

Annex 1 – Requirements Specification

Version dated 21 May 2019

1 Introduction

This document contains the requirements that are imposed on the System that is to be delivered.

1.1 Terminology

Wherever a term or abbreviation is used that appears in the glossary section of the Tendering Guidelines, the meaning defined in that section will be used. These terms are written with capital letters.

In addition, terms are used from the PREMIS and OAIS vocabulary lists. The Tenderer is expected to be familiar with the professional field and should thus know these concepts.

For the definition of other professional terms, Wikipedia (as consulted on 11 February 2019) is leading.

In the following requirements, the terms “User”, “Administrator” and “GUI” are used.

An “Administrator” refers in this document to a User who is authorised for management tasks. Administrator is a collective name for various types of Administrators, each with their own responsibilities.

A GUI here refers to a graphical user interface offered by the System, for instance through a web browser, without the User having to enter commands or expressions.

When reading the requirements, it is always important to allow for the fact that:

1. Where a User is involved, the requirement is met if the functionality is offered both
 - a. through an API; and
 - b. through a GUI or other user interface, for instance a command line interface.
2. Where an Administrator is involved, the requirement is met if the functionality is usable through a GUI or other user interface, for example a command line interface, unless it is explicitly stated that only a GUI is permitted.
3. Where the requirement states explicitly that a GUI is to be used (worded as “through a GUI”), then only a GUI is permitted.

1.2 Structure of the tables

The requirements are in tables of three columns:

- The left-hand column contains the code for the requirement. This is used to allow reference to the requirement. If a requirement is split into sub-requirements, then each sub-requirement has its own code, consisting of the code of the overarching requirement plus a letter. The codes of the sub-requirements are given in the middle column.
- The middle column describes the requirement. If it states here that the System has specific functionality, then the requirement is only met if that functionality can be delivered, including support, documentation etc. as per the ‘Other non-technical requirements’ (Section 10) and the conditions stated in the Tendering

Guidelines. Explanations and examples are printed in italics and, taken strictly, are not part of the requirement.

- The right-hand column contains either “KO” or a number. The indicator “KO” means that it is a knockout requirement, i.e. the System can only be accepted if this requirement is met entirely and literally. If a number is given in the right-hand column, then this represents the maximum number of points that the Tenderer can score if the System complies with the requirements. This is then a non-mandatory requirement that the System does not necessarily have to comply with.
For some requirements, there is also a “Y/N” or a “(*)” in the right-hand column. The meaning of this is explained in the following paragraphs.

1.3 How to respond to the Requirements Specification

For every requirement where there is a “Y/N” in the right-hand column, the Tenderer must indicate whether the System complies with the requirements entirely and literally, without further explanation.

For every requirement in which there is no “Y/N” in the right-hand column, the Tenderer must give an explanation in one or both of the following ways:

- by describing how the System meets the requirement.
- by referring to a specific part or parts of the documentation of the System, where there is a description of how the System complies with the requirement. In that case, the corresponding documentation must also be delivered or must be available online.

The explanation must give KB a sufficiently clear picture to assess the response as described in 1.4.

There is an immediate preference for the System complying with the requirement ‘out of the box’. If this cannot be done, there should be an indication of whether:

1. The Tenderer is integrating the System with its own products or third-party products to be delivered by KB.
2. The Tenderer is extending the System with customised solutions.

For both options, any costs associated with customisation or integration must be included in the price offered.

1.4 Method of assessing the response

The explanations that the Tenderer has given will be assessed by KB.

- If they are knockout requirements, KB may conclude based on the explanation is that a requirement has not been met, even if the Tenderer has stated that they do comply with the requirement. In that case, the Tender will be set aside.
- For the non-mandatory requirements, KB will determine a score based on the explanation. If there is a “(*)” in the right-hand column, then that same indicator is used in the central column to specify how the score for

that requirement is calculated. If there is no “(*)”, the score is calculated as follows:

- 100% of the maximum score if the System meets the requirement properly in all aspects.
- 75% of the maximum score if the System meets the requirement properly, but not in all aspects.
- 50% of the maximum score if the System meets the requirement to some extent.
- 25% of the maximum score if the System meets the requirement to a very limited extent.
- 0 if the System does not meet the requirement.

The said “aspects” are both quantitative (does the System comply with all elements of the requirement?) and qualitative (does the System provide a user-friendly and proper solution for the requirement?).
If the Tenderer is complying with the requirement by developing customised solutions, no more than 50% of the maximum score can be achieved.

One of the requirements is that all questions in the final Section of this Requirements Specification must be answered by the Tenderer. The answers to these questions do not affect the score, nor the contract award, unless KB can demonstrate the incorrectness of one or more of the answers. In the latter case, the Tender will be set aside.

1.5 Contents

The requirements are subdivided into sections according to the various functional components distinguished by the OAIS model.

Requirements that do not fit in any of these components have been included in Section 9 and 10.

Questions to which KB wants answers are listed in Section 11.

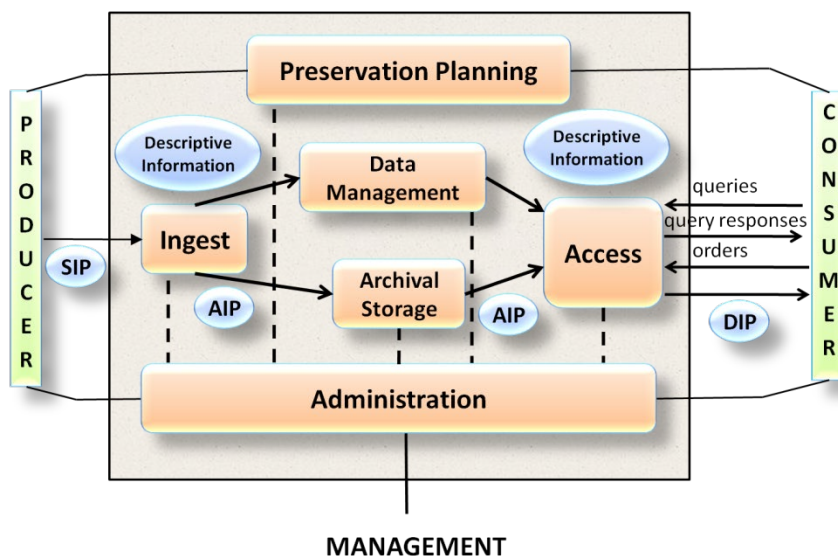


Figure 1: OAIS Functional Entities and Information Packages

- Section 2 – General Requirements
- Section 3 – Administration
- Section 4 – Ingest
- Section 5 – Preservation Planning
- Section 6 – Data Management

- Section 7 – Archival Storage
- Section 8 – Access
- Section 9 – Other technical requirements
- Section 10 – Other non-technical requirements
- Section 11 – Questions

‘Administration’ is the first functional element of this document, because it is often referred to from the other components.

