility	Date	Version
Sharon Hanley	Commercial Director	12/06/2025	1.0
Max Gattlin	Product Director	12/06/2025	1.0
Dr Keith Grimes	Chief Digital Health Officer & CSO	12/06/2025	1.0

Approved by

This document must be approved by the following people:

Name	Title/Responsibility	Date	Version
Curistica Ltd	Data Protection Officer	12/06/2025	1.0

Table of contents

SECTION 1 – Screening

1. Is a DPIA required?

a. Summary of how data will be used and shared

Surgery Assist V5 is a web-based digital assistant that helps users navigate non-clinical administrative healthcare tasks such as booking appointments, accessing pharmacy services, and using tools like the NHS App. It operates 24/7 and improves both patient independence and provider efficiency.

The platform offers two modes:

- **"Decision Tree" Assistant** – A structured, option-based tool guiding users through predefined decision trees based on established administrative workflows.
- **"AI" Assistant** – An optional, LLM-powered "AI" interface that responds to free-text queries using only verified information from the decision tree system. While responses are not fully testable due to their probabilistic nature, they are safety-assured under DCB0129 guidelines and undergo extensive testing before deployment.

The AI assistant is not the default and requires deliberate user activation. Available services may vary by GP practice.

For the purposes of this DPIA, we will only consider the automatic collection of data.

The collection of personal data through the feedback mechanisms, which require explicit consent of the user, has been assessed and deemed to be small-scale with limited processing, therefore not requiring a formal DPIA. Further rationale is provided in Appendix A.

b. Description of the data

☐	Personal data
☐	Pseudonymised data
☑	**Anonymous data**

SECTION 2 – Data purpose, use and benefits

2. What are the purposes for using or sharing the data?

The data collected is used for:

1. **Providing the service:** Surgery Assist requires processing of user inputs, which may include personal and/or health data for proper functioning and to provide benefit for users.
1. **Risk Management and Safety Assurance:** In line with DCB0129/0160, data processing is required for the purposes of assuring and reviewing the digital clinical safety profile of the service and executing responsibilities in responses to a safety incident as laid out in the Clinical Risk Management Plan. Data is required to be processed to identify any instances where hazards identified in the hazard log may have occurred, either identified before (proactive) or after (retroactive) has occurred.
    a. For the decision tree assistant this includes but not limited to identifying pathway failures and technical issues/bugs.
    b. For the AI assistant this includes but is not limited to continuously evaluating the relevance and accuracy of the LLM's retrieval against approved knowledge bases, identifying model inaccuracies, hallucinations, or inappropriate matches.
2. **System Stability and Performance Monitoring**:
    a. Monitoring the functionality, stability, and availability of the digital assistant and the integrated Large Language Model (LLM) feature. This ensures that the service operates reliably across user sessions and devices.
    b. Identify and rectify bugs, technical issues, or operational errors.
3. **Product Improvement and Optimisation:**
    a. Analysing trends in user queries and model outputs to refine the knowledge base accessed by the decision tree and LLM components (e.g., update or remove outdated material).
    b. Improve the natural language search capabilities to better serve user needs within the approved administrative scope.
    c. Enhance the user interface and experience by understanding user preferences and query patterns.
4. **Usage Analytics:** Aggregated analysis of how users interact with the service (e.g., common query types, success rates) to inform service reporting and strategic improvements. This includes understanding how users transition between structured menus and free-text search modes.
5. **Product support:** If errors or complaints are raised, historical prompt-response logs can be used to investigate the incident, provide meaningful support to clients and evidence compliance with service

guarantees and data protection principles.

3. What are the outcomes and benefits for individuals and society?

Non-clinical administrative tasks constitute a substantial burden on the clerical capacity of healthcare providers when performed over the phone or in person. These include but are not limited to appointment booking, making referrals, repeat medication enquiries and checking of test results. The demand for these services is high and is often a source of poor user experience with healthcare services.

Established digital pathways exist for many of these administrative tasks which allow them to be completed by users/patients themselves, without requiring interaction with provider staff, yet many users/patients are either unaware they can, do not know how to access or complete the task digitally.

Surgery Assist surfaces, highlights and guides users to accomplish administrative healthcare tasks via established and available digital pathways through an interactive digital assistant model, helping them to setup, access and interact with the digital services available to them to accomplish their tasks. In doing so, Surgery Assist not only helps users, but also helps healthcare providers manage patient demand more efficiently

Services which users can access digitally vary by practice, but usually include online appointment booking, accessing local community and pharmacy services, access (via drop-off) to digital healthcare products (NHS App, online consultations, symptom checkers), all available 24/7,

Surgery Assist is integrated into healthcare providers cloud telephone systems, websites, and via QR codes displayed on posters and on waiting room display screens.

Further information is available at the product website; Surgery Assist (formerly Surgery Assist) - The #1 Automated Assistant for General Practice.

SECTION 3 – Data types, sources and linkage

4. Can anonymous data be used? *If not, explain why*.

☑	**Yes**
☐	No
☐	Unsure

5. Which types of personal data are collected and for what purpose?

☐	Forename	☐	Physical description, for example height	☐	Photograph / picture of people
☐	Surname	☐	Phone number	☐	Location data
☐	Address	☐	Email address	☐	Audio recordings
☐	Postcode full	☑	GP details: • ODS code	☐	Video recordings
☐	Postcode partial	☐	Legal representative name (personal representative)	☑	Other: • Device • OperatingSystem • Browser
☐	Date of birth	☐	NHS number	☐	None
☐	Age	☐	National insurance number		
☐	Gender	☐	Other numerical identifier		

- **Dataset structure and data captured, including purpose outlined in:**
  - Decision Tree Model: Appendix B
  - AI Model: Appendix C

6. Which types of Special Category data are collected and for what purpose?

Type of data		Purpose
☑	**Information relating to an individual's physical or mental health or condition, for example information from health and care records**	This is required for the functioning of the AI digital assistant in order to respond to user queries. The AI digital assistant records logs of all interactions with users as part of our commitments as outlined in Section 2. It is reasonably expected therefore that health data may be captured if a user enters it during their interaction with the digital assistant, for example "I have back pain". This prompt and consequent response is then recorded in the logs. This data however would have no direct personal data counterpart (e.g. no name, address, contact details etc..) associated with it and it would not be possible to identify an individual from the high level data (browser type, device type and ODScode) collected.

