GIVE – newspaperproject Primeur

Request for Quotation

Background document v.1.0

English translation

meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

Europese Unie

EFRO
EUROPEES FONDS
VOOR REGIONALE
ONTWIKKELING

Disclaimer

Transparency and knowledge sharing are core values for meemoo. That is why we share our main tender files from the domains of digitization and archiving. This way everyone can see how we work or get inspiration. We would like to include this warning and disclaimer:

This document has been drawn up in the context of a specific project, with a specific objective, timing and budget, which in turn are rooted in all kinds of circumstances, meemoo's vision on digitization and so on. During or after the execution of the project, this vision, circumstances or other context elements may change or have changed. Furthermore, this document may also contain errors. The person who copies these documents in whole or in part is and remains fully responsible for assessing the consequences for his interests. It is best to obtain sound legal advice, especially - but not exclusively - for copying the legal provisions.

Meemoo cannot be held liable in any way for any disadvantage that any party may incur by making available or copying all or part of these documents.

Table of contents

meemoo vzw | Ham 175 | B-9000 Gent | BTW BE 0644.450.380 | RPR Gent, afdeling Gent
www.meemoo.be | info@meemoo.be | T: +32 9 298 05 01

1 Vision on digitization

1.1 Vision

Good digitization offers an answer to the classic challenges related to access and conservation. On the one hand, part of the Flemish heritage is threatened by conservation problems such as the physical deterioration of the carriers. If we want to save this part of our heritage, we need to digitize it. On the other hand, digital access to content offers many new possibilities. Therefore, it is necessary that there is sufficient digital content available. Today that content is digital born, but analog content from the past must be digitized as well. This demand is particularly high for newspapers, as these sources enjoy a higher level of engagement than other digitized heritage materials to date.

1.2 Meemoo and its partners

Digitization of heritage is a technical and expensive challenge. The influence of both factors can be significantly reduced by centrally coordinating the operation and thus achieving economies of scale. Meemoo takes on this role by bringing different parties together, launching digitization tenders and leading the digitization projects itself.

Meemoo believes that mass digitization can go hand in hand with high quality, emphasizing the importance of professional and transparent project management and good communication and understanding between all parties involved. The following are the parties within the digitization projects of meemoo:

1.2.1 The 'content partners'

The administrators of the Flemish cultural heritage are, as it were, the customers of meemoo. Meemoo offers them digitization, sustainable storage and access. Meemoo also initiates dissemination initiatives with their content, always fully respecting the ownership, copyright, commercial or ethical rights of the materials. Meemoo leads this digitization project for them, but the content providers retain all rights on the material. Meemoo therefore takes a commitment and has a responsibility towards them. When choosing a digitization partner (service provider), meemoo will take this responsibility into account by extensively involving the content providers in the decision-making processes and by opting for a high-quality quotation. In this project, the content partners involved are the following:

Name of the content partner	Estimated newspaper titles	Estimated newspaper pages
ADVN	30	40.000
Amsab-ISG	14	85.000
Bibliotheek Hasselt Limburg	6	50.000
Erfgoedbibliotheek Hendrik Conscience	7	115.000
KADOC	16	105.000
KU Leuven Bibliotheken	1	115.000
Openbare Bibliotheek Brugge	2	5.000
Universiteitsbibliotheek Gent	2	115.000
Total	**78**	**630.000**

Fig. 1 – involved content partners with their estimated titles and pages

1.2.2 Co-coordinating partner: Flemish Heritage Libraries

For the entire Primeur project, meemoo works together with a specialized partner from the cultural heritage sector. The non-profit organization Flemish Heritage Libraries (hereinafter VEB) is the national service organization for institutions in Flanders and Brussels that manage library heritage. The organization strengthens professional and sustainable heritage activities in these institutions through services such as databases, tools and projects. As an expertise network, the organization stimulates knowledge sharing, connection and collaboration. In this way, it contributes to well-preserved and freely accessible library heritage that can provide pleasure, knowledge, insight and inspiration to every citizen in our diverse society. The VEB is strongly committed to digital services. Since its foundation, the organization has been developing standardized, central accesses to library heritage in Flemish collections in collaboration with more than 100 preservation institutions. These databases are an indispensable springboard for overarching digitization projects for library heritage. The preservation of the enormous Flemish collections of historical newspapers has been high on the VEB's agenda for years: large parts of this source, indispensable for our history, are seriously threatened by autonomous decay that continues steadily despite all good (precautionary) care. As administrator (since 2009) of the umbrella newspaper catalog Abraham, main partner of the newspaper digitization project News of the Great War and executor of the projects Abraham 2020 and Nieuwe Tijdingen, the VEB gained in-depth expertise on Flemish newspaper heritage, newspaper digitization, registration of newspaper collections of various heritage institutions, and data management.