1.6 Batch Processes

Many actions in the System are carried out by means of Batch Processes. We identify three types of Batch Processes:

- Ingest processes: these processes convert SIPs into AIPs.
- Publication processes: these processes convert AIPs into DIPs.
- Administrative processes: these processes mostly handle AIPs and/or Metadata, but never SIPs or DIPs.

It is not a requirement that the System distinguishes between these three types. The System could unify these three concepts into a single kind of Batch Process in which all the above is possible.

Because they are all Batch Processes, there is a great deal of overlap in the requirements imposed. The requirements for the administrative processes are discussed in Section 3. Sections 4 and 8 contain the specific requirements regarding the ingest processes and publication processes and there will be references to the relevant requirements in Section 3 to indicate what the overlapping requirements are.

2 General requirements

GE-1	The System should adhere to the basic principles of OAIS. Describe how the System follows the OAIS principles in no more than 1000 words, excluding pictures.	1000
------	--	------

3 Administration

3.1 Navigation, searching and collecting

AD-1.1	An Administrator can navigate through the structure of a Publication to all Metadata and Files.	KO
AD-1.2	An Administrator can execute searches based on complex queries in which all Metadata that is present can be searched. A complex query is a query in which several search criteria are combined using Boolean operators. It must be possible to make such a query through a GUI.	KO

AD-1.3	An Administrator can execute searches based on complex queries (see AD-1.2) in which all data recorded in the audit log can be searched. <i>Requirement AD-7.1 gives more information about the audit log.</i>	1000
AD-1.4	An Administrator can save a search as described in AD-1.2. An Administrator can execute, modify or delete a saved search.	100
AD-1.5	An Administrator can create a Set by performing a search as described in AD-1.2. In that case the Set is the same as the result of the search. <i>A Set is intended to serve as input for Batch Processes; see AD-2.3.</i>	KO
AD-1.6	An Administrator can create a dynamic Set based on a search. When retrieving this Set, the result consists of all objects that meet the search criteria at the moment of retrieval.	100
AD-1.7	An Administrator can enter a list of identifiers and save the Publications or Files concerned as a Set.	1000
AD-1.8	An Administrator can request the list of existing Sets, examine the contents of a Set, add Publications or Files to a Set, delete Publications or Files from a Set, and delete Sets.	KO
AD-1.9	An Administrator can save a search as described in AD-1.2 or the result of such a search as a Collection (an AIC). <i>A Collection is intended for grouping Publications into compilations that are stored for longer periods.</i>	100
AD-1.10	An Administrator can create a dynamic Collection based on a search. When retrieving this Collection, the result consists of all objects that meet the search criteria at the moment of retrieval.	100
AD-1.11	An Administrator can enter a list of identifiers and save it as a 'Collection'.	1000
AD-1.12	An administrator can retrieve Collections and modify and delete them through a GUI.	KO
AD-1.13	An Administrator can add Metadata to a Collection. The administrator can retrieve this Metadata and modify and delete it through a GUI.	500
AD-1.14	A Publication can belong to more than one Collection at a time.	KO Y/N

3.2 Starting and configuring Batch Processes

AD-2.1	An Administrator can add, retrieve, modify and delete a Batch Process Configuration. When adding or modifying, the Administrator can indicate which actions have to be carried out in what sequence. <i>For example, the configuration of the Batch Process for a publication process may comprise Metadata transformation, making derivatives and possibly taking further actions. The actions that are carried out and the sequence for these actions are defined in the batch configuration.</i>	KO
AD-2.2	Actions that an Administrator must as a minimum be able to include in a Batch Process Configuration are: • Checks, as described in IN-6. • Detailed checks, for example as described in AD-5.1 and AD-5.2. • Modifying and/or creating Metadata. • Modifying and/or creating Files. • Conversions of file formats as described in AD-4.3. It is possible to use an internal or external Tool for the actions described above.	KO

AD-2.3	An Administrator can execute a Batch Process based on a Batch Process Configuration combined with a Set. <i>Sets have been described in AD-1.5 ff.</i>	KO
AD-2.4	One of the possible results of a Batch Process can be that a new Collection is created or that an existing Collection is modified. <i>Collections have been described in AD-1.9 ff.</i>	KO
AD-2.5	An Administrator can schedule a Batch Process after which the System will automatically start the Batch Process at one or more scheduled moments, or at regular intervals.	KO
AD-2.6	An Administrator can retrieve an overview of the current Batch Processes and their statuses.	KO

3.3 Error handling for Batch Processes

AD-3.1	For each Batch Process, an Administrator can define which Administrator the System will send a message to if an error occurs during that Batch Process. <i>'Sending a message' can be effected in various ways. It may for example be done by e-mail, but other forms of warning are also possible, for example through the System's GUI. Describe which forms the System supports.</i>	1000
AD-3.2	An Administrator can retrieve the errors that have occurred. The System collates large numbers of messages into groups, <i>for example for each type of error, so that there is a clear overview of the error messages.</i>	1000
AD-3.3	An Administrator can flag individual errors and entire groups of errors as 'handled', after which they will no longer appear in the list of errors. <i>If the error "File <<filename>> has a size of 0 bytes" has occurred 100 times, for example, the Administrator can flag them all as having been 'handled' in one go.</i>	1000
AD-3.4	An Administrator can define which sorts of errors they want to flag automatically as 'handled' and/or they can set messages about specific types of errors to be ignored during a Batch Process so that they no longer receive error messages for those types of errors.	1000
AD-3.5	An Administrator can make a Set of any combination of errors and groups of errors. This Set then consists of all Publications or Files that these errors relate to.	1000
AD-3.6	An Administrator can offer Publications for which an error occurred during the Batch Process to the Batch Process again. The Administrator can modify the Publications first.	KO

3.4 Change

AD-4.1	An Administrator can add Files to a Publication through a GUI. This then creates a new Version of the Publication. <i>Example: adding a missing page. If necessary, the Administrator can change the sequence of the Files; this is after all structural Metadata (see DM-15 for being able to modify Metadata).</i>	KO
AD-4.2	An Administrator can replace Files through a GUI. This then creates a new Version of the Publication. The original File is retained in the old Version of the Publication. <i>Example: replacing a failed scan of a page.</i>	KO
AD-4.3	The System can convert Files from one format to another.	KO