☐	Biometric information in order to uniquely identify an individual, for example facial recognition	
☐	Genetic data, for example details about a DNA sample taken as part of a genetic clinical service	
☑	**Information relating to an individual's sexual life or sexual orientation**	The AI digital assistant records logs of all interactions with users as part of our commitments as outlined in Section 2. It may be expected therefore that information relating to an individual's sex life can either directly or indirectly be recorded if they enter a prompt relating to services offered – e.g. "I think I have an STI" or "What services are there for trans women?". This prompt and consequent response is then recorded in the logs. This data however would have no direct personal data counterpart (e.g. no name, address, contact details etc..) associated with it and it would not be possible to identify an individual from the high level data (browser type, device type and ODScode) collected.
☐	Racial or ethnic origin	
☐	Political opinions	
☐	Religious or philosophical beliefs	
☐	Trade union membership	
☐	Information relating to criminal or suspected criminal offences	
☐	None of the above	

7. Who are the data subjects that can be identified from the data?

☐	Patients or service users
☐	Carers
☐	Staff:
☑	**Wider workforce:** whilst individual staff cannot be identified, the GP practice is identifiable and therefore the workforce at that practice may be (particularly if it is a small practice).
☐	Visitors
☐	Members of the public
☐	Other

8. Where will your data come from?

Data is collected from the interaction of service users with the Surgery Assist service.

9. What is the expected volume of data?

The expected data volume is dependent on practice size. The current average demand is 30 users per 1,000 registered patients per practice per month.

10. Is data being linked?

☐	Yes
☑	**No**
☐	Unsure

a. As a result of linkage will it become possible to identify individuals who were not identifiable in the original dataset?

☐	Yes
☐	No
☑	**N/A**

SECTION 4 – Use of Artificial Intelligence (AI)

11. Is an AI system being used to collect or process data?

In this scenario the term “AI” is an umbrella term for a range of algorithm based technologies that solve complex tasks by carrying out functions that previously required human thinking.

☑	**Yes**
☐	No
☐	N/A

12. What type(s) of AI system is being used?

☐	Machine Learning (ML)
☑	**Natural Language Processing (NLP)**
☐	Computer Vision
☐	Robotics and Control Systems
☐	Speech Recognition and/or Speech Generation
☑	**Generative AI**

13. Why is AI being used - what is the unique capability that it provides which makes it beneficial for users?

The use of AI allows for natural language processing of free-text user inputs, thus allowing them to directly access relevant materials without the need to step through multiple decision tree flows, saving them time and resulting in a quicker outcome.

The Generative AI component allows for responses to be tailored to the input, thus making the experience more personal and impactful.

In addition, the use of AI also allows interactions with the digital assistant in other languages, thus opening access to users who may otherwise have not been able to use the digital assistant.

14. Are you using a commercially available Foundation Model or a custom model?

Please provide information on the source of training data, collection method, quality assessment and any processing done to address quality issues.

☐	Commercially available foundation model with no changes.
☑	**Commercially available foundation model with changes:** Microsoft OpenAI Azure Foundry model. The original training data is unknown as this is commercially sensitive material. For the purposes outlined in this DPIA, the model is heavily constrained and only permitted to provide responses from data provided to it by the specific practice it is being deployed in. This data has been vetted and provided by the practice and then further reviewed by Hanley following a standard operating process.
☐	Custom Model:

15. What output is the AI expected to provide?

☐	Prediction
☑	**Recommendation:** The AI digital assistant provides a recommendation to the user based on their input. The recommendation can only be from a previously defined, vetted and quality assured list as well as any additional options provided by the practice - those options would be equally assured.
☐	Classification
☐	Transformation (e.g. Voice to Text):
☐	Other

16. What impact is the output expected to have on the individual?

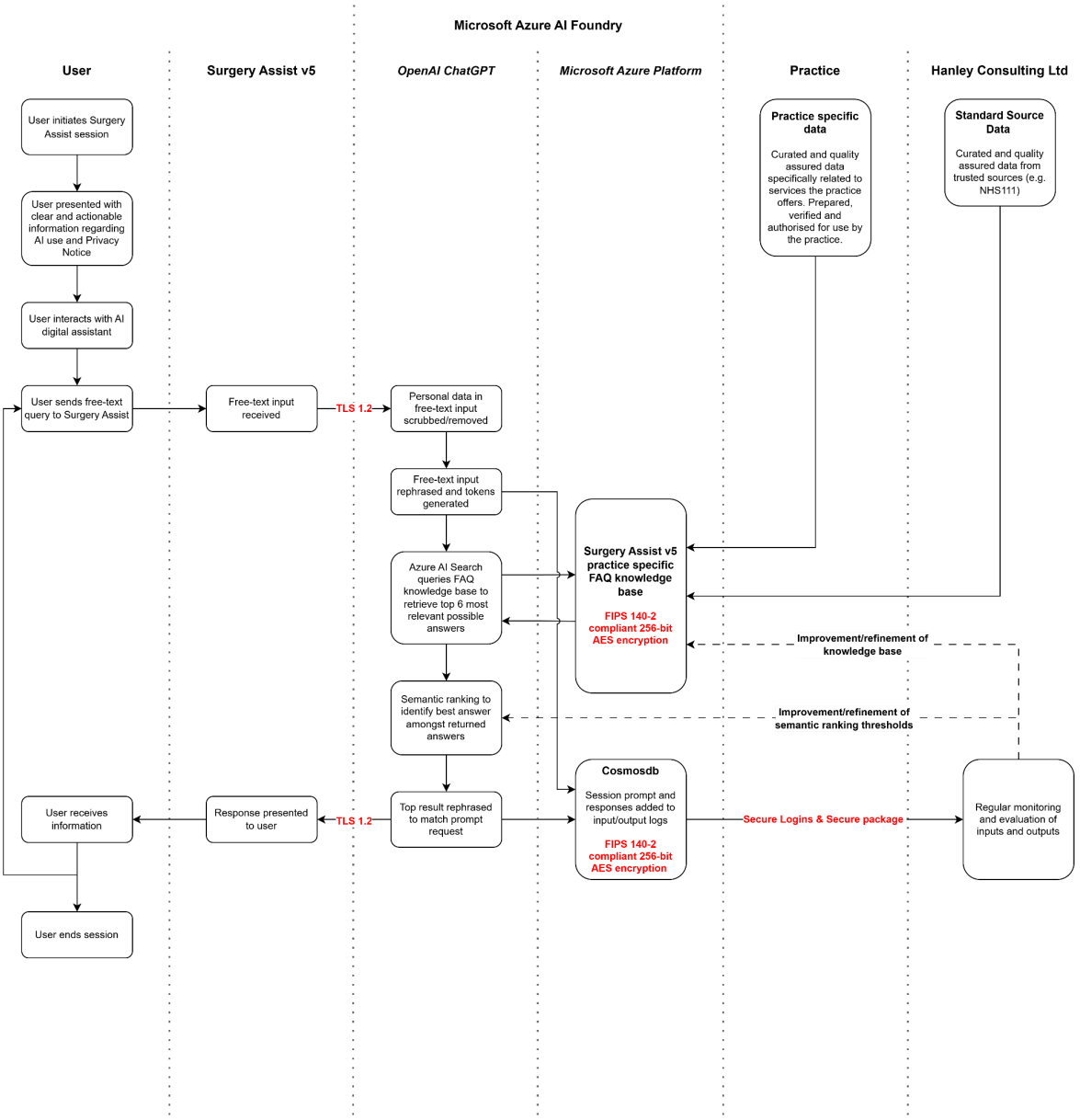
☐	Catastrophic - relating to life and death decisions
☐	Significant - legal effect or impacting someones liberty/rights
☐	Moderate - denial of service, targeting of a message
☑	**Low - direction to information or action:** The user is provided with a recommendation to access a service - the user can then judge themselves whether or not to undertake the recommended action or not. The user *always* has the option to contact the practice and speak to someone, therefore the overall risk is considered low/very low.
☐	Other

