



# UNIVERSITY OF KENTUCKY

## Purchasing Division

### Written Questions and Answers

Enterprise Data Warehouse (EDW) Modernization Implementation

RFP UK-2262-23

Closing Date: 07/06/2022

Today's Date: 06/23/2022

No.	Question	Answer
1	Whether companies from Outside USA can apply for this? (like, from India or Canada)	Yes, however, company must be registered to do business in Kentucky. See section 6.15 of the RFP. None of the UKHC data is permitted to leave the United States.
2	Whether we need to come over there for meetings?	We do not require traveling for meetings. The current intent is to perform the project the engagement in hybrid mode.
3	Can we perform the tasks (related to RFP) outside USA? (like, from India or Canada)	Yes, only for tasks that do not require sending data outside North America or accessing UKHC systems outside North America.
4	Can we submit the proposals via email?	No. See Section 3.6 of the RFP.
5	Page 2 Will University accept reasonable, good faith redlines to purchasing terms?	See section 3.5 of the RFP.
6	Page 7: Section 2.1 Who are the stakeholders?	The stakeholders comprise of representatives across; IT, Clinicians, (Operations and Finance) Administrators and Research.
7	Page 10: Section 2.2 You state a goal "anticipating/course correcting with predictive models/NLP at the point of care". Does this mean you are requiring offerors extract data in real-time from Epic, run it through predictive models/NLP and then deliver results into Epic in real-time? If not, what does 'at the point of care' mean and how do you envision this working architecturally?	Offeror to propose the plan and approach that accomplishes the availability of data and analytical insights that impact the quality and efficiency of care in real-time and near real-time during the course of a patient stay. Epic shall be at least one of the data source.
8	Page 15: Section 3.6	Yes

	Is it acceptable to ship 1 package so long as both proposals contained therein are in separate, sealed envelopes that are clearly marked as described in this section?	
9	Page 15: Section 3.6 Does UKHC require USB-A or USB-C drives?	To put your proposal on it a normal USB Flash drive will work.
10	Page 15: Section 3.6  For an abundance of clarity, can UKHC confirm that the technical proposal envelop contains transmittal letter, exec summary & overview, criteria 1, 2, 4, and 5 and that the financial proposal envelop contains only criteria 3? Do any items from the technical proposal need to be replicated for the financial proposal?	That is correct. No, no technical needs to be included with the financial proposal.
11	Page 15: Section 3.6  For an abundance of clarity, can UKHC confirm that section 8 and sections 7.1-7.2 are the required items that make up criteria 3?	Criteria 3 is listed in Section 4.7 and is the Financial Proposal, including section 8.0.
12	Page 16: Section 3.11  Is proprietary information held confidential, so long as it as marked as such?	University Legal makes the determination as to what is confidential within the proposals.
13	Page 22: Section 4.5.D  As a private company we do not provide prospectus / company annual reports. Will UKHC waive any requirement around this item?	Please provide a letter of explanation within your response.
14	Page 22: Section 4.5.F  This is a hypothetical and must be waived since it is not possible to accurately respond to. Can you confirm response to this will be waived?	No
15	Page 25: Section 4.6  Under "Enterprise Analytics and ML Ops Platform" UKHC lists customized use-cases. Will the offeror be creating the	Assets like " <i>Enterprise-Level Balanced Scorecard</i> ", " <i>Capacity &amp; Throughput Management</i> ", " <i>Population Health</i> " and " <i>IoMT Analytics</i> " exists in various stages of the development life-cycle.