AD-4.4	An Administrator can install Tools so that the System can carry out format conversions. These Tools can be used in all Batch Processes. <i>Format conversions are sometimes required to keep Files accessible in the longer term and to fit in with new developments in technology. Being able to install and use Tools means that this functionality is extensible and new conversions can be supported.</i>	1000
AD-4.5	The System offers an API through which an external system can carry out the actions described in AD-4.1 and AD-4.2. Carrying out these actions is a Batch Process, for which the requirements in AD-2.5, AD-2.6 and AD-3.1 through AD-3.6 apply.	KO

3.5 External Tools in Batch Processes

The glossary defines a Tool as a delimited piece of functionality. This can be a function that the System offers as standard, but can also be a function that has been added to the System by plugging in a third-party component or a component that has been developed by KB. If the System offers a function as standard, we refer to it as an internal Tool and otherwise as an external Tool. The capability of using external Tools is essential to ensure the extensibility and future-proofing of the System.

AD-5.1	The System offers the capability of using an external Tool for identifying the file format in a Batch Process, such as a Tool that determines the PRONOM Unique Identifier (PUID) of a File. <i>The PUID is more specific than the MIME type and can be used for communicating with the Digital preservation community.</i>	KO
AD-5.2	The System offers the capability of using external Tools that make it possible to check within a Batch Process whether a File complies with the requirements that the specifications of the file format impose on such a File. <i>This check can reveal issues with long-term storage of Files. The Tool checks for example whether a File that the identifier of the file format has flagged as being an EPUB 3.1 File genuinely complies with all the requirements that the IDPF imposes on an EPUB 3.1 File.</i>	KO
AD-5.3	The System offers a possibility for using an external Tool in a Batch Process for extracting technical metadata, based on the properties of the File. <i>An example of technical metadata: the height and width of an image.</i>	KO
AD-5.4	The System offers the possibility for using an external Tool in a Batch Process that can process Files with file formats that the System does not support as standard.	KO

3.6 Dashboard

AD-6.1	Every Administrator can configure their own dashboard. The dashboard is a part of the GUI on which various Monitoring components are visible at the same time, as selected by the Administrator. <i>Examples of Monitoring components are:</i> • <i>An overview of error messages that the Administrator in question has to handle (see AD-3.2).</i> • <i>A list of all running Batch Processes (see AD-2.6).</i> • <i>For every Batch Process that is dealing with a series of Publications: the number of Publications or Files that have been processed or are still outstanding.</i> <i>See also RA-6 for the positioning of reports on the dashboard.</i>	1000
--------	---	------

3.7 Logging

Logging usually means that log records are written to a text File, but other forms of logging are also permitted, for example in a relational database. At least two types of logging are required: audit logging (AD-7.1) and system logging (AD-7.2 through AD-7.3). In addition, there is a requirement for the access logging to be retained in a separate log (AD-7.4). Maintaining event Metadata is not covered by the logging, but it has been described in DM-24.

AD-7.1	The System keeps a record in an audit log of which User implements what changes (including additions and deletions) to both the Files and the Metadata. <i>Searching for this information is described in AD-1.3.</i>	KO
AD-7.2	The System maintains a system log. This is a log of what the System is doing. An Administrator can examine the system log in the form of an unformatted text File.	KO
AD-7.3	The system log distinguishes between at least three levels of logging information: • error: information about the errors that occur. • warning: information about the warnings that have been issued by the System. • information: information about all actions carried out. The System logs at all levels as standard, or if not, the Administrator can set it up so that the System does log at all levels.	KO
AD-7.4	The System maintains a separate log of all cases of access being refused to Users or external systems plus the reason why access was refused depending on the logging level. For every action, this contains at least the date and time, account name, IP address of the source and whether access was granted or refused. An Administrator can examine this log in the form of an unformatted text file, separately from the system log described in AD-7.2.	500
AD-7.5	The above-mentioned logs (AD-7.1, AD-7.2 and AD-7.4) cannot be modified or deleted via the System, or otherwise this may only be done by Administrators authorised specially for the purpose.	KO

3.8 Technical administration

AD-8.1	An Administrator can make a backup of the System and can restore the System later using that backup. After the recovery, the System will once again function properly, in the same condition as when the backup was made. The System reports any inconsistencies between the data in the System and the data in the Archival Storage.	KO
AD-8.2	An Administrator can install the System in multiple environments at the same time. <i>An environment could for instance be a virtual server. The production environment is for regular use and additionally there are a development environment, a test environment and an acceptance environment, so that new versions of the System can be developed and tested before they go into production.</i>	KO
AD-8.3	Installing the System can be part of an automated process. <i>For delivering a version from one environment to another, the Administrator uses a script in which installing the System is part of that script.</i>	500

AD-8.4	An Administrator can copy a selection of the data from the production environment to another non-production environment (for example for testing purposes). When doing so, the System enforces the integrity of the data on the following points: • Referential integrity: the referential integrity constraints in the physical data model are observed so that references are maintained. • Consistency: the selection of data in the non-production environment is the same as the same selection in the production environment.	1000
AD-8.5	An Administrator can take the System down and start it up again later, without inconsistencies arising in the data. After being rebooted, the System will once again function as well as before it was taken down. The Administrator does not have to carry out any extra actions to get the System running properly again.	KO

3.9 Other requirements for administration

AD-9.1	An Administrator can retrieve information about all Files that are being processed or have been processed by the System, namely: • Files that have been supplied (in a SIP) but have not yet been ingested. • Files that have been ingested and are therefore in the Archival Storage in an AIP. • Files that are derived from the above-mentioned Files, <i>for example by conversion to another format or by manual correction of an OCR file.</i> • Files that have been delivered as part of a DIP (see Section 8). The information that can be retrieved includes the location of the File and which of the four statuses mentioned above are applicable to the File.	100
AD-9.2	The System sends a clear error message when an error occurs in any of the technical components of the System. An Administrator can define which Administrators the System sends such error messages to. An Administrator can also examine such issues through the GUI. <i>For the meaning of the phrase 'send messages', please refer to the explanation with AD-3.1.</i>	1000
AD-9.3	Services that the System can supply to an external system give clear, documented response codes. <i>Example: an OAI-PMH service.</i>	KO
AD-9.4	An Administrator can modify style sheets and schemas and the System will then use these modifications in Batch Processes that utilise these style sheets and schemas, while the rest of the System continues to run. Describe the actions that are needed for effecting a modification of a style sheet or a schema in (for example) the ingestion process. <i>Style sheets and schemas that are used in the transformation and validation of Metadata (for example using XSLT style sheets, XSD schemas and Schematron) are part of the configuration of an ingestion process.</i>	100
AD-9.5	The System offers the capability of generating an SHA-512 checksum of a File in a Batch Process.	KO