17. Is the output fully automated or does it have human review?

☑	**Fully automated**
☐	Human supervised / Human-in-the-loop

☐	Not applicable

18. Rationale Explanation: summarise how the AI functions in terms of the inputs, processing and outputs.



The flow is outlined graphically above.

The user enters their query into the chat window with the AI assistant, this may be something like “I have back pain” or “I need my prescription refilled”.

The digital assistant will then send this query to the OpenAI model which will analyse and interpret the request and determine the most likely question which needs to be answered. Any personal data is stripped out of the input prompt at this point.

The AI will then review the FAQ Knowledge Base available to it to find the 6 most appropriate answers relating to the query.

The knowledge base has been built from quality assured standard source data (from Hanley Consulting Ltd) and documents/data provided by the client (the GP practice) directly.

Once the results are returned, they are semanticly ranked to determine which most closely matches with the request the user made. This is explored further at Semantic ranking - Azure AI Search | Microsoft Learn.

Responses are scored from 0 to 4, where the higher number indicates a closer match.

If a score is greater than 1.4 it is deemed to be relevant and provided to the end user.

If the best available score is ranked below 1.4 semantic rank, then the user is prompted to redefine their input.

Full logic is provided in Appendix D.

All input/output pairs are logged, anonymously, and are regularly reviewed as part of our monitoring process to continuously ensure the safety and effectiveness of the AI digital assistant.

19. **Responsibility explanation:** who is involved in the development, management and implementation of an AI system, and who to contact for a human review of a decision.

Hanley Consulting Ltd own and manage the AI and lead the technical development of the digital assistant.

Curistica Ltd led the clinical safety and data privacy aspects of the AI digital assistant and provide continuous oversight in terms of data collection, processing, clinical risk and incident management/response.

Curistica Ltd provide Clinical Safety Officer and Data Protection Officer services to Hanley Consulting and have been embedded within the team from very early in the product lifecycle providing input into both clinical risks and their management as well as data protection considerations.

Microsoft Azure hosts the AI digital assistant.

Any queries regarding outputs should be directed to support@hanleyconsulting.co.uk.

20. **Data explanation:** What data is being used and how is it used to determine the output?

The flow is outlined in Section 5, please also see response to Q18.

21. **Fairness explanation:** What are the steps taken across the design and implementation of an AI system to ensure that the decisions it supports are generally unbiased and fair, and whether or not an individual has been treated equitably?

The AI model is agnostic with respect to any demographic or social characteristic, that is, even if that data is put into the input “e.g. I am a black Caribbean man seeking….” that additional piece of information is not considered when generating responses. In addition, whilst we are unable to comment on the original data on which the Foundation Model was trained, we can be assured that the data on which the AI acts on, i.e. the “standard” repository and the practice specific content has been curated in a manner that is “fair”. Testing done during pre-deployment has not identified any difference in response when individuals add in demographic or social information which may lead to discrimination(i.e. protected characteristics).

22. **Safety and Performance explanation:** What are the steps taken across the design and implementation of an AI system to maximise the accuracy, reliability, security and robustness of its decisions and behaviours?

Surgery Assist v5 has DCB0129 compliance in line with NHS Standards. What this means is that it has passed a comprehensive risk management process which, following a stepwise approach, has identified, assessed and mitigated (where appropriate) any safety and reliability risks during its development, this includes where applicable to the AI element of the platform. This process has been further validated by extensive testing including input/output pairs in order to ensure that responses from the AI are safe and accurate. In addition, Hanley Consulting Ltd has committed to continued monitoring and improvement of Surgery Assist following deployment which includes, but is not limited to, continued review and evaluation of the input/outputs from the AI system and tuning/refining the model performance. These tasks are more fully outlined in our Clinical Safety Case and within this DPIA.

23. **Impact explanation:** What are the steps taken across the design and implementation of an AI system to consider and monitor the impacts that the use of an AI system and its decisions has or may have on an individual, and on wider society? Specifically address the following:

    - **Allocative harms:** result of a decision to allocate goods and opportunities among a group. The impact of allocative decisions may be loss of financial opportunity, loss of livelihood, loss of freedom, or in extreme circumstances, loss of life.
    - **Representational harms:** occur when systems reinforce the subordination of groups along identity lines. For example, through stereotyping, under-representation, or denigration, meaning belittling or undermining their human dignity.

Surgery Assist does not, through its function, produce outputs which may cause allocative or representational harms; either through intended use or reasonably expected unintended use.

This is by design, the system is socially and demographically agnostic; that is to say it does not, ever, consider social or demographic characteristics when generating outputs, even when those details are included in the input - this has been further explored above.