1.2.3 The 'Service providers'

The executor(s) of the digitization, usually designated via a tender procedure. Further named as the service provider (SP).

2 This project

2.1 Wider context

meemoo vzw | Ham 175 | B-9000 Gent | BTW BE 0644.450.380 | RPR Gent, afdeling Gent
www.meemoo.be | info@meemoo.be | T: +32 9 298 05 01

Primeur is the newspaper digitization project within the general GIVE-project (Coordinated Initiative for Flemish Heritage Digitization). This project is part of a broader range of digitization projects of the Flemish recovery plan. Primeur is part of the Flemish Resilience Recovery Plan and is being realized with the support of the European Regional Development Fund. With this project, meemoo is taking a step towards preserving the paper heritage in a sustainable way and making it digitally accessible in addition to the audiovisual heritage.

Research by the VEB has shown the importance of the digitization of newspaper material, as newspapers are among the most frequently consulted digitized heritage materials. The request for consultation comes from a wide audience (researchers, educators, local historians, students, etc.).

The combination of the poor paper quality on which many of these publications are printed and the high degree of consultation, puts this archival source under serious threat. Even when the storage conditions are optimal, the natural decay of the paper due to acidification is high. The paper becomes brittle and the risk of damage when the material is manipulated is very high. Digitization is therefore absolutely necessary to ensure the preservation of the content of these documents for the future and to make and keep them accessible. Appendix 4 'description of the material' discusses the precise condition of the material to be digitized.

Parallel with Primeur, the VEB and meemoo are also working together on Nieuwe Tijdingen, a three-year project in which they determine the most suitable conditions for a program for the preservation of the entire Flemish newspaper heritage. Among other things, Nieuwe Tijdingen is realizing a plan for long-term and large-scale newspaper digitization with dozens of collection-managing partner organizations. It brings together existing digital newspaper collections and realizes high-quality access to digitized newspapers. Nieuwe Tijdingen therefore ensures that the historical newspapers that are digitized within Primeur can also be optimally made available, together with newspapers digitized in the past or in the future.

2.2 Scope of the project

2.2.1 Digitization within the context of GIVE-project 'Primeur'

In the context of the GIVE project in which material from the Flemish cultural collections is made digitally available and accessible, this project aims to realize a large-scale digitization of vulnerable newspapers. Within the Primeur project, the aspect of digitization serves multiple purposes:

- In view of the natural decay of the carriers themselves – which ultimately have a self-destructive character – the creation of a digital copy that is as faithful as possible to and can eventually replace the analog original.
- Making all inventoried material digitally accessible and searchable in a smooth and user-friendly manner embedded in existing or new platforms of both meemoo, the content providers, and other organizations that, if possible, use the principle of 'open data' to transfer the material to different sources, and make it accessible.

All this must be done within the budget and timing limits set by the Board of Directors of meemoo, and with the full agreement and cooperation of the co-coordinator VEB and the collection administrators of the carriers, the so-called content partners. In this project, the group of content providers involved consists of about eight cultural heritage institutions (see fig. 1).

2.2.2 Additional digitization batch(es)

In addition to the digitization of newspapers as part of Primeur, meemoo and the VEB foresee possible additional newspaper digitization as part of the Nieuwe Tijdingen project (see also section 2.1).

This possible additional digitization will amount to approximately 50,000 newspaper pages and will be carried out in collaboration with 3 to 4 content partners in Flanders who are not part of the Primeur

EFRO EUROPEES FONDS VOOR REGIONALE ONTWIKKELING

meemoo vzw | Ham 175 | B-9000 Gent | BTW BE 0644.450.380 | RPR Gent, afdeling Gent
www.meemoo.be | info@meemoo.be | T: +32 9 298 05 01

partners. If this digitization is requested, it must be performed after the digitization of all Primeur batches. The physical condition of the pages and bindings is at least as good as that within Primeur (see appendix 4).

This possible extension is mentioned in this document as a mandatory option OP03 under section 4.9.

2.3 Phasing of the project and timeline of digitization

	2021		2022						2023	
Phase	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
1. Preparation										
2. Selection										
3. Tender										
4. Registration and packaging										
5. Digitization										
6. Quality check										
7. Ingest and processing of the files										
8. Completion										

fig. 2 - high level timeline of the project

The actual digitization will start in phase 5 'Digitalization'. The files are then recorded/saved on the digital storage infrastructure and the archive environment of meemoo and made available to the partner institutions.

For the digitization phase (phase 5), the planned timing is given below as an indication, which meemoo wishes to adhere to in the implementation of the project. For the timing of the tendering procedure, reference is made to the 'Overview document':

- Second/third week of April 2022: First consultation with tenderer(s).
- Last week of April 2022 Further details of the collaboration, start of the test phase.
- First week of June 2022: Start of the pilot phase.
- Beginning of July 2022: Start of the production phase.
- End of July 2023: Expected end of the production phase.

meemoo vzw | Ham 175 | B-9000 Gent | BTW BE 0644.450.380 | RPR Gent, afdeling Gent
www.meemoo.be | info@meemoo.be | T: +32 9 298 05 01

The final phase is the final dissemination of the digitized material through channels that will be developed by meemoo, through existing websites such as Abraham, online catalogs and access platforms of the content partners and through other channels, e.g. educational websites.