4 Ingest

IN-1	The System offers various methods that producers can use for supplying Publications, which then go into the ingestion process, namely: IN-1a FTP and SFTP, both as a push method (the producer uploads the Publications using the FTP or SFTP protocol) and as a pull method (the System downloads the Publications via an FTP or SFTP server belonging to the producer). IN-1b External storage: the System reads the Publications from a storage device connected to the network, <i>for example an external hard disk</i>. IN-1c OAI-PMH: the System uses an OAI-PMH service from the producer. IN-1d One or more other delivery methods (to be specified). *) <i>Every other delivery method gains 10 points, up to a maximum of 100 points.</i>	100 Y/N 100 Y/N 100 Y/N 100*)
IN-2	The System can check the checksum supplied with all the delivery methods offered, at least MD5 and SHA512 checksums. <i>This constitutes a check that no Files have become corrupted during the delivery process, as described in OAIS, Quality Assurance.</i>	100 Y/N
IN-3	An Administrator can add, retrieve, modify and delete an ingestion process. The ingestion process is configurable. This means that the Administrator can state the conversions, checks and any other actions that may need to be required when adding and modifying. <i>Because the ingestion process is configurable, changed requirements and new types of Deliveries can be implemented quickly.</i>	KO
IN-4	The User can handle the ingestion of digital objects that are used for documentation. The User can relate Publications to documentation and can retrieve the documentation that is related to any Publication. Documentation can be stored in any format. <i>For example: that what is referred to in OAIS terms as 'Representation Information'. Examples: (1) A submission agreement. (2) The XML schema used and delivered by the producer.</i>	100
IN-5	The System must check whether a Publication being supplied has previously been ingested, without any changes made to the Files, the Persistent Identifier or other Metadata, in which case the Publication should not be ingested a second time and this fact should be recorded in the ingestion log.	1000
IN-6	An Administrator can include any combination of at least the following checks in the ingestion process, as indicated in IN-3: • Checking that Files are not 0 bytes in size. • Limited identification of the file format, based on the content of the File <i>and not for example based on the extension. The System recognises the MIME type, for example.</i> • Checking that Files are not corrupt. • Determining whether Files are encrypted. • Virus checking. • Checksum confirmation. <i>The aim is to check whether the producer adheres to the delivery agreements that have been made with that producer. If not, the producer will be informed and the producer should be able to resolve the issue and resubmit the Delivery.</i>	1000

IN-7	The System checks the references of Publications to objects in other systems. Checking the reference means verifying that the object is present in the other system. <i>Using this functionality, it is possible to check whether all references to information in external systems – within KB or elsewhere – are still correct, or for example whether a Publication that is referred to as being in the Institutional Repositories has not disappeared from that university. The references are recorded in the Metadata of the Publications.</i>	100
IN-8	The System can transform the valid Metadata supplied by the producer into the System's own data model, preferably by means of XSLT. <i>*) If XSLT transformations are possible, the System scores the maximum number of points. If only other transformations are possible: 100 points. If no transformations are possible: 0 points.</i>	1000 *)
IN-9	An ingestion process can be executed as a Batch Process.	KO
IN-10	The requirements AD-2.5, AD-2.6, AD-3.1 through AD-3.6 and DM-24 are applicable to all ingestion processes. <i>This means that the Administrator is informed of any errors, has sufficient information to analyse and deal with these errors, that the Administrator can retrieve an overview of the current processes, and that the relevant events will be recorded as Metadata.</i>	KO Y/N
IN-11	Errors during the ingestion process that are related to a part of the Publications must not hinder the progress of that process for other Publications. Only fatal errors which are errors of a nature such that further processing is impossible and/or meaningless, cause the entire process to stop or halt. <i>An example of a fatal error would be if an essential system component such as the database is no longer available. An example of a non-fatal error would be an error in the processing of a single Publication, for example an XML file that is invalid. The Administrator is notified of all errors; this follows from IN-10.</i>	1000
IN-12	The Administrator can configure the ingestion process in such a way that when an update of a previously supplied Publication is submitted by the producer, this update will be stored as a new Version of the previously supplied Publication.	1000
IN-13	When a Version of a Publication has passed an ingestion process successfully, all components of the AIP are stored in the Archival Storage. All Files and all Metadata that are part of a Version of the Publication are deemed to be part of the AIP. The AIP manifest is stored in a metadata format such as METS. <i>An AIP contains all Metadata plus the Files that are needed for understanding the AIP. That information is preserved by exporting it to archival storage. This is part of the exit strategy. A standard metadata format such as METS is used because the AIP is then independent of the format and the system. We are making no statements here about the physical layout of an AIP in the Archival Storage.</i> <i>The AIP manifest contains all Metadata that is related to the Files and Publications (in OAIS terms, 'Content Data Objects') in an AIP. This Metadata comprises in OAIS terms Representation Information, Preservation Description Information, Package Description and Packaging Information.</i>	1000
IN-14	A User can retrieve the configuration that was used for the ingestion from the Metadata for any ingested Publication. <i>This can for example be realised by recording which version of which configuration was used along with the Publication.</i> <i>See IN-3 and IN-6 for the requirements relating to configuring ingestion processes.</i>	100

IN-15	An Administrator can modify the configuration of an ingestion process and the System can then execute the modified ingestion process while the rest of the System continues to run. Describe the actions that are needed for effecting a modification of an ingestion process. <i>Such changes are described in IN-3. A small modification, such as adding an extra check, must be possible without having to restart the System, let alone reinstall it.</i>	100
IN-16	The System can execute multiple ingestion processes in parallel, both as instances of the same configuration and as instances of different configurations. <i>This is so that the System does not have to wait until one Publication is fully ingested before starting on the next Publication.</i>	100
IN-17	An Administrator can configure an ingestion process in such a way that the System carries out certain checks or conversions in parallel. <i>This is so that other ingestion actions for the same Publication can continue when waiting for data from another system during the ingestion of a Publication.</i>	100