Nevertheless, the system is monitored regularly, as part of the ongoing post-deployment evaluation process to ensure that there is no deviation from this position.

24. How are users being informed and given access to information about the use of AI?

Information should include:

- The use of AI.
- The existence for automated decision making if it is producing legal or similarly significant effects.
- meaningful information about the logic involved; and
- the significance and envisaged consequences for the individual.

Surgery Assist has two modes, a decision tree and AI mode.

The decision tree mode is the default. Users must actively choose to use the AI mode.

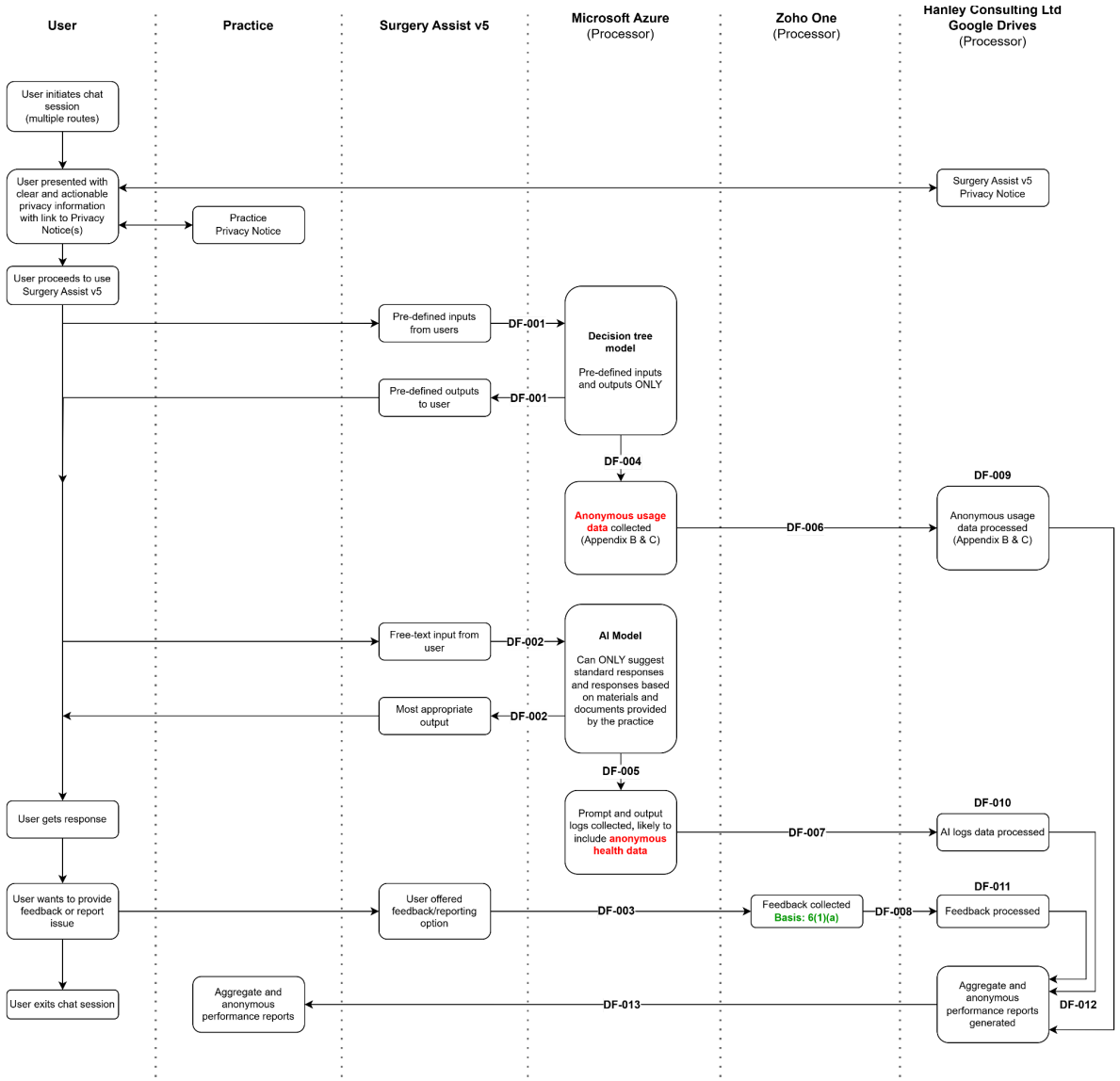
Before users are provided access they are informed through on-screen prompts that they will be interacting with an AI agent.

The Privacy Notice and the DPIA will also be publicly available for users to review if they wish to understand more about the logic, with links to the former through the Surgery Assist interface. The DPIA link will be embedded within the Privacy Notice.

Given the very low risk of impact derived from the use of AI, this is deemed a reasonable approach to satisfy users' Right to be Informed.

SECTION 5 – Data flows

25. Describe the flows of data.



26. Confirm that your organisation's information asset register (IAR), record of processing activities (ROPA) or your combined information assets and flows register (IAFR) has been updated with the flows described above.

☑	**Yes**
☐	No
☐	Unsure

27. Is data being shared internationally?

☐	Yes
☑	**No** - all storage and processing servers are within the UK.
☐	Unsure

a. If yes, give details, including any adequacy assessments, safeguards or measures put in place to protect the data whilst outside of the UK.

SECTION 6 – Intended use and legal basis

28. Under Article 6 of the UK General Data Protection Regulation (UK GDPR) what is the lawful basis for processing personal data?

☐	(a) We have consent
☐	(b) We have a contractual obligation
☐	(c) We have a legal obligation
☐	(e) We need it to perform a public task
☐	(f) We have a legitimate interest
☑	**Other: Not applicable, no personal data captured.**

29. If special category data is being collected and processed, what is the lawful basis under Article 9 of the UK GDPR?

☐	(a) We need it to comply with our legal obligations for employment
☐	(b) We need it for legal claims, to seek legal advice or judicial acts
☐	(c) We need to comply with our legal obligations to provide information where there is a substantial public interest, as set out in this list
☐	(d) We need it to comply with our legal obligations to provide or manage health or social care services
☐	(e) We need it to comply with our legal obligations for public health
☐	(f) We need it for archiving, research and statistics where this is in the public interest
☐	Other
☑	**Not applicable**

30. What is the legal basis for using and sharing this health and care data under the common law duty of confidentiality?