	metrics, dimensions, data models, visualizations and documentation for these use-cases from scratch? Or do all of these assets exist today for these use cases and offeror will rebuild these assets into the modernized EDW?	Given the differences in the platforms, we expect the effort to include creating the dimensions, data models, visualizations and documentation for new and existing assets being rebuilt in the modernized EDW.
16	Page 36: Section 6.26  Will UKHC confirm that pre-existing content, IP, tools, etc. that offeror brings to accelerate execution of the proposed engagement will remain under the ownership of offeror?	Offeror may leverage their pre-existing tools and practices as accelerators to the execution of the engagement if such tools are exclusively developed and refined outside of the engagement.
17	Page 31: Section 6.12  Can you confirm that UKHC will appropriate funding for the entirety of the initial term of the proposed engagement in advance of any project work commencing?	No, we are not able to confirm full upfront payment in advance of work commencing. Payment shall be made on a scheduled basis as determined by subsequent SOW contract.
18	Page 7: Section 2.1  Is the offeror solely responsible for the setup and security of a fresh Azure environment, or will the modernized data platform be deployed within UKHC's existing Azure account in collaboration with UKHC Azure administrators?	We have an Azure tenant, the EDW modernization shall be deployed in a fresh Azure UKHC subscription that Offeror will be accountable and responsible for setting up. Offeror will work with UKHC Azure team in performing the set-up. The UKHC Azure team will be provide details of current set-up and facilitate infrastructure requests and signing-off on set-up.
19	Page 7: Section 2.1  If offeror will be collaborating with UKHC Azure administrators to deploy the modernized data platform on an existing Azure account, what are the divisions of responsibility between UKHC and offeror personnel?	See #18
20	Does UKHC plan to support any real-time or near-real-time analysis use cases with this data platform? If so, can you elaborate on the latency requirements and source systems for real-time data (if known)?	Yes, UKHC plan to support any real-time or near-real-time analysis use cases with this data platform. Latency requirements ranges 1-15 minutes. Source systems include but not limited to Epic and Philips Devices.
21	Page 7: Section 2.1  The RFP lists "self-servicing capability (beyond dashboarding)" as a desired	This is in reference to code-free self-service analytics beyond dashboarding.

	outcome. Is this limited to users comfortable with SQL, Python, R, etc., or are you also looking to support <i>code-free</i> self-service analytics beyond dashboarding tools?	
22	Multiple pages and sections Has UKHC implemented a standardized data visualization solution (Tableau, Power BI, Qlik)? If not, should the Offeror's proposed solution include deployment and integration of a data visualization solution?	UKHC has implemented Tableau as our standard visualization solution and are open to leveraging Power BI.
23	Page 12: Section 2.4 Is there a specific supplier diversity requirement for this bid?	No.
24	Is there a metadata management system in place? If yes can you share a brief on the same	No.
25	Does the system have a documented data model, business glossary, documented list of business processes?	We have partial documentation of the current EDW not including the data models.
26	Is there a data cleansing or data correction process defined currently? Is there a process to measure the quality of the data?	No.
27	What are the Business Units (BU's) consuming the data	Ambulatory Operations, Hospital Quality, Hospital Operations, Billing and Budgeting, Performance Services, KCR and Research.
28	What is the user density for the consumption of the data in future architecture	We expect 1200+ active users of the future system.
29	Is the data for internal consumption or external user will also have access	Primarily internal with some external users as determined by the " <u>Enterprise Health Data Trust (Exchange)</u> " and similar use-cases.
30	What is the expected data volume to be processed in future state data platform?	The future-state is expected to process 16TB of data with 2-4TB processed daily.
31	What is the data volume of delta data	Less than 1TB.
32	What is the ingestion frequency required for the future state data platform?	Real-time, near real-time, daily, weekly, monthly and ad-hoc.
33	What is the report and dashboard data refresh frequency required for the future state data platform	Reports refresh frequencies are primarily daily.
34	What is the reporting tool of preference, as tableau is already in use?	Tableau (willing to introduce Power BI when if it is the best option on a case-by-case basis).