5 Preservation Planning

PP-1	An Administrator can record preservation risks about Files and then retrieve, modify and delete them later. <i>Some issues that are found during the checks can be classified as a preservation risk. These are issues that could possibly occur in the future and must be monitored further.</i>	1000
PP-2	An Administrator can make a report about the preservation risks.	100
PP-3	The System has a technical registry where technical characteristics of the file formats are stored. An Administrator can add, retrieve, modify and delete information to the technical registry. <i>Examples of technical characteristics are magic numbers and the application that can be used to open Files of that particular file format.</i>	KO
PP-4	An Administrator can export the contents of the technical registry, i.e. store it in a readable and structured File outside the System. <i>This is part of the exit strategy.</i>	KO

6 Data management

DM-1	The System can store all data that is defined in the KB Data Model (see Annex 8 “KB Data Model”). Give an indication of the extent to which this model can be applied in your System and where the potential bottlenecks are. Use a maximum of 1,000 words for your explanation (excluding images). <i>The System's Data Model does not have to be identical, but it must be possible to put the data of KB in the new System without loss of information.</i>	1000
DM-2	Annexes 8b, 8c and 8d are examples of XML files from KB's current DAMPS. Four blocks of data have been highlighted in them. These are: 1. A block with structural Metadata; see the KB data model ENTITY: Object/file – RELATION: Object/file <> Object/file – structural map sections and the XML files in a. Annex 8d; b. Annex 8b. 2. A block with preservation Metadata; see the KB data model ENTITY: Event and the XML file in Annex 8c. 3. A block with Metadata supplied by the publisher; see the KB data model ENTITY: Intellectual Entity/Publication - Descriptive metadata - SUPPLIED METADATA and the XML file in Annex 8d. 4. A block with descriptive Metadata stored as the supplied Metadata (for instance as it is supplied from a catalogue system); see the KB data model ENTITY: Intellectual Entity/Publication - Descriptive metadata - CATALOG RECORD and the XML file in Annex 8c. For the elements, including the attributes in these blocks for each element and attribute, state in Annex 8a how the information in the Data Model of your System is stored by stating: - The location where the element or attribute is stored in your Data Model; - The element in which it is stored; and - How the value as stated in the XML file is stored.	5500
DM-3	A User can enter, examine, modify and delete the following data about Publications as individual data elements, without expanding the System's Data Model: DM-3a Relationships with other Publications and parts of other Publications (<i>various elements are for example Files and groups of Files</i>) DM-3b Relationships with objects in other systems (see also IN-7) DM-3c Hierarchical relationships between various levels of a Publication, <i>for example the levels 'journal', 'year', 'issue' and 'article'.</i> DM-3d the type and subtype for each relationship, <i>for example as per 'Relationship Type' and 'Relationship Subtype' in the Preservation Schemes on http://id.loc.gov/vocabulary/preservation.html.</i>	100 100 100 100
DM-4	An Administrator can add new Metadata elements and validation rules to the System's Data Model by configuring them. <i>An example of a validation rule would be a regular expression for a Metadata element.</i>	100
DM-5	Adding validation rules to the Data Model can be done by the Administrator in a standardised language, <i>for example XSD or Schematron.</i> State what validation languages are supported by the System.	100

DM-6	The System offers the same functionality for new Metadata elements as for Metadata elements that are in the Data Model as standard. Describe any restrictions. <i>Searches can for example be done in the same way on these new elements as on standard elements.</i>	100
DM-7	A User can retrieve Publications at several logical levels. A User can also retrieve only the Metadata for one of these stated levels, in so far as it is present. <i>Examples of logical levels are Intellectual Entity, Representation, groups of Files and the individual Files that make up the Publication.</i>	KO
DM-8	Every logical level of a Publication can have its own Persistent Identifier.	KO
DM-9	A User can query which Versions of an object are present. An object is in this case a Publication or part of a Publication.	1000
DM-10	A User can retrieve every Version of an object. If the User does not state a specific Version, the System supplies the most recent Version. See DM-9 for the definition of an 'object'.	1000
DM-11	A User can search for a Publication using the Publication's Persistent Identifier. If the publication has several Persistent Identifiers, the User may then use each and any of those identifiers to search for the Publication. This can be done using a Persistent Identifier supplied by the producer, including a DOI, an ISBN or an ISSN.	KO
DM-12	The System assigns a Persistent Identifier in the NBN format to every Publication, using an external Tool if necessary. <i>Example: The Administrator can configure a step in a Batch Process that calls an external Tool to be built by KB, that supplies an NBN. The NBN is stored in a metadata field defined by the Administrator. Please refer to the glossary for more information about NBN.</i>	KO
DM-13	DM-13a A User can modify the Metadata of a Publication. The capabilities for modifying Metadata include as a minimum adding, modifying and deleting values as well as modifying the sequence of the Files by changing the structural Metadata. DM-13b The System records every modification action on the Metadata of a Publication automatically as a new Version of that Publication. <i>This is intended for correcting errors in the Metadata, for example the date of publication, altering them for a new situation, for example altering the changed access rights for a defined level (Publication, File, part of a File, group of Files, etc.) in the Metadata or for adding to them, for example stating the ISNI of an author or a relationship between two Publications. Modifying the Metadata in a Batch Process has been described in AD-2.2.</i>	KO KO
DM-14	The system records why the Metadata was modified (if specified by the User), by whom and when.	1000
DM-15	The User can add and modify Metadata (as described in DM-13) through a GUI, for which the User does not have to understand a formal notation. Describe the possibilities for changing Metadata through a GUI and state functionalities that benefit the usability. <i>An example of a formal notation would be XML.</i>	100