☐	Implied consent
☐	Explicit consent
☐	Section 251 support
☐	Legal requirement
☐	Overriding public interest
☑	**Not applicable**

a. Please provide further information or evidence.
N/A

SECTION 7 – Data storage and security

31. Is information being collected?

☑	**Yes**
☐	No

a. How is the data being collected?
Data is collected automatically by the platform (Microsoft Azure) during use by the end-user and through interaction of the user with Surgery Assist v5.

32. Is information being stored?

☑	**Yes**
☐	No

a. How will information be stored and kept secure?

Asset number	Asset description	Controller	Data	Geographic location	Storage Location	Retention Period	Security Measures	Update Frequency
IA-001	Anonymised usage data	Hanley Consulting Ltd	GP/hospital details, Other	In the UK	Hanley Consulting Ltd owned Microsoft Azure environment	3 years	Role based access controls (RBAC), Encryption at rest, Password protection, Multi-factor authentication	Real-time
IA-002	Log of interaction of user with AI digital assistant including input prompts and output responses.	Hanley Consulting Ltd	GP/hospital details, health, sexual activity, sexual orientation	In the UK	Hanley Consulting Ltd owned Microsoft Azure environment	3 years	Role based access controls (RBAC), Encryption at rest, Password protection, Multi-factor authentication	Real-time
IA-003	Feedback from users using the Surgery Assist platform	Hanley Consulting Ltd	Name, Contact details	In the UK	Hanley Consulting Ltd owned Zoho One environment	3 years	Role based access controls (RBAC), Password protection	Real-time

33. Is information being transferred?

☑	**Yes**
☐	No

a. How will information be transferred and used for the purposes outlined in Section 1?

Flow number	Short description	How is the data transferred?	How often is the data transferred?	What technical measures are in place to protect the data flow?	What organisational measures are in place to protect the data?
DF-001	Two-way interaction of User with Surgery Assist v5 decision tree model	Electronic upload/download	Real-time	Encrypted transfer	Contract
DF-002	Two-way interaction of user with Surgery Assist v5 OpenAI model	Electronic upload/download	Real-time	Encrypted transfer	Contract
DF-003	Patient Feedback on Surgery Assist v5	Electronic upload/download	Real-time	Encrypted transfer	Contract
DF-004	Processing of anonymous usage data	Electronic upload/download	Real-time	Encrypted transfer	Contract
DF-005	Processing of user interaction with OpenAI model and creation of logs	Electronic upload/download	Real-time	Encrypted transfer	Contract
DF-006	Transfer of anonymous usage data to Hanley Consulting Ltd Google Drives	Electronic upload/download	Weekly	Securely packaged, Secure logins	Contract
DF-007	Transfer of anonymous AI processing logs to Hanley Consulting Ltd Google Drives	Electronic upload/download	Weekly	Securely packaged, Secure logins	Contract
DF-008	Transfer of collected feedback to to Hanley Consulting Ltd Google Drives	Electronic upload/download	As required	Secure logins, Securely packaged	Contract
DF-009	Processing of anonymous usage data by Hanley Consulting Ltd	Electronic upload/download	Weekly	Secure logins, Securely packaged	Contract
DF-010	Processing of anonymous AI processing logs by Hanley Consulting Ltd	Electronic upload/download	Weekly	Secure logins, Securely packaged	Contract
DF-011	Processing of collected feedback by Hanley Consulting Ltd	Electronic upload/download	As required	Secure logins, Securely packaged	Contract
DF-012	Creation of aggregate and anonymous performance reports	Electronic upload/download, Email	Weekly	Securely packaged, Secure logins	Contract
DF-013	Transmission of performance reports to Practice	Electronic upload/download, Email	Monthly	Secure logins	Contract

SECTION 8 – Data retention and deletion

34. How long will the data be used for?

Usage metrics: 3 years
AI logs: 3 years
Feedback: 3 years

35. How long will the data be retained for?

Usage metrics: 3 years
AI logs: 3 years
Feedback: 3 years

36. What will happen to the data at the end of this period?

Action		Details
☑	**Secure destruction (for example by shredding paper records or wiping hard drives with evidence of a certificate of destruction)**	Data processors (Microsoft Azure, Google Drive, Zoho One)
☐	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	

SECTION 9 – People's rights and choices

37. How will individual rights be complied with?

Individual right	How you will comply (or state *not applicable* if the right does not apply)	
The right to be informed The right to be informed about the collection and use of personal data.		We have assessed how we should inform individuals about the use of data for [Surgery Assist . We consider the communications methods below meet this obligation because of the nature of the interaction with the service, the expectation of the user with respect to the manner they would be informed and the necessity given the likely impact.
	☑	Privacy notice(s) for all relevant organisations - available on Privacy Policy - Hanley Consulting Ltd
	☐	Information leaflets
	☐	Posters
	☐	Letters
	☐	Emails
	☐	Texts
	☐	Social media campaign
	☑	DPIA published (best practice rather than requirement)
	☐	Other
	☐	Not applicable
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.	No personal information collected. Any information collected cannot be attributed to any identifiable individual either, however a chat *session* belonging to a user may be retrievable. Where a user provides sufficient information for us to reasonably identify and authenticate their session, we will consider requests for access, rectification, or erasure. Where this is not possible and the data cannot be linked to an individual, these rights may not apply (in line with Article 11 UK GDPR). Information required for such an action must include at a minimum - Exact date and time of session/conversation - ODScode - Inputs made into the model (e.g. text input)	