35	Is there a requirement for real time data ingestion, if yes, what will be the possible mode of ingestion	Yes. Offeror shall recommend best practice based on experience including modes such as streaming and API integration.
36	Is there a list of systems available, required to be integrated to the future state data platform	Primarily Epic today, a comprehensive list is not available.
37	Please provide details on what kind of Unstructured/Semi-Structured data sources will be leveraged.	We require the capability to the capability to ingest and process documents, notes, charts, voice, images, xml, json and web data that may contain, clinical notations, research data, study data, diagnostics, metadata and or patient feedback.
38	Is the cloud platform setup, management and configuration responsibility of the vendor?	Initial set-up and configuration is the responsibility of the Offeror. Offeror is responsible for successful transition of ongoing support and management to UKHC.
39	Is there a requirement for workspaces for different BUs/data analyst to support their analytical needs, if yes is there a requirement to ingest the resultant dataset to the workflow.	Yes, primarily for the Research and Performance Services units who will require their result sets available for ingestion to the workflow.
40	<p>For Phase 2:</p> <p>Please provide more details on existing data landscape for migration perspective</p> <ul style="list-style-type: none"> <li>- Is all data managed in current Netezza system required to be migrated to new platform?</li> <li>- What is the volume of data that needs to be migrated?</li> <li>- Approximately how many ETL routines to be converted/re-written in target platform?</li> <li>- If the target data visualization platform is different than Tableau, how many Tableau reports to be migrated to target platform?</li> <li>- Does current data environment comply with Epic Data platform or Epic is newly implemented and only new data platform is supposed to comply with its capabilities?</li> <li>- Please provide Architecture of current platform depicting source systems, data flow, downstream systems etc.</li> </ul>	<p>All 16TB of the current Netezza DB will be migrated to the modernized Platform majority of which will be for storage and research with 4-8TB for daily operations.</p> <p>The current EDW is comprised of approximately 300 ETL DataStage jobs.</p> <p>The Epic Cogito platform services operational and some analytical reports. The EDW is complementary, ingesting Epic and other sources to support research, dashboarding and advance analytics.</p> <p>The current EDW architecture is a hybrid containing point to point ETL and source to mart ETLs. Sources include flat files, xmls, Netezza and SQL DBs. ETL lands sources to Netezza where store-procedures and queries surface processed data for visualization.</p>
41	For Phase 3: Can we assume only implementation of azure components	The scope includes implementing the Azure components for ML Operations and

	(Infrastructure, tools & process) pertaining to Advance Analytics & Machine Learning use cases is in scope. Scope does not include building & deploying ML models for any use cases?	deploying at least one use-case to demonstrate the value, utility and adoption of the ML infrastructure. This engagement goes beyond the deployment of tools, the tools must be adopted and integrated into operations with visible demonstrated value.
42	For Phase 3: If implementation of ML use cases in scope, do you have the hypothesis developed or you are looking for Vendors help in assessing and developing hypothesis as well?	We expect the Offeror to evaluate the best fit high-value hypothesis that delivers measurable impact to UKHC mission.
43	For Phase 3: Are you looking for Vendors help in evaluating and implement platforms/tools for AutoML and MLOps as well?	We have a list of preferred tools based on an evaluation conducted in 2021. We expect the Offeror to share recommendations on any new developments or alternatives that may better fit our needs as implementation nears.
44	What is the evaluation criteria for the program?	<p>UKHC EDW modernization is a transformational initiative aimed at accomplishing outcomes such as:</p> <ul style="list-style-type: none"> <li>• added capabilities to efficiently transfer, store, integrate, process and prepare data to present actionable intelligence regardless of the variations in data formats, transfer mechanism or speed;</li> <li>• accelerated efficiency with delivering advanced analytics (such as Machine Learning and NLP);</li> <li>• sophisticated self-servicing capability (beyond dashboarding) to securely meet user needs;</li> <li>• optimize capabilities to support UKHC's research mission in a more strategic way;</li> <li>• provide sophisticated architecture for performance services/analytics;</li> <li>• successful transition of the modernization solution to UKHC; and</li> <li>• creating the expertise within UKHC to support and independently build on the modernization path.</li> </ul>
45	Is there an on-prem architectural diagram available to understand the source systems and their connectivity?	No.