DM-16	The System checks whether the changes to the Metadata comply with the Data Model, including the validation rules that are defined in the Data Model. The System gives feedback on all discrepancies: • a warning in the GUI, if the change is being done through a GUI; • an error message to the Administrator if it is a change made by a Batch Process. Requirements AD-3.1 through AD-3.6 are applicable to these error messages. <i>Examples: Mandatory Metadata missing for a Publication. A Persistent Identifier does not comply with the format for that Persistent Identifier.</i>	1000
DM-17	The System can implement Metadata changes that are supplied from another system where such changes comply with the System's Data Model. Any such action can be started both automatically and manually by a User. <i>For Publications of which the Metadata is also stored in other systems, such as in any cataloguing system, there should be a method of synchronising the Metadata with DAMPS when the Metadata is modified in the other system. Synchronisation is done in just one direction, from the other system to DAMPS. The data may be supplied through an API (see OT-4) or retrieved by DAMPS.</i>	1000
DM-18	An Administrator can query, create and modify authority files and controlled vocabularies. The Administrator can define in the Data Model that certain Metadata attributes may only have values that appear in such an authority file or controlled vocabulary. <i>This makes it possible for the System to check whether a Metadata attribute that is supposed to refer for instance to an event or to an Agent does actually contain a valid value.</i>	100
DM-19	An Administrator can delete a Publication. <i>This is only done in exceptional cases and always in consultation, according to specific work instructions.</i>	KO
DM-20	When a Publication is deleted, the Administrator can choose whether the Metadata should also be deleted or whether the Metadata should be retained. <i>This is sometimes needed is a Publication has been added incorrectly (in which case the Metadata is also deleted) or if it is determined that it is not – or no longer – permitted to retain the Publication (in which case the Metadata is retained).</i>	100
DM-21	An Administrator can configure it so that the System must use the 'four-eyes' principle when deleting Publications. In that case, the System enforces another Administrator to sign it off when an Administrator wants to delete a Publication. It is then possible using the role-based authorisation model as described in OT-7 to make sure that each pair of eyes belongs to a different role.	1000
DM-22	The System offers the capability of checking whether relationships between objects in the System are valid, i.e. that the related objects are present in the System. Objects are in this case Publications, Files and documentation (see IN-4). <i>There may be multiple types of objects.</i> <i>When two objects are related to each other and one of the objects is deleted, the relationship is no longer valid.</i>	100
DM-23	The System processes Files with filenames in which special characters appear in the same way as Files with filenames in which only alphanumeric characters appear. Alphanumeric characters are the characters 0 through 9, A through Z and a through z. Special characters are all UTF-8 characters other than • the alphanumeric characters, • the control characters (see https://en.wikipedia.org/wiki/Unicode_control_characters), • and characters that the operating system does not permit in filenames. <i>Examples of special characters are '&', '<', 'ü', 'Ø' and space.</i>	KO Y/N

DM-24	The System records every non-trivial action carried out on a Publication as event Metadata. This also applies for actions executed by Tools or other systems connected to the System. Trivial actions such as examining the data of a Publication do not have to be recorded as an event. Because events are Metadata, requirement IN-13 also applies here, stating inter alia that Metadata is part of the AIP. The System records information for every event such as the type of action, the Agent who carried out the action and the date and time on which this was done.	KO
-------	--	----

6.1 Reports

For concrete examples of the reports that KB needs, see section 11.3 Q-3.2.

RA-1	A human User can define reports and then have them generated by the System.	KO
RA-2	With the help of a GUI, a User can define reports based on all available Metadata and the audit log. The User can then have the defined reports generated by the System. <i>The minimum content of the audit log is described in AD-7.1.</i>	100
RA-3	The System can supply every report in the file formats CSV, PDF and XML. List all the file formats that are possible for reports.	100
RA-4	A User can state that the System must generate a specific report periodically and specifies the frequency. The System then generates the report automatically at the given frequency.	100
RA-5	Once a periodic report has been generated, the System can: a) send it to a specific e-mail address; b) send it to a specific FTP or SFTP server or store it at a network location; c) show it on the User's dashboard.	100
RA-6	An Administrator can place reports on their dashboard (see AD-6.1) and determine how they are laid out. The dashboard contains reports or graphs that are relevant for the Administrator, providing information at a glance for the day-to-day administration. The System may also place the reports on a different dashboard from the Monitoring dashboard that was mentioned in AD-6.1.	100

7 Archival Storage

KB has its own storage system for Archival Storage that the DAMPS must use.

AS-1	The System uses a storage system (Object Storage) for storing Files that is compatible with Amazon S3. <i>State in the explanatory notes whether extensions are also supported.</i>	KO
AS-2	The System can use a storage system for storing Files that is compatible with at least one of the protocols NFS and CIFS.	KO
AS-3	An Administrator can configure the System in such a way that the System makes use of multiple storage systems. <i>There could for example be a different storage system for AIPs than for all other Files (SIP's, DIP's, configuration files).</i>	100

8 Access

AC-1	The way in which DIPs are made is configurable within the System. An Administrator can add, modify and delete DIP configurations. In a DIP configuration, the Administrator specifies, among other things, the Metadata format to be delivered, the types of Files that are delivered and what Collection or Collections the DIP belongs to.	KO
AC-2	The System can redeliver all delivered Publications (stored Files and Metadata) from a given producer back to that producer. <i>Some institutions want to store the Files themselves (again) that they give KB to keep, for example after a disaster or redelivery after digitisation.</i>	100
AC-3	An Administrator can configure one or more publication processes for each publication environment. When configuring a publication process, the Administrator states which DIP configuration must be used for which Publications and what delivery method is applicable. <i>This ensures – depending on the publication environments (e.g. delpher.nl) – that another type of DIP can be delivered (e.g. other Metadata fields or another Metadata format) and that a different delivery method can be used (e.g. placing on a particular disk, or placing them where they can be opened using OAI-PMH).</i>	500
AC-4	The publication process is a Batch Process, for which DIPs are created and to which the requirements in AD-2.1 through AD-3.6 apply. <i>Not all the actions stated in AD-2.2 are relevant for publication processes. An action could be executing a Tool that creates those DIPs.</i>	500
AC-5	The System has its own publication environment that is suitable for direct access. <i>This could replace the KBGA either entirely or in part in the future.</i>	100

9 Other technical requirements

OT-1	The System provides an OAI-PMH interface that other systems can use for retrieving data from the System. It must be possible to configure this in a way that defines which Metadata elements can be retrieved using OAI-PMH. <i>State in the explanatory notes what possible restrictions there may be on querying this data. The KB assumes that all metadata can be fetched from the System, in the same granularity as they were ingested.</i>	100
OT-2	The System provides an SRU interface that other systems can use for searching for data in the System. It must be possible to configure this in a way that defines which Metadata elements can be searched through using SRU. <i>State in the explanatory notes what possible restrictions there may be on searching through this data. KB assumes that all metadata in the System can be searched, in the same granularity as they were ingested.</i>	100
OT-3	The System provides an IIIF Image API that other systems can use for retrieving images from the System.	100 Y/N
OT-4	The System provides an API that allows other systems to add Files and Metadata to the System and to update Files and Metadata. <i>State in the explanatory notes what possible restrictions there may be.</i>	100