The right to rectification The right to have inaccurate personal data rectified or completed if it is incomplete.	No personal information collected. Any information collected cannot be attributed to any identifiable individual either, however a chat *session* belonging to a user may be retrievable. Where a user provides sufficient information for us to reasonably identify and authenticate their session, we will consider requests for access, rectification, or erasure. Where this is not possible and the data cannot be linked to an individual, these rights may not apply (in line with Article 11 UK GDPR). Information required for such an action must include at a minimum - Exact date and time of session/conversation - ODScode - Inputs made into the model (e.g. text input)
The right to erasure The right to have personal data erased, if applicable.	No personal information collected. Any information collected cannot be attributed to any identifiable individual either, however a chat *session* belonging to a user may be retrievable. Where a user provides sufficient information for us to reasonably identify and authenticate their session, we will consider requests for access, rectification, or erasure. Where this is not possible and the data cannot be linked to an individual, these rights may not apply (in line with Article 11 UK GDPR). Information required for such an action must include at a minimum - Exact date and time of session/conversation - ODScode - Inputs made into the model (e.g. text input)
The right to restrict processing The right to limit how their data is used, if applicable.	Individuals may opt out by not using the service.
The right to data portability The right to obtain and re-use their personal data, if applicable.	No personal information collected. Any information collected cannot be attributed to any identifiable individual either, however a chat *session* belonging to a user may be retrievable. Where a user provides sufficient information for us to reasonably identify and authenticate their session, we will consider requests for access, rectification, or erasure. Where this is not possible and the data cannot be linked to an individual, these rights may not apply (in line with Article 11 UK GDPR). Information required for such an action must include at a minimum - Exact date and time of session/conversation - ODScode - Inputs made into the model (e.g. text input)
The right to object The right to object to the use and sharing of personal data, if applicable.	Individuals may opt out by not using the service.

38. Will the national data opt-out need to be applied?

☐	Yes
☑	**No**
☐	Unsure

Explanation: No confidential personal information is being collected.

39. Is automated decision making (ADM) being employed?

☐	Yes
☑	**No**
☐	Unsure

a. **Where the effect of the automated decision on the individual is substantial, how will an individual's right not to be subjected to a decision solely made by automated means be upheld?**

b. **Is special category data being used as part of automated decision making?**

☐	Yes
☑	**No**

40. Detail any stakeholder consultation that has taken place (if applicable).

None

SECTION 10 – Other organisations

41. What organisation(s) decide why and how the data is being used and shared (controllers).

Controller	Role
Hanley Consulting Ltd	Controller

42. What organisation(s) are being instructed to use or share the data (processors).

Processor(s)	Role and relationship
Microsoft Azure	Hosts Surgery Assist v5 and collects, processes and stores data from it.
Google Drive	Stores data.
Zoho One	Collects and stores user feedback

43. What organisations have been subcontracted by any processors to handle data:

Sub-processor(s)	Role
N/A	None

44. What due diligence measures and checks have been carried out on any processors used?

Due diligence measures		Details (leave blank if not applicable)
☑	**Data Security and Protection Toolkit (DSPT) compliance**	Microsoft: 8JH14 Google LLC: 8JE14
☐	**Registered with the Information Commissioner's Office (ICO)**	Microsoft: Z6647359 Google LLC: Z6647359
☐	Digital Technology Assessment Criteria (DTAC) assessment	
☐	**Stated accreditations**	

☑	**Cyber Essentials or any other cyber security certification**	Microsoft: SOC1/2/3, Cyber Essentials Plus, G-Cloud, ISO 27001, 27017, 27018, 27701 Google: SOC1/2/3, Cyber Essentials Plus, Cloud Security, ISO 27001, ISO 27017, ISO 27018, ISO 27701
☐	Other checks	

SECTION 11 – Risks and Mitigations

45. Risk assessment table

Risk ref no.	Description	Initial risk score	Mitigations	Residual risk score
01	Loss of Surgery Assist Usage data	2	Regular backups by data processors (in-place already)	1
02	Inadvertent sharing of Surgery Assist usage data	2	Maintain security precautions, encryption and secure packaging of all data transfers (in place already)	1
03	Users not aware of systematic collection of data	4	All clients to be informed of and have clear signposting to privacy policies on their websites where Surgery Assist is being deployed and within the Surgery Assist interaction window (in place already)	2
04	Users entering direct personal data (e.g. name, address, phone number) into AI digital assistant and it being recorded/visible in logs. This then also potentially allows health data to be linked to an individual. **Likelihood**: Moderate	8	Two Mechanisms: - Users are warned before accessing the AI digital assistant that they should not enter personal information and must agree to this before they can proceed. - Microsoft Azure platform has a PII redaction tool which removes direct PII from the user input if detected *before* it is passed onto the LLM for processing. In this way it is not recorded in logs.	2

05	**“Singling out”** - users with unique or very rare health conditions may be identifiable in combination with their ODS code if that health condition is recorded in the AI input/output logs. **Likelihood**: Very low	2	A data subject may become identifiable if an individual has access to the AI logs AND: 1) The individual also has personal knowledge of the data subject, their location and their condition. 2) The data subject has made their condition and location public knowledge. This is a highly unlikely/edge case scenario as both situations require access to the AI logs which are secured by RBAC, secure logins and MFA and restricted to Hanley Consulting Ltd staff only. Realistically only the immediate family, friends and GP practice staff would be in a position to have personal knowledge of the DS and they would not have access to the AI logs. In the second scenario, the data subject has manifestly made that information public and therefore falls under Article(9)(e). Overall, the likelihood of singling out occurring is incredibly small/negligible.	2
06	**“Linkability”** - users may be identifiable via additional data sources **Likelihood**: Very low	1	There is no scenario where an individual may be identified via the content of the recorded data (save for the scenarios identified in Risk Ref 05). However, a data subject may be able to identify their own conversations if they have the following information: - The practice name/ODS code. - Device/Browser details - The date of the session - The exact time (hh:mm:ss) of the session - The content of the session (e.g. what topics were discussed) In this instance a specific conversation record may be identifiable and execution of Rights granted under UK GDPR may be considered. However this may not be practical as the number of concurrent users with the same/similar requests may mean that multiple, rather than a single, matched record is retrieved.	1

Risk scoring table

		Impact (I)				
		Negligible (1)	Low (2)	Moderate (3)	Significant (4)	Catastrophic (5)
Likelihood (L)	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

46. Detail any actions needed to mitigate any risks, who has approved the action, who owns the action, when it is due and whether it is complete.

Risk ref no.	Action needed	Action approver	Action owner	Due date	Status e.g. outstanding/ complete

SECTION 12 – Review and sign-off

Reviewer sign-off	
Reviewer name:	Sharon Hanley
Reviewer job title:	Managing Director
Reviewer contact details:	sharon@hanleyconsulting.co.uk
Date of review:	12/06/2025
Comments:	
Date for next review:	12/06/2026

Approver sign-off	
Approver name:	Curistica Ltd
Approver job title:	Data Protection Officer
Approver contact details:	dpo@curistica.com
Date of approval:	12/06/2025
Comments:	

Appendix A

Rationale for Feedback collection not requiring a full DPIA.