46	What is the volume of data coming in on a daily basis?	See #40 and #31
47	How many GREFs are being run?	Please clarify "GREFs".
48	What is the total number of Epic users? How many concurrent users are there in the EDW?	Epic is fully implemented at UKHC and available to all UKHC clinicians, administrators and professionals.  The EDW may have nearly 100+ concurrent users based on active users.
49	What is the format of the source files, is schema drift happening over time in these files?	Majority of the sources are flat-files or relational databases with little to no schema drifting happening over time.
50	Do all the source systems support incremental load?	No.
51	The initial full load from on-prem to Azure Data Lake will happen from the source systems or from Netezza?	This has not been finalized but we are leaning in the direction of having the initial migration happening from Netezza.
52	Do all the source systems provide a direct access?	No. Most do and others are delivered through our MFT tool.
53	Are there any on-premises applications that need to connect to the cloud database?	Yes.
54	Will there be access to the old system once the new system is in use?	Yes, read only access for applications will exist once the new system is in use. Note that the new system may consist of legacy ETL and DB tools as technical debt.
55	How often will historical data need to be accessed?	Longitudinal data going back two years may be accessed frequently on a daily basis. Beyond two years for history, most queries will be Ad-hoc in support of research needs.
56	Is there scope of new reporting tool like Power BI which is native to Azure?	Yes.
57	Can development happen from locations outside of the United States?	Yes.
58	How many total tables/views etc. are in the total scope?	To be determined. Total current count is in the order of 10,000 of which we expect nearly 2000 being required.
59	Is data retention part of migration. If yes is data retention policy identified?	No.
60	Is reporting through Azure or part of the scope?	See #56.
61	Is there a list of transformation logic modules/pipelines ? If yes, how many?	No list. We have nearly 300 DataStage Jobs.
62	What carrier is being used for internet connectivity? Is connectivity to Azure needed?	N/A.

63	Will UKHC postpone the due date for submissions due to postponement of question and answers and proximity to the holiday?	No.
64	Could you share your architecture diagram of the current EDW Platform?	This is not available, see #40.
65	What is the UKHC's Data Model Strategy? Do you intend to: a) retain & optimize the current model, b) use Native Azure CDM Model, c) use custom build model, d) any other model	We intend to do a combination of all predominantly retaining and optimizing existing models.
66	Could you provide more clarity on the EDW Modernization foundation attributes like IoMT, streaming data, SDOH, Registries, various third party sources? Are these functionalities currently a part of the ecosystem or is it a future state vision for the roadmap? Also, please provide details on the below. - What is the current deployment for IoMT & streaming data? - What type of SDOH data UKHC is currently ingesting and processing? - What is the number of registries UKHC has subscribed to?	These are disparate functionalities in various stages of maturity in the current EDW. We seek to optimize, unify and mature as part of the future state.
67	Could you share the following details on UKHC's MLOps platform? - What is the current MLOps technology landscape & models? - Are there any existing MLOps tools/frameworks that are being used today? - count of ML Ops models?	We do not have an MLOPs framework. Our Data Scientist leverages tools such as; Python (numpy, pandas, scikit-learn, keras, tensorflow etc.)
68	What is UKHC's current implementation on Health Data Trust? How data is shared and accessed by the partner entities E.g. API, data share et al.	This is a future state requirement. The current EDW provides user-based role access to enable partner affiliate organizations Netezza access based on UKHC ID.
69	What are the tentative project timelines that UKHC has identified for the 4 phases mentioned? <Refer Pg-8 in RFP>	We expect each phase to run 3-4 months in duration for a total project timeline of 9 -16 months.
70	Does UKHC want the vendors to price out the overall implementation in phases?	Yes. Cost shall be submitted with breakdown in each phase as well as estimated level of involvement at each phase.
71	Which cybersecurity tools suite is UKHC using for protecting data threats today?	Will not publish this answer.