OT-5	Does the System provide any APIs other than those listed in OT-1 through OT-4, and if so which? <i>*) Every additional API gains 10% of the maximum number of points, up to a maximum of 100%.</i>	100*)
OT-6	The System provides an SDK that allows the System to be extended. Describe what a software developer can do with this toolkit.	100
OT-7	A User can only use the functionalities of the System for which they have been authorised beforehand by an Administrator. To that end, the Administrator can define roles and state for each role what authorisations are associated with it, then assigning Users to the roles. <i>The System therefore applies authentication and authorisation when other systems make use of the APIs that the System provides.</i>	KO
OT-8	The continuity of the System does not depend on systems and Files that are outside the control of KB; in other words, if these fail, the System continues to meet the 'must have' requirements. <i>The System thus continues to function if the external systems or resources are not available. An example could be the schemas of the Library of Congress that are not accessible because the servers cannot be reached. Links to external systems that KB has realised for itself are not covered by this requirement.</i>	KO
OT-9	An Administrator can install the System on servers of KB that are located in a data centre of KB.	KO
OT-10	All the data traffic between the servers themselves and between servers and client systems is encrypted at the application level.	100
OT-11	The System uses one of the following database platforms: • Oracle Database version 12.1 or newer. • MySQL version 8 or newer.	KO
OT-12	The System is able to run in an environment based on Red Hat Enterprise Linux 7.5 or newer, or on Microsoft Windows Server 2019 or newer.	KO
OT-13	The System is able to run on virtualised server hardware.	KO
OT-14	The System has been developed in line with at least the OWASP secure coding practices (https://www.owasp.org/images/0/08/OWASP_SCP_Quick_Reference_Guide_v2.pdf) or coding guidelines with comparable content. <i>State which secure coding guidelines have been used. Send them with the tender or refer to a website.</i>	500
OT-15	Users log in on the System using their KB accounts. This is achieved through a link to Microsoft Active Directory (ADFS with SAML or OAuth).	KO
OT-16	Should passwords be stored in the System itself despite the link described in OT-15, then those are stored as a one-way hash.	KO

10 Other non-technical requirements

10.1 Documentation

ON-1.1	All documentation about the System is in English or Dutch.	KO Y/N
--------	--	-----------

ON-1.2	The following documentation is available: • A user manual. • Documentation that specifies how the System can be connected to other systems. Example: specifications of the APIs provided by the System. As a minimum, the input and output of APIs is clearly described, including the possible error messages. • Documentation about the System's Data Model. • Technical specifications, including about what minimum hardware and software configuration is needed for the System to function properly. • Technical management documentation: ○ An installation procedure for an environment on which the System has not yet been installed. ○ An installation procedure for an environment on which a previous version of the System has been installed. ○ A description of the available configuration options. ○ A description of the tasks that the Administrator must carry out in order to let the System function without problems.	KO
ON-1.3	Documentation is available that specifies how the System can be extended with new Tools. <i>Example: specification of an SDK.</i>	100

10.2 Volume, performance and scalability

ON-2.1	The System can ingest 6,000,000 Files per month, with an average size of 1 MB. <i>Describe what is needed in order to achieve this.</i>	KO
ON-2.2	The System functions well with more than 1 billion Files and is scalable for growth as defined in ON-2.1 for 10 years. If the System has a theoretical maximum in terms of the number of Files, state this in the explanatory notes.	11000
ON-2.3	Given an IT infrastructure that complies with KB's requirements (see requirements AS-1 through AS-3 and OT-9 through OT-13) and the specifications to be drawn up by the Tenderer, and given the required volume of content (ON-2.1 and ON-2.2), it must be possible for 10 Users to work on the System simultaneously with acceptable response times. The term 'acceptable response times' will be taken to mean at least the following: • Retrieving the Metadata for a Publication based on the system-specific identifier: less than 3 seconds. • Opening an image of 10 MB: less than 5 seconds. • Storing a change to a single Metadata field of a Publication, which thus includes creation of a new Version of the Publication: less than 5 seconds. <i>For example, if 5 Users are waiting for a report, 5 other Users can open image files of 10 MB within 5 seconds.</i>	100
ON-2.4	When scaling up the system as defined in ON-2.2, it is possible to ensure that the response times and the time taken by the Batch Processes and by report generation do not increase by placing additional servers with similar specifications to the existing servers ("scaling out"). *) <i>When scaling up can only be achieved using servers with different specifications ("scaling in"): state those specifications. An alternative such as this scores 50% of the number of points. If performance-neutral scaling up is only possible to a certain maximum volume, specify this maximum.</i>	100*)

10.3 Other

ON-3.1	There is a community of Users of the System who come from various organisations. Describe how this community is organised.	100
ON-3.2	The community described in ON-3.1 has an effect on the direction in which the product develops. The Tenderer offers the possibility of submitting changes to the System on behalf of the community.	100
ON-3.3	The System stores all Files and Metadata that belong with a Publication in the Archival Storage in such a way that every Publication in principle without needing additional information can be restored to a package of Files and all associated Metadata. <i>This is the core of the exit strategy.</i>	1000
ON-3.4	When installing a Tool, information about that Tool is recorded, namely at least the name and version of the Tool. This can be done automatically by the System or manually by an Administrator.	KO
ON-3.5	The System can store every File, irrespective of the file format, in the Archival Storage, including as a component of a Publication.	KO
ON-3.6	Multiple Users can modify data in the System simultaneously, but the System ensures that one User does not overwrite the data that another User has entered. Explain how this will be achieved. The term 'data' is not referring only to Publications, but genuinely to all data that can be modified by a User, i.e. also for example Batch Process configurations and other configuration data.	100
ON-3.7	An Administrator can set up the GUI for each group of Users in such a way that the GUI is matched to the requirements of the User type for each type of User. State how these settings improve the usability of the System for the various groups of Users.	100
ON-3.8	The Tenderer provides consultancy services to support the implementation of the System.	KO
ON-3.9	The Tenderer provides training to the KB staff concerned. The training courses and the training materials are matched to the various user groups within KB.	KO
ON-3.10	The Tenderer is certified to ISO 14001 or an equivalent. <i>KB may ask the Tenderer to provide proof of this.</i>	100 Y/N

10.4 Requirements relating to the services provided

As explained in the Tendering Guidelines, the finalised content of the service level agreement (SLA) will be agreed with the Contractor after the contract is definitively awarded. In this Requirements Specification, we have listed a number of requirements below that must be addressed properly in the SLA, within the context defined here.