Surgery Assist is a digital assistant for primary care that supports patients to self-serve and carry out administrative activities by signposting individuals to available online services, websites and apps.

The platform does not routinely collect, process or store any personally identifiable data during normal operation and users are generally unable to enter personal data through normal use of the Surgery Assist platform.

The Surgery Assist feedback function is therefore not deemed to require a full DPIA as we do not undertake processing which likely results in high risk to the rights and freedoms of individuals under UK GDPR or under European Guidelines.

Surgery Assist is also not operating as an innovative technology under the ICO's definition and is not collecting novel forms of data.

The only scenario where user personal data may be left is if users inadvertently leave PII or special category data (health) when entering information on the feedback form through a misunderstanding of the purpose of the form.

To mitigate this there are clear notices when using the forms and formal assessment is made within our Clinical Safety Hazard Log.

Users may however choose to leave their contact details (and must explicitly consent to this) if they wish to be contacted by us to go through their feedback with them.

To date, we have no recorded instance of any inadvertent submission of PII or health data via our feedback mechanisms. This is in the context of Surgery Assist recording over 174,000 interactions with clients and receiving 294 individual feedback submissions through our platform.

In summary, the platform is not designed for and we do not collect any personal data routinely. Any personal data collected is either consented or inadvertent, the latter which we have mitigated as best we can. Therefore any unintended data collection or processing of PII or special category data will be small-scale by nature.

This therefore does not constitute a "high risk" to the rights and freedoms of individuals, negating the need for a DPIA.

Appendix B

Decision tree digital assistant dataset

Data	Format	Collection Method	Who collects?	Who is it sent to?
Conversation ID	Numeric	Randomly generated	Microsoft Azure	Hanley Consulting Ltd
Referrer URL	URL	HTTP Header	Microsoft Azure	Hanley Consulting Ltd
Operating System (iOS/Android/Linux/)	String	HTTP Header	Microsoft Azure	Hanley Consulting Ltd
ODS code	Alphanumeric	Chatbot used	Microsoft Azure	Hanley Consulting Ltd
Chat status (open/closed)	Boolean	By Model (Closed)	Microsoft Azure	Hanley Consulting Ltd
Chat started	date-time	API	Microsoft Azure	Hanley Consulting Ltd
Chat closed	date-time	API	Microsoft Azure	Hanley Consulting Ltd
Chat duration	date-time	Calculation	Microsoft Azure	Hanley Consulting Ltd
Chatbot Name	String	Chatbot used	Microsoft Azure	Hanley Consulting Ltd
Timestamp	date-time	Automatically generated	Microsoft Azure	Hanley Consulting Ltd

Appendix C

“AI” digital assistant dataset

MessageLog Table: (Stores details about each individual message exchange)

- **PartitionKey:**
  - **Description:** A key used by Cosmos DB to distribute data across partitions for scalability and performance. Derived from data like ConversationId or UserId (UserNumber).
  - **Capture Method:** Automatically generated based on defined partitioning strategy when writing the record to Cosmos DB.
- **RowKey:**
  - **Description:** A unique identifier for this specific message log entry within its partition in Cosmos DB.
  - **Capture Method:** Automatically generated when writing the record to Cosmos DB.
- **Timestamp:**
  - **Description:** A timestamp automatically generated by Cosmos DB when the record (message log entry) was last modified.
  - **Capture Method:** Automatically generated by Azure Cosmos DB.
- **UserInput:**
  - **Description:** The raw text query or message entered by the user into the chatbot interface.
  - **Capture Method:** Captured directly from the chat input field upon user submission.
- **TranslatedUserInput:**
  - **Description:** The user's input message after being translated (e.g., into English) if the surgery has enabled multi-language interactions and translation is performed. May be null if no translation occurred.

    - **Capture Method:** Generated by Microsoft Azure Language service integrated into the chatbot system, using UserInput as the source.
- **UserLanguage:**
    - **Description:** The detected language of the UserInput (e.g., 'en', 'fr').
    - **Capture Method:** Determined by Microsoft Azure Language service within the application from the UserInput.
- **AssistantAIResponse:**
    - **Description:** The final text response generated by the Surgery Assist system (utilising various Microsoft AI services) and presented to the user.
    - **Capture Method:** Generated by the Surgery Assist system (utilising various Microsoft AI services such as Azure OpenAI) based on the processed user input and selected from multiple candidates from the Azure storage table source data (AzureResultX).
- **AzureResult[1-6]:**
    - **Description:** Potential responses generated by the Surgery Assist system (utilising various Microsoft AI services), retrieved from Azure Cognitive Search index based on semantic matching with the user's query (interpreted as the 6 possible responses matched against the source data (e.g. ‘FAQs’ stored in Azure storage tables) AzureResult1 might be the top-ranked match.
    - **Capture Method:** Retrieved by the Surgery Assist system (utilising various Microsoft AI services) querying a knowledge base (like the FAQ data source hosted on Azure) based on the user's input.
- **AzureResult[1-6]Score:**
    - **Description:** The relevance or confidence score (semantic ranking) associated with each corresponding AzureResult[1-6], indicating how well that result matched the user's query according to the search/ranking algorithm.
    - **Capture Method:** Calculated by the Azure AI Search service during the retrieval process.
- **RephrasedResponse:**
    - **Description:** A version of the AssistantAIResponse that might have been automatically rephrased for clarity, tone, or conciseness by another AI component. May be null if no rephrasing occurred.
    - **Capture Method:** Generated by Azure OpenAI text generation/rephrasing model using the initial AssistantAIResponse as input.
- **TranslatedResponse:**
    - **Description:** The AssistantAIResponse translated into the UserLanguage if the original response was generated in a different language (e.g., English) and the user requires translation. May be null if no translation occurred.
    - **Capture Method:** Generated by Microsoft Azure Language services using the AssistantAIResponse (or RephrasedResponse) as the source.
- **ODSCode:**
    - **Description:** Organisational Data Service (ODS) code or a similar identifier specific to the organisation the chatbot belongs to (e.g., identifying a specific GP surgery or healthcare provider)

