AI Buyer's Guide assessment template

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Purpose**

The exciting promise of artificial intelligence to transform areas of health and care is becoming a reality. However, deciding on whether a specific AI product is right for your organisation is no easy thing. Whilst prospective vendors are proliferating, discerning between their technologies is difficult.

This assessment template is designed to be used together with the [NHSX AI Lab’s Buyer’s Guide to AI in Health and Care](https://docs.google.com/document/d/1HGnPmiSjdfEJVWKfWxg1E1hFEQVkV9YqWLPNHk-Mofo). The guide sets out the important questions you need to consider in order to make well-informed buying decisions about “off-the-shelf” AI products. The sections of this assessment template align exactly to the order of questions in the guide. Explanation of and guidance on the questions set out in this template is offered in the guide.

The template offers a standardised structure for organisations/systems to set out their best answers to the questions posed in the guide. A completed template would serve as a detailed brief on a specific product, for key decision-makers to consider. Different people from across your organisation/system will need to input on different questions; the numbering and sub-numbering in the template should simplify the coordination exercise required.

Please get in touch with us at the AI Lab ([ailab@nhsx.nhs.uk](mailto:ailab@nhsx.nhs.uk)) if your organisation has procured, or decides to procure, an AI product. We are collecting these details so that the Lab can be a central point for sharing use cases, successes and challenges. Please also get in touch about your experience of using the guide and this template.

**Acknowledgments**

The AI Lab is extremely grateful to Haris Shuaib and the AI Board at Guy's and St Thomas' NHS Foundation Trust (GSTT), who first devised and developed an assessment template, and have been happy to share their work.

| **0.0** | **Background information on product** |
| --- | --- |
| *0.1* | *Vendor / manufacturer’s name:* |
| *0.2* | *Name of product:* |
| *0.3* | *Short description of product:* |
| *0.4* | *Intended users of product:* |
| *0.5* | *Anticipated timescale for potential implementation in your organisation:* |
| *0.6* | *Main point/s of contact within your organisation for liaising with vendor:* |

| **1.0** | **Problem to be solved** |
| --- | --- |
| **1.1** | **Challenge-driven** |
| *1.1.1* | *What is the problem you are trying to solve?* |
| *1.1.2* | *What is the rationale for choosing AI to solve your problem?*  *What is it about AI - over and above other solutions - that makes it a powerful choice?* |
| *1.1.3* | *What is the appropriate scale for addressing your challenge - i.e. organisational, system, regional or even national?* |
| **1.2** | **Credible business case** |
| *1.2.1* | *What is the baseline you are looking to improve, and what metrics matter in measuring this improvement?* |
| *1.2.2* | *What do you expect the quality improvements and/or savings and efficiencies to be for your organisation?* |

| **2.0** | **Regulatory standards** |
| --- | --- |
| *2.1* | *What is the intended use of the product? What can it be used for and under what conditions can it be used? What can it not be used for?* |
| *2.2* | *If the product is defined as a medical device, does it have CE marking?*  *What is the product’s risk classification, and do you agree with this designation?* |
| *2.3* | *If the product carries out regulated clinical activity independently of clinicians, has it been registered as a service through the Care Quality Commission (CQC)?* |
| *2.4* | *If the product is categorised as operational healthcare software, has the manufacturer developed it in line with ISO 82304?* |
| *2.5* | *If the product is categorised as healthcare software in general, have you asked to see documentation enabling you to monitor the product manufacturer’s compliance with DCB 0129?* |

