

COLLECTION CARE DATABASE**DEADLINE FOR TENDER RESPONSES: 5PM, 5 MAY 2017****CLARIFICATION QUESTIONS AND RESPONSES**

The National Archives has received a number of clarification questions relating to this opportunity. Please find below a list of those questions and their associated responses.

Q1: *Migration: Mapping fields: TNA field title to New Database field title would be required where data would be imported into the system. Our expectation would be to provide our field titles in a spreadsheet and have the mapping done on your side (i.e. 'Accession number' = 'ID number', 'Creation Date' = 'Object Date'). We can possibly provide a custom template for the TNA in our system using your terms (depending on budget priorities), in which case we would need to know 1. The field titles, 2. Agree an order of fields, 3. The character limit for each field and 4. Type of data that would be stored (date, text, integer, currency, dimensions, signature etc) and examples.*

A1: Please see file example named 'Proposed Field Titles for Conservation Documentation' (attached as a separate document).

Q2: *Migration: Data import vs. doc attachment. Microsoft Access, Excel, Word. We should be able to upload Microsoft Access and Excel data (once the fields are mapped for each). This may or may not be easy, depending on the data stored and how complicated your data and relationships are and if this data needs to be stored in the same way in our system. Can you provide your database schema, and the template/ or example of the data you store in excel?*

A2: Yes, we can provide examples of data stored in Excel. (Please see examples attached as separate documents).

Q3: *What percentage of data is in Word format approximately?*

A3: A very small percentage is in word, there are approximately 1,800 documents with approximately 1-6 pages each.

Q4: *Would simply attaching Word docs as historic reports be sufficient?*

A4: Attaching the word document would be sufficient.

Q5: *Do the Access and Excel data sets include basic object details for the word reports, so that step 1 could be importing basic details, and step 2 attaching word docs to the object records?*

A5: The Access and Excel data would have the catalogue reference number that would also be present in the Word document; however this isn't always in the title but in the actual document itself.

Q6: *Migration: with images stored on separate drives. Do the images have ID numbers as their title?*

A6: The images do have ID numbers as their title but with a "." instead of a "/". So for example "FO 371/21" as an image would be saved as "FO 371.21"

Q7: *What are the file sizes of the images?*

A7: Generally the images are 3.5mb-6mb in size. The occasional TIFF file will be larger for Loans and Exhibitions; however we do not store these in the long term.

Q8: *Do the images have metadata? Do you expect anything to be done with the metadata if so?*

A8: There is no metadata.

Q9: *Do images have to be stored in their original formats?*

A9: Images don't have to be stored in original format, as long as they can still be viewed easily. Please ensure you state in your tender response which viewer(s) you intend to use and whether you anticipate any compatibility issues with such viewer(s).

Q10: *What resources would be available on your end to facilitate the progression of this project?*

A10: There are 2 members of staff within the Collection Care Department who have time allocated to facilitate the delivery of this project.

Q11: *Is the data currently held in Excel and Microsoft Access mapped to item entries as defined in section 5.3, page 6 of the ITT? I.e. mappable to annotations, attachments, links, item condition, collection, catalogue reference (identifier) etc?*

A11: The data held in Excel and Microsoft Access are identifiable using the catalogue reference number, it is not always mappable to documents or images, due to document titles not always using the catalogue reference number. However, the catalogue reference number would be present in the word document. Where any items are not easily mappable, we will work with the Appointed Supplier to rectify as far as possible.

Q12: *Can you provide a sample of the way data is held in Word?*

A12: Please see examples attached as separate documents.

Q13: *You mention you currently have not more than 1 TB of data. How many individual items with individual identifiers does this represent?*

A13: There are approximately 1950 individual items on the Access database, this does not represent the individual identifiers because for example if a series was treated and had 100 pieces, it may have only be entered as FO 371/ 1-100, instead of each reference individually (FO 371/1, FO 371/2, FO 371/3, etc.).

The UNFIT LIST Excel sheet has 9340 individual items, the majority of which are individual identifiers.

There are 39,078 images in 2,359 Folders – most are marked by individual identifier, but some are not.

There are approximately 1800 word documents of about 1-6 pages, not all with individual identifiers. However the identifier would be contained in the text of the Word document.

Q14: *Over the initial 3-year term what is likely to be the maximum number of users with (a) read & write permissions and (b) read-only permissions?*

A14: The maximum number of read and write users will be 50. Ideally we would like to be able to add read-only users as required, perhaps by sending a link; we estimate no more than 25 read-only accounts at any one time.

Q15: *What is the oldest version of each type of browser (I.E., Chrome, Firefox, etc.) that will be mandatory for the solution to support, or a statement that you will deploy modern browser interfaces to all users?*

A15: The current browsers are Internet Explorer v11, Chrome v56.0.nnn and at a pinch, Firefox v52.0. This is the minimum versions the product should support. TNA will also move from Windows 7 to Windows 10 within the lifetime of the contract, therefore compatibility with the new MS browser in W10 'Edge' is desirable. However, Edge is not mandatory as Edge as it appears IE will be around for some time.

Q16: *Please explain how relevant data is stored in Word, is it free-form narrative and unstructured or tabular – is a consistent formal glossary of terms or codes adopted? (is a spade always a 'spade' or shovel/digging tool, etc). Please provide examples of data stored in Word that are representative of majority of cases and an estimate of (number of pages) of Word format data records.*

A16: The data stored in Word, is mainly free form narrative and unstructured, there are some documents with tables. There are approximately 1800 number of documents with approximately 1-6 pages. We have no codes adopted and there is not a consistent formal glossary. Examples are provided as separate documents.