    - **Capture Method:** Set by the admin user in the frontend of the application configuration, based on the environment where the chatbot is deployed.
- **TotalTokens:**
    - **Description:** The number of tokens processed by the underlying language model (e.g. GPT3.5-Turbo) for generating the response to the UserInput. Used for monitoring usage and potential cost calculation.
    - **Capture Method:** Returned by the AI language model API call and logged by the system.
- **MessageTimestamp / CustomTimeStamp:** (There might be redundancy or specific use cases for each)
    - **Description:** A specific timestamp indicating when the user's message was received or processed by the application logic. Could differ slightly from the database Timestamp. CustomTimeStamp might be set explicitly by the application at a specific point in the workflow.
    - **Capture Method:** Logged explicitly by the chatbot application server logic upon receiving or processing the message.
- **[FieldName]@type:**
    - **Description:** Specifies the data type of the corresponding field (e.g., UserInput, AzureResult1Score) as stored within Cosmos DB (e.g., 'String', 'Double', 'DateTime').
    - **Capture Method:** Metadata field, added implicitly or explicitly during data serialisation/storage into Cosmos DB to define the schema.

ConversationLog Table: (Stores summary information about each complete conversation session)

- **PartitionKey:**
    - **Description:** Cosmos DB partition key for the conversation record, using the UserNumber or ODSCode.
    - **Capture Method:** Automatically generated by the application logic based on defined partitioning strategy when creating the conversation record.
- **RowKey:**
    - **Description:** Unique identifier for this conversation record within its partition, set by the ConversationId.
    - **Capture Method:** Set by the application logic, typically using the unique ID generated at the start of the conversation.
- **Timestamp:**
    - **Description:** Timestamp automatically generated by Cosmos DB when the conversation record was last modified (e.g., updated at the end of the chat).
    - **Capture Method:** Automatically generated by Azure Cosmos DB.
- **ChatbotName:**
    - **Description:** An identifier for the specific chatbot instance or version being used.
    - **Capture Method:** Set by the admin user within the frontend application configuration.
- **ODSCode:**
    - **Description:** Organisational Data Service (ODS) code as in the MessageLog, identifying the context of the conversation.

    - **Capture Method:** Set by the admin user in the frontend of the application configuration, based on the environment where the chatbot is deployed.
- **LandingPage:**
    - **Description:** The URL of the specific web page where the user initiated the chat session.
    - **Capture Method:** Captured by the frontend chat client script from the browser's current page URL (window.location.href) when the chat starts.
- **Device:**
    - **Description:** The type of device the user is using (e.g., 'Desktop', 'Mobile', 'Tablet').
    - **Capture Method:** Extracted by the system (frontend script) by parsing the browser's User-Agent string.
- **OperatingSystem:**
    - **Description:** The operating system running on the user's device (e.g., 'Windows', 'macOS', 'iOS', 'Android').
    - **Capture Method:** Extracted by the system by parsing the browser's User-Agent string.
- **Browser:**
    - **Description:** The web browser used by the user (e.g., 'Chrome', 'Firefox', 'Safari', 'Edge').
    - **Capture Method:** Extracted by the system by parsing the browser's User-Agent string.
- **ReferrerUrl:**
    - **Description:** The URL of the web page from which the user navigated to the LandingPage (if available). Helps understand traffic sources.
    - **Capture Method:** Captured by the frontend chat client script from the browser's document.referrer property. May be empty for direct traffic or due to browser privacy settings.
- **ChatStarted:**
    - **Description:** Timestamp indicating when the conversation officially began (This is the timestamp of the first user message or the chat initiation event).
    - **Capture Method:** Logged by the system when the conversation object is created or the first message is received, associated with this conversation ID.
- **TokenUsage:**
    - **Description:** The aggregated total number of AI language model tokens used across all message exchanges within this entire conversation.
    - **Capture Method:** Calculated by the system by summing the TotalTokens from all related entries in the MessageLog table for this conversation.
- **ChatStatus:**
    - **Description:** The final status of the conversation (e.g., 'Completed', 'AbandonedByUser', 'TimedOut', 'AgentTransfer').
    - **Capture Method:** Set by the system based on detecting the end condition (e.g., user closes window, inactivity timer expires, explicit 'end chat' action).

- **[FieldName]@type:**
    - **Description:** Specifies the data type of the corresponding field in Cosmos DB (e.g., 'String', 'Boolean', 'DateTime').
    - **Capture Method:** Metadata field added during data serialisation/storage into Cosmos DB.

Appendix D

Score Comparison

Pass the results to the **ScoreComparisonService** for evaluation using the following logic: ***Condition 1: All Scores Above 2.50***

- **Explanation**: All answers have high relevance.
- **Sub-Conditions**:
  - If the top and sixth best answers have a score difference < 0.1 → list all 6and ask for clarification.
  - If the top six are close in score → list 6 and ask for clarification
  - If the top five are close in score → list 5 and ask for clarification.
  - If the top four are close in score → list 4 and ask for clarification
  - If the top three are close in score → list 3 and ask for clarification. . .
  - If the top two are close → list 2 and ask for clarification.
  - Otherwise → select the top answer.

Condition 2: All Scores Below 1.40

- **Explanation**: All answers have low relevance.
- **Action**: Respond with a message indicating the system couldn't find an appropriate answer and ask the user to rephrase.

Condition 3: Scores Between 1.40 and 2.50

- **Explanation**: Scores are moderate; results are somewhat relevant. • **Sub-Conditions**:
  - If the top and sixth best answers are close (< 0.3) → list all 6 and ask for clarification.
  - If the top five are close in score (< 0.3) → list 5 and ask for clarification.
  - If the top four are close in score (< 0.3) → list 4 and ask for clarification
  - If the top three are close in score (< 0.3) → list 3 and ask for clarification. . .
  - If the top two are close (< 0.3) → list 2 and ask for clarification.
  - Otherwise → select the top answer.

Condition 4: Mixed Scores (Some > 2.50, Some < 2.50)

- **Explanation**: Filter out low-scoring answers.
- **Action**: Apply logic from other conditions to remaining high-scoring results.

5. Handling Score Comparison Results

- **No Relevant Results**:

  "I'm sorry, either I don't have an appropriate response to your question, or perhaps I have misunderstood. Could you please rephrase the question?"

- **Multiple Relevant Results**:

  Ask the user to select from the top results for clarification.

- **Single Best Result**: Pass the `BestTextResponse` directly to the user after translation and rephrasing.