| **3.0** | **Valid performance claims** |
| --- | --- |
| *3.1* | *Does the prediction generated by the AI model result in an output that supports practical action?* |
| **3.2** | **Model performance metrics** |
| ***3.2.1*** | ***If classification model:*** |
| *3.2.1.1* | *What are the sensitivity and specificity metrics of the model?*  *Does the trade-off between these metrics give you confidence, given the context of your use case?* |
| *3.2.1.2* | *What are the positive predictive value and negative predictive value metrics of the model?*  *Does the trade-off between these metrics give you confidence, given the context of your use case?* |
| *3.2.1.3* | *Is there an issue of class imbalance to take into account?* |
| *3.2.1.4* | *What is the model threshold? Does the choice of threshold correspond to the use case?* |
| *3.2.1.5* | *What is the Area Under the Curve (AUC) metric of the model?* |
| ***3.2.2*** | ***If regression model:*** |
| *3.2.2.1* | *What is the Root Mean Square Error (RMSE) of the model?* |
| *3.2.2.2* | *What is the Mean Absolute Error (MAE) of the model?* |
| *3.2.2.3* | *What is the R-Squared (R²) value of the model?* |
| *3.2.2.4* | *How much of an issue are outliers for your use case dataset, and how does this influence which of the metrics above should be prioritised?* |
| **3.3** | **Model validation** |
| *3.3.1* | *What are the results from validation tests, to understand the model’s predictive performance on data it hasn’t seen before?*  *Was the validation internal or external?* |
| *3.3.2* | *Has the separation of training and validation data been clearly documented?* |
| *3.3.3* | *Do you understand the characteristics of the validation dataset and what it was used to test for?*  *Was the validation dataset:*   * *Similar to the original dataset in terms of its population and setting?* * *Different to the original dataset in terms of its population and/or setting?* * *Representative of the same or new populations over time?* * *Different to the original dataset on account of technical reasons (e.g. images taken on different scanners)?* |
| *3.3.4* | *Was the validation dataset sampled fairly and representatively, and did it incorporate edge cases?* |
| **3.4** | **AI safety** |
| *3.4.1* | *How does the vendor evidence model robustness?*  *Can the model make reliable predictions, given that data is subject to uncertainty and errors?*  *Does the model remain effective even in extreme or unexpected situations?* |
| *3.4.2* | *How does the vendor evidence model fairness?*  *What measures are in place to prevent the model from discovering hidden patterns of discrimination in its training data, reproducing these patterns and making biased predictions as a result?* |
| *3.4.3* | *How does the vendor evidence model explainability?*  *Can predictions made by the model be explained in terms that both a trained user of the product and a patient/service user would understand?* |
| *3.4.4* | *How does the vendor evidence model privacy?*  *Is the model resilient against attempts to re-identify individuals whose data was contained in the model’s training set?* |
| **3.5** | **Comparative performance** |
| *3.5.1* | *How does reported model performance compare to the current state (i.e. how things are currently done without use of the AI product)?* |
| *3.5.2* | *Is it possible that the seemingly obvious comparator current state may not be the best place to look for potential value offered by the AI product? Are there any less obvious comparators?* |

| **4.0** | **Will the product work in practice** |
| --- | --- |
| **4.1** | **Evidence base for effectiveness** |
| *4.1.1* | *What is the evidence base for demonstrating the product’s effectiveness?*  *Is the standard of this evidence sufficiently robust, taking into account the function and associated risk of the product?* |
| **4.2** | **Insight from other organisations** |
| *4.2.1* | *What insight is available on the product’s effectiveness in other health and care settings?* |
| **4.3** | **Deliverability** |
| *4.3.1* | *If significant changes to your organisation’s ways of working are needed to realise the benefits promised by the product, is this possible?* |
| *4.3.2* | *If implementation of the product will cause short-term disruption, how will you manage this?* |
| *4.3.3* | *If you are replacing an older system with the new technology, have you factored in time, costs and potential complications of dealing with a legacy system?* |
| *4.3.4* | *Have you considered starting off with a pilot project with a tightly defined scope and set of success metrics, before scaling?* |
| *4.3.5* | *What artefacts does the product produce?*  *E.g. Does it produce additional data or files? Does it trigger an alert? If so, what kind of alert?* |
| *4.3.6* | *Does the product record and make available operational data - e.g. processing time, product usage?* |
| **4.4** | **Usability and integration** |
| *4.4.1* | *How will the product interface with different technology systems that are implicated in your deployment, and how will you ensure clear and reliable workflows?* |
| *4.4.2* | *Have you asked the vendor for their software architecture diagram?* |
| *4.4.3* | *Does the product make use of open standards to promote interoperability?* |
| *4.4.4* | *If you want to automatically access the product’s internal data, have you considered whether the product has an Application Programming Interface (API)?* |
| **4.5** | **Data compatibility** |
| *4.5.1* | *What are the product’s data requirements, and how will it ingest this data for processing?* |
| *4.5.2* | *Does your organisation have the data needed, in the right format?*  *What are the sources and types of data needed?* |
| *4.5.3* | *Can your organisation’s data be labelled and stored in the right way?* |
| *4.5.4* | *How reliable is the quality of this data?* |
| **4.6** | **Data storage and computing power** |
| *4.6.1* | *What are the data storage and computing power requirements of the product?*  *How much data will the product need/generate and how long will the data be stored for?* |
| *4.6.2* | *If your project will use cloud-based servers, are you clear about where these are based?* |
| *4.6.3* | *If data storage and computing infrastructure is not provided by the vendor, can your organisation cover the associated costs?* |
| *4.6.4* | *As your use of the product scales and data processing requirements increase, do the infrastructure costs increase in a linear or exponential way?* |
| **4.7** | **Auditing and evaluation** |
| *4.7.1* | *Have you considered how you will audit and evaluate the product and its implementation?*  *Have you factored this into your costs?* |

| **5.0** | **Support from staff and service users** |
| --- | --- |
| **5.1** | **Staff** |
| *5.1.1* | *Have you directly involved staff who will be end-users of this prospective product in the procurement exercise?* |
| *5.1.2* | *Which staff groups have you engaged and gathered input from regarding this procurement?* |
| *5.1.3* | *How confident are you of widespread clinical/practitioner and operational support for the product? What will you do to cultivate this?* |
| *5.1.4* | *Will your vendor supply any induction or training that is needed in your organisation?* |
| **5.2** | **Service users** |
| *5.2.1* | *How compelling a story can you tell about the expected improvement in health and care outcomes?* |
| *5.2.2* | *How will you communicate with patients/service users about how the AI product is being used, how their data is being processed, and, where relevant, how an AI model is supporting decisions which impact on them?* |