Q17: *Section 5.2 of the ITT refers to the integrity of your datasets. Please advise whether the data is considered clean or whether a significant data cleansing process will be required and is this to be part of this ITT or are alternative measures in place.*

A17: Some data cleansing may be required for the data in the Access database, but not the other sets, this would be a requirement of the ITT.

Q18: *Section 5.3 of the ITT indicates you require your data management application to be an image manipulation application. Is it considered an appropriate response to allow images to be downloaded for manipulation in a 3rd party application, not part of the respondent's tender, and uploaded with appropriate version control?*

A18: To minimize the number of applications we are using, we would prefer the ability to do this within the system.

Q19: *We are in the process of applying for ISO 27001 certification. Must we hold ISO 27001 now? (We are not certified for Ministry of Defence List X, or Cyber Essentials Plus).*

A19: The requirement for one of the three standards is mandatory. If certification was attained prior to contract award, then it would be considered.

Q20: *Is it envisaged the supplier will migrate current data, or that the supplier will provide processes to facilitate CCD doing this en-masse?*

A20: It is envisaged the supplier will migrate current data.

Q21: *Can CCD confirm they use a single Access database, and how many records (out of the 11 million maintained by TNA) this database holds?*

A21: We use a single Access database which currently has 1,948 records. However this number is likely to be higher, as currently a series may have been entered as FO 371/1-100, and not individually such as FO 371/1, FO 371/2, FO 371/3.

Q22: *What information do the Word and Excel files contain? Will these be imported as attachments or does the system need to extract the contents and import into the database?*

A22: The Word and Excel files contain information regarding condition, treatment, status and survey information. They can be imported as attachments, with some of the data from Excel extracted.

Q23: *Can CCD provide the structure of the Access database, along with dummy/test data? If Word and Excel files contain data, please can CCD provide samples of these too?*

A23: We can provide examples of Word and Excel files (attached separately). Unless required, we prefer not to provide a structure of the current database as we want to move away from this structure.

Q24: *Do image file names correspond to Catalogue Reference identifiers? If not, please can CCD advise how these files are named and related to CR identifiers?*

A24: The images file names are labelled in this format FO 371.3005 or FO 371 3005, whereas on the Catalogue this would be FO 371/3005

Q25: We understand originating government department, please can CCD give examples of series, piece, item and sub-item, and maximum numbers required for each range?

A25: Examples of different references

	Department	Series	Piece	Item	Sub item	Sub item
A 9/1	A	9	1			
ADM 1/5439	ADM	1	5439			
ASSI 5/35 part 1	ASSI	5	35	Part 1		
ASSI 5/35 part 1	ASSI	5	35	Part 3 end		
CP 52/3/25/2/2	CP	52	3	25	2	2
FO 371/104564	FO	371	104564			
FO 1093/1	FO	1093	1			

Maximum numbers for each range can vary greatly and we would need the ability to input numbers in to the 100,000's.

Q26: Does CCD currently operate any barcode or RFID labelling system?

A26: No.

Q27: Annotating Images: Should this be destructive (i.e. the image is saved with annotations embedded graphically), non-destructive (i.e. the annotations are saved separately), or mixed?

A27: Mixed would be preferable.

Q28: Activity Workflow Templates: Please can CCD give examples of the activity tasks required for template workflows, and what features may be required for bespoke workflows?

A28: Please see the document 'Example Exhibition Workflow/Action List' (provided as a separate document).

Q29: Reporting: Please can CCD provide samples of common and/or required reports?

A29: Currently common reports are regarding document condition and treatment carried out. Currently these will be produced manually, please see 'Example-Condition report' (provided as a separate document). The current system is able to produced a report, however it exports it in a disorganised manner (please see 'Example Conservation Documentation report from Current System', provided as a separate document).

Q30: *Synchronisation with cloud storage: Please can CCD clarify the aim of this requirement? Is it for backup of the database and attachments to TNA cloud storage, or for uploading content to the supplier's cloud storage?*

A30: Currently CCD does not have cloud storage, so for now we would need the supplier to provide the cloud storage.

Q31: *Active Directory Login: The Expressions of Interest Stage document states Active Directory Login is desirable but this is not mentioned in Invitation to Tender – Stage 1. Can CCD confirm if this point is still of interest?*

A31: It is something desirable but not essential, if it could be achieved within our budget that would be great.

Q32: *Timescales: What timelines do CCD envisage for modifications, testing, data migration and live running?*

A32: For modifications and testing, July to September with a system in place by the end of the third quarter in 2017 at the latest (October-December). It would be desirable to have data migration alongside this, but implementation of a system is the priority.

Q33: *What are the key operating systems that the database will work with? Is it primarily Windows?*

A33: The key operating systems will be Windows 7, possibly migrating to Windows 10 within the term of the contact.

Q34: *Are the drives that the data will be stored on internal or external?*

A34: We want the solution to be Cloud Based storage.

Q35: *How is the database currently ran?*

A35: Currently the database is an Access Database stored on an internal networked drive.

Q36: *How much involvement will your IT team have with the set up this database, will it be worked on internally or is a professional service engagement more likely?*

A36: We want a system that is supported by the supplier and requires minimum set up from our IT department.

Q37: *Can we have a sample of your data schema?*

A37: We are able to provide a sample of data (please see documents provided separately).

Q38: *What do you mean by “Develop bespoke template workflows, e.g. for Loans and for Exhibitions (to allow tracking of tasks)”?*

A38: We would like the ability to input our own template for managing projects, specifically for the Loans and Exhibitions. (Please see ‘Example Exhibition Workflow/Action List’, provided separately).

Q39: *What are the up time requirements?*

A39: We have no specific requirements, however require the greatest uptime possible. Please include your uptime statistics in your tender response.