72	Page 8; Section 2.1 Do you require the four phases to run sequentially?	No, phases may run concurrently depending on UKHC resource availability. We expect phase one and two may start sequential and all subsequent phases run in parallel.
73	Page 8; Section 2.1 Does UK have an overall timeline expectation for completion of the project?	Yes. The transformation is expected to run no longer than 18 months.
74	Page 8; Section 2.1 Do UK and the vendor mutually confirm the use cases to be deployed during Phases 2 and 3?	Yes.
75	Page 8; Section 2.1 As this project progresses, it is possible that new external data sources may become available to UK. Is the expectation that a finite number of external data sources would be in scope initially with a process for adding future data sources later?	Yes. As new sources are required, we expect the engagement to provide the knowledge transfer that empowers UKHC staff to make and support modifications to the system including identifying and onboarding new data sources.
76	Page 9; Section 2.1 Do the Phase 3 use cases include the development and deployment of NLP or ML models?	Yes.
77	Page 10; Section 2.1 Will UK be able to allocate time for current staff to participate in the learning? Does UK require assistance to back-fill current roles to ensure current operations are not impacted?	Yes.
78	Page 10; Section 2.1 Are you prepared to hire new staff with requisite skills that may be required for certain functions, e.g., Azure administration, machine learning model development, etc.?	We plan to train existing staff and hire if sufficient justification is presented. The engagement is expected to identify justification for required roles in the early parts of phase 1.
79	Page 10; Section 2.2 Is there an established governance structure (e.g., Steering Committee) in place for reporting status and addressing issues? Can you share the UK executive roles/titles included in the structure?	There will be an EDW modernization committee chaired by the Executive Director of Enterprise Data Management and includes, Physician Champion, Nursing Champion, Research Leadership, Chiefs, VPs and IT Leaders (Directors and above).
80	Page 29; Section 6.2 Will UK be required to garner financial approval for this work on an annual basis, or will approval for the full amount	Funding will require justification annually regardless of financial approvals to provide allocation for the full project estimate.

	be secured before the project work begins?	
81	Please provide the UKHC Modernization Roadmap for us take under consideration. a. If the architecture is not part of the roadmap, please provide the developed architecture.	Please refer to section 2.1 and the “EDW Modernization RFP Requirements” attachment.
82	Does UKHC expect all phases to be executed within a certain timeline?	See #69
83	Scale questions: a. How many active data consumers in the existing and future analytics environment? b. Can UKHC provide numbers of sources and statistics surrounding those sources? (# of tables, storage size, record counts, etc.)	We estimate approximately 1200+ active data consumers for the future environment.  Epic shall be a major source. The environment needs to be future-proof and flexible evolve with the needs of UKHC.
84	Will a semantic layer be created as part of this project?	To be determined.
85	Is UKHC open to separate SOW's per phase?	Yes.
86	UKHC IT Support a. How many UKHC developers will be working on the team? What are their skillsets presently? b. Is there a plan to scale the team to support the new platform? c. Can training for UKHC developers be done at the end of phases?	We have 3-4 UKHC developers planned to be part of the project team. They have a background in DataStage, Netezza and Epic Cogito development. They are receiving formal training in Azure courses (ex. AZ 900, DP 203 and DP 300).  We are seeking a partner to execute on a train as you go model that provides UKHC staff with experiential learning while building the modernized EDW.
87	Is UKHC open to using data modeling toolsets?	Yes.
88	How much of the DevOps engineering will be managed by the Offeror?	Offeror is responsible for establishing the practice and transition it to UKHC staff.
89	Will infrastructure be managed by Offeror? Can the infrastructure be managed as code?	Infrastructure will be managed by UKHC after the build is completed and stabilization by Offeror.  Infrastructure can be managed as code if Offeror prove this to be the most suitable, efficient and cost-effective approach.
90	What will the process be for Azure security and governance? Is that in place	Policies exists. UKHC is open to additional recommendations for which we will reserve the right to implement.