| **6.0** | **Culture of ethics** |
| --- | --- |
| *6.1* | *Are you confident that your AI project, and the product in question, is:*   * *Ethically permissible?* * *Fair and non-discriminatory?* * *Worthy of public trust?* * *Justifiable?* |
| *6.2* | *Have you assessed your project against the principles of the Data Ethics Framework? Are there any areas of the project that need revisiting as a result?* |
| *6.3* | *Have you carried out a stakeholder impact assessment? What are the key insights from it?* |

| **7.0** | **Data protection and privacy** |
| --- | --- |
| *7.1* | *Will you be able to create a data flow map that identifies the data assets and data flows pertaining to your AI project?* |
| *7.2* | *Will you be able to develop a data processing contract - otherwise known as an information sharing agreement - with the vendor?* |
| *7.3* | *Is your organisation’s use of data for this project covered under its data privacy notice?* |
| *7.4* | *What will be in place in terms of data protection to mitigate the risk of a patient or service user being re-identified - in an unauthorised way - from the data held about them?* |
| *7.5* | *In cases where you will be processing personally identifiable data, will you be able to complete a Data Protection Impact Assessment?* |

| **8.0** | **Ongoing maintenance** |
| --- | --- |
| **8.1** | **Vendor’s responsibilities** |
| *8.1.1* | *Is the vendor providing a managed service for the product?* |
| *8.1.2* | *What is the vendor’s approach to product and data pipeline updates?*  *Who pays for these?* |
| *8.1.3* | *What is the vendor’s plan for mitigating adverse events - i.e. if the AI product fails or is compromised?* |
| *8.1.4* | *What is the vendor’s plan for addressing performance drift?*  *Have you agreed a suitable margin of acceptable drift? Does performance need continuous monitoring or is an interval audit sufficient?* |
| **8.2** | **Your organisation’s responsibilities** |
| *8.2.1* | *If you are not buying into a managed service, do you have the IT capability in-house?* |
| *8.2.2* | *Can your organisation develop a sufficiently robust understanding of relevant data feeds, flows and structures, such that if any changes occur to model data inputs, you can assess any potential impacts on model performance - or signpost questions to the vendor?* |
| *8.2.3* | *Are you clear about your organisation’s reporting requirements vis-a-vis adverse events?* |
| *8.2.4* | *What are the vendor’s expectations of your organisation sending back data to support their iteration of the model or development of other products?*  *Have you clarified what the vendor means by model iteration and development, and have you ensured that your information governance arrangements address this?* |
| **8.3** | **Decommissioning** |
| *8.3.1* | *On decommissioning the product, what will happen to any data that is stored outside of your organisation’s systems? Will it be deleted, or archived?* |
| *8.3.2* | *How will you ensure that you have access to any data or analysis you require that is due to be deleted or archived?* |
| *8.3.4* | *On decommissioning the product, how will you ensure that the vendor’s access to any part of your organisation’s infrastructure is revoked in full?* |

| **9.0** | **Compliant procurement** |
| --- | --- |
| *9.1* | *Have you clearly documented and justified instances of your organisation talking to / inviting specific vendor/s to bid for the project?* |
| *9.2* | *If you are being offered a product for free, what steps have you put in place to ensure that you remain compliant with public procurement guidelines?* |

| **10.0** | **Robust contractual outcome** |
| --- | --- |
| **10.1** | **Commercials** |
| *10.1.1* | *Are you clear about exactly what you are buying?*  *E.g. Is it a lifetime product? Is it a license? What is the accompanying support package?* |
| *10.1.2* | *Have you set out a clear specification and service level agreement?*  *Do these secure the quality, availability, flexibility and performance of service that you need?* |
| *10.1.3* | *What provisions are in place for contract termination and handover to another supplier?* |
| *10.1.4* | *To what extent will you be able to publish details of your contract?* |
| **10.2** | **Intellectual property** |
| *10.2.1* | *How will you ensure that any agreement with your prospective vendor is fair, in the sense that it recognises and safeguards the value of the data you are sharing?* |
| **10.3** | **Liability** |
| *10.3.1* | *With regard to product liability, is the vendor providing any indemnities, and are they clearly set out in the contract?* |
| *10.3.2* | *Is it clear what is considered as product failure versus human error in using the product?* |
| *10.3.3* | *What is the extent of cover your own indemnifier or insurer can provide in the event of product failure or human error? Do you need to purchase additional cover or extend existing cover?* |
| *10.3.4* | *Does your contract and information governance documentation clearly set out what measures you expect the vendor to have in place vis-a-vis compliance with data protection regulation?* |