Generic requirements relating to the services provided

ON-4.1	All communication will take place in English or Dutch.	KO
ON-4.2	KB can make agreements about the term within which the requested consultancy services are provided.	KO
ON-4.3	If the context of the agreed services means that a visit is needed to physical facilities in the buildings themselves, the Contractor will provide the services in question on site. In this case, KB will ensure the accessibility of the locations.	KO
ON-4.4	Agreements can be made about access to systems belonging to KB by the Contractor's members	KO

	of staff for specific support or for consultancy objectives. If necessary (for example when a SaaS solution is used), a Data Processing Agreement can also be signed.	
ON-4.5	What is known as an escrow agreement under Dutch civil law can be signed between the Contractor and KB. This is an agreement in which the source code can be made available to KB in the event of severe delivery problems on the part of the Contractor that are acute and structural in nature (e.g. bankruptcy) in order to ensure continuity for KB. For more information about the meaning of this, please refer to https://nl.wikipedia.org/wiki/Escrow-overeenkomst .	KO
ON-4.6	Agreements can be made to the effect that the System must not cause any data losses as a consequence of e.g. bugs in the System. This applies to both the System of the Contractor and external tools supported by the Contractor. The Contractor cannot be held liable for incorrect use and incorrect configuration of the System where this is caused by KB itself.	1000
ON-4.7	The Contractor provides support for the exit strategy if migration to a new System is required at the end of the contractual period. The Contractor provides advice to ensure a smooth transition from the old System to the new System. This advice covers information about how the data and Metadata from the old System can be migrated.	1000

Requirements relating to incident management and problem management

ON-4.8	KB can submit incidents 24/7 by e-mail or through an online ticketing system.	KO
ON-4.9	KB can make agreements with the Contractor about advice on the resolution (temporary or structural solutions) of incidents, problems and notifications, making associated agreements about possible prioritisation levels for these incidents, problems and other notifications. Depending on the priority level, a response time and a target time for a solution will be agreed. When an incident is submitted, KB determines the priority level.	KO
ON-4.10	KB may make agreements with the Contractor about the escalation procedure to be followed in an SLA to be agreed later.	KO

Requirements relating to change management and release management

ON-4.11	The Contractor issues regular updates and makes them available to KB so that the System remains up to date in terms of technological developments and other developments relating to digital sustainability.	KO
ON-4.12	The Contractor will not simply delete functionality but will announce this well in advance if this is functionality that is made available through the System to other external systems. The Contractor will support at least one major release prior to the current version.	KO
ON-4.13	If serious security issues are diagnosed, KB will be informed about them immediately. The Contractor will then state what actions KB could take in order to limit the security risk and will make issuing modified software a high priority where relevant.	KO
ON-4.14	The Contractor will provide a Release Calendar.	KO
ON-4.15	If the support for any one of the components listed below expires during the term of the contract and it requires an upgrade, the Contractor will ensure that DAMPS can handle the new version or can be made suitable for it. • O/S platform	KO

	• Database platform • Third-party tools: only where the Contractor supports them as part of the solution and where they are needed for DAMPS to function.	
--	--	--

Requirements relating to SLA management/reporting

ON-4.16	KB can make agreements about what is referred to as the 'Service Window' (whereby the time zone of Amsterdam will be used as the default) or about the opening times, distinguishing between the window during office hours and any extended opening times beyond that for specific and/or urgent situations.	KO
ON-4.17	KB can make agreements about an escalation process that will allow at least the following categories (not an exhaustive list) to be distinguished: a) the urgency of the report, b) discussion about the delivery (delivery time, SLAs and KPIs), and c) discussion about the results achieved for projects and services.	KO
ON-4.18	The Contractor and KB will discuss outstanding issues regularly, at moments to be defined further.	KO
ON-4.19	Periodic reports can be produced (e.g. monthly, quarterly and annually) about all notifications, broken down into support requests and consultancy.	KO

11 Questions

QU-1	All questions in this Section have been answered truthfully.	KO Y/N
------	--	-----------

11.1 Administration

Q-1.1	What options does the System provide for checking the checksums of Files that have been stored and what algorithms are supported for this? How can an Administrator carry out this action?
Q-1.2	The external Tools that the System uses must be tested regularly, for example when there are new versions of software or when certain Files are not processed correctly by the Tools. The tests must verify that the Tools are fit for purpose and that they respond appropriately to errors. The System supports these tests. Describe how – for example after updates to Tools are installed – the effectiveness of the Tools can be tested in a test environment with a standard test set of Files, so that updates to Tools do not introduce problems and the results are as expected.

11.2 Ingest

Q-2.1	Copying and moving data can entail risks for the data. The System must therefore copy and move data as little as possible. How is it possible to reduce the number of copy actions as far as possible? What copy actions are unavoidable and why?
Q-2.2	Does the system provide a web GUI allowing producers to upload Publications that then progress automatically into the ingestion process?

11.3 Access

Q-3.1	What reporting is delivered as standard with the System?
Q-3.2	The reports needed by KB include those listed below. Based on RA-2, we assume that the System is capable of delivering these reports. Do you agree with this? If not, what are the problems? • The following information for each producer: ○ an overview of all their Publications in our KB collections, ○ the numbers processed in the past quarter or past year, ○ the total number of Publications in DAMPS, ○ the date of the last Delivery. • The following information in a single report for every combination of producer plus publication type: ○ the numbers processed in the past quarter or past year, ○ the total number of Publications in DAMPS. • The following information in a single report for every combination of producer plus publication type: ○ the sum of the sizes of all Publications, ○ the sizes of the smallest File, the smallest Publication, the largest File and the largest Publication, ○ the average number of Files per Publication, ○ the average size of a File. • An overview of lapsed embargoes that will have lapsed after a given date (defaulting to today). • An overview of trigger events and how many Publications from which producers are involved in this. A trigger event occurs when a producer stops issuing a newspaper or magazine. • An overview of the file formats and how many Files are in each format. • An overview of encrypted Files. • An overview of all manual and other changes that have been made in a given time interval, for example last week. • An overview of the number of unresolved errors that occurred during an ingestion process. • An overview of Publications that have been in the ingestion process for longer than a week and have not yet been processed. • Preservation risk reports. • An overview of the various ingestion processes and their configurations. The style sheets and schemas used are particularly important. • An overview of the Files with the information as described in AD-9.1.

11.4 Other technical questions

Q-4.1	What component creates the bottleneck for performance?
-------	--

11.5 Other non-technical questions

Q-5.1	ON-3.9 states <i>“The Tenderer provides training to the KB staff concerned. The training courses and the training materials are matched to the various user groups within KB.”</i> Given the Tenderer’s experience, what specific experience and expertise do the Users and Administrators at KB need, in addition to following the said training course, in order to be able to use, configure and administer the System successfully? Please also list any specific knowledge or experience that is not strictly necessary but that will ensure that Administrators and Users will be capable sooner of carrying out their tasks.
-------	--