	already or will governance policies need to be established?	
91	<p>What are the concepts and technologies that make up the current UKHC EDW?</p> <p>a. Section 4.6 - Therefore, Offeror must demonstrate capabilities and expertise with concepts and technology that make up the current UKHC EDW as well as the technology and principles required to transition to UKHC's targeted modern enterprise data platform.</p>	<p>Understand of Inmon and Kimball data modelling approaches. Understanding of Epic Cogitio, IBM DataStage, Netezza and Data Modeling.</p>
92	<p>Section 2.1</p> <p>For Phase 3, what are the data science/AI/ML use cases to be targeted?</p>	<p>To be determined in collaboration with Offeror.</p>
93	<p>Section 2.1</p> <p>Are there existing AI/ML models which will need to be migrated to MS Azure? Please provide details.</p>	<p>To be determined in collaboration with Offeror.</p>
94	<p>Section 4.6E</p> <p>Are there preferred modalities or tools, such as an Learning Management System(LMS) platform, that we should consider for training delivery?</p>	<p>Our preference is to provide practical experiential learning with supporting LMS artifacts. UKHC leverages the Success Matters LMS.</p>
95	<p>Section 4.6F</p> <p>Is retiring of the on-premise solution part of the project? Are there any expectation to co-exist the solutions for specific time?</p>	<p>Part of the project goal is to retire the legacy ETW tools. However, both solutions will need to co-exists to facilitate a successful transition and manage risks associated with those processes that require redesign and or refactoring in order for them to run cost-effectively and performant in the cloud.</p>
96	<p>Section 4.6</p> <p>Will the current Data model(s) in EDW expected to be retained in future state EDW or should be replaced with something new.</p>	<p>Current EDW Data Models will be retained in the future state and may require some modification to best leverage the cloud.</p>
97	<p>Section 7.1</p> <p>What factors/features were considered and compared when making selection for MS Azure services (14 listed in section 2.1)?</p>	<p>These are selected based on EDW modernization baseline architecture. The architecture is built considering; EPIC Integration Path</p> <ul style="list-style-type: none"> <li>• Self Service Capabilities</li> <li>• Scalability</li> <li>• Deployable as a Managed Service</li> <li>• Ecosystem Integration</li> <li>• Cost to Implement</li> </ul> <p>Availability of Skillsets</p>

100	Section 7.1 As part of the review, is UKHC open to recommendations for additional MS Azure services and/or other complementing technologies (e.g. Security, DevOps and Governance) not part of MS Azure?	Yes.
101	Section 7.1 What resourcing capacity (FTEs) and Azure competency (certification levels) does UKHC plan to commit for support and operations and organizational change management (including training support)?	Please see #86
102	Section 7.1 As part of EDW modernization, Is there a plan to change the Gateway , from IBM DataPower, to something else? Any technologies evaluated which will be compatible with MS Azure?	No.
103	Section 7.1 Does the Research Analytical Data Platform (e.g. RDW) currently exist? Any migration expectations?	No. Today we have a large Netezza DB available with historical data to support Research.
104	Section 7.1 How many data source and target systems are on-prem vs cloud based?	Today virtually all sources and targets are on-prem.
105	Section 7.1B What privacy compliance framework(s) are in scope for the EDW implementation (e.g., GDPR, CCPA, etc.)	HIPAA/HITRUST
106	Section 7.1B Does UKHC already have technology solution that can facilitate the process of identifying, profiling and cataloging data and metadata?	Not actively enabled.
107	Section 7.1B Are there any Business-2-Business or Business-2-Customer requirements for providing access to EDW?	Please review " <u><i>Enterprise Health Data Trust (Exchange)</i></u> " example use-case.
108	Section 7.1B Are there expectation to use other visualization tools (except Tableau) to report out of the MS Azure data platform?	See #22
109	Section 7.1B	Estimated 18-29

	How many downstream applications are expected to consumer data from MS Azure Data platform?	
110	Section 7.1B Are there any expectations to develop visualization solutions using UKHC business intelligence tools or other applications?	Yes. See #22.
111	Section 7.1B What technology solution(s) does UKHC use for enterprise Identity and Access Management? Will MS Azure native capability be used ?	MS Azure native capabilities will be used. IAM to be disclosed only to the selected Implementation Partner.
112	Section 7.1B What is the current state framework/architecture for Cyber Security? Any considerations for the modern EDW solution?	To be disclosed only to the selected Implementation Partner.
113	Section 7.1B Does UKHC already have documented data storage and archival consideration including retention and disposal requirements?	Yes.
114	Section 7.1C Does UKHC currently have a Cloud Security Posture Management (CSPM) technology solution in place?	To be disclosed only to the selected Implementation Partner.