2.4 The carriers

From August 2021 to November 2021, the Flemish Heritage Libraries completed the first inventory phase for this project. At the eight repositories, the newspapers nominated for digitization were mapped

The appendix **'Description of the material'** contains a list of the selected newspapers for digitization and some of their characteristics, including some guiding statistics. When preparing this request for quotation, meemoo did not have the exact quantities of the carriers. Therefore, the quantities are indicative. They do not bind meemoo in any way.

3 Business case: the digitization process from start to finish

Below is an outline of what the digitization process looks like from start to finish, in order to make it clear to the candidates in which stages and in what context the work of the tenderer (in that case: service provider) will be. This process is detailed in consultation with the service provider after the award of the contract, without extending the contract.

3.1 Registration and barcoding

EFRO EUROPEES FONDS VOOR REGIONALE ONTWIKKELING
meemoo vzw | Ham 175 | B-9000 Gent | BTW BE 0644.450.380 | RPR Gent, afdeling Gent
www.meemoo.be | info@meemoo.be | T: +32 9 298 05 01

All carriers that will be digitized are registered with and by the content partner. In the first place, this means that the newspapers receive a barcode (A-barcode) on a farde around the carrier (in case of loose editions) or on a flyleaf (in case of bound editions), and that a number of characteristics per carrier are included in the database for registration and logistics.

All newspapers are placed in a collection box, where the collection box is given its own barcode (b-barcode). The content partner carries out this work with the support of the VEB, and meemoo facilitates it by offering a registration database, clear instructions, packaging, endpapers, barcodes, etc. If possible, existing data from the collection management systems of the content partners, the information from the Abraham database and the catalog of Belgian newspapers is used and the information from the Abraham database, catalog of Belgian newspapers.

3.2 Packaging

All carriers that will be digitized are packaged at and by the content provider. Both loose newspapers and bound newspapers (with and without covers) are packed in a collection box. A collection box contains at least one carrier. The collection boxes are given a barcode so that meemoo (and also the content partners) know where the carriers are at all times.

These collection boxes with loose newspapers and the newspaper bundles are transported in a transport container, supplied by the service provider. The collection boxes and newspaper bundles are placed in the transport containers just before transport by the service provider under the supervision of the content partner.

3.3 Transport of the carriers to the SP

The various transports and their content are planned by meemoo in collaboration with the content partners and the service provider, so that the content partners are always aware of where their carriers are and when they can be expected back. Before the packaged carriers are loaded into the transport vehicle of the service provider, they are protected against damage during transport with further packaging using material provided by the service provider. When the carriers leave the content provider's repository, the content provider adjusts their status in the database for registration and logistics accordingly.

3.4 Digitization

When the carriers arrive at the service provider's facilities, the service provider adjusts their status in meemoo's registration and logistics database accordingly. The service provider digitizes the material according to the agreed standards and formats as stated in the tender, and also creates secondary files that are necessary for logistical follow-up, long-term storage, quality control and accessibility, such as XMLs and checksums. He hereby gains access to the data recorded on the carriers during registration with the content partner, while identification is possible through the barcoding. Between the transport and the digitization as well as between the digitization and the return transport, the carriers are stored at the service provider in circumstances that do not harm the carriers any more than this would have been in case of the carriers remaining at the content partners themselves.

3.5 Transport of the carriers: way back

EFRO EUROPEES FONDS VOOR REGIONALE ONTWIKKELING

meemoo vzw | Ham 175 | B-9000 Gent | BTW BE 0644.450.380 | RPR Gent, afdeling Gent
www.meemoo.be | info@meemoo.be | T: +32 9 298 05 01

The service provider returns the carriers in the same packaging to the place where they were collected. When the carriers leave the service provider's facilities, the service provider adjusts status in meemoo's registration and logistics database accordingly. The various transports are planned by meemoo in collaboration with the content partners and the service provider, so that the content partners are always aware of where their carriers are and when they can be expected back. When the carriers arrive at the content partner's repository, the content partner adjusts their status in the meemoo registration and logistics database accordingly.

3.6 Transport of the files to meemoo

The service provider delivers all files to meemoo according to the agreed procedure and timing (both the image files and the secondary files). Meemoo initially checks all these files for their presence. The service provider also stores all files on its own servers as a backup for six months. After these six months, the service provider will destroy all backup copies of the files. In addition to the secondary files, the service provider also sends meemoo a general report on the progress of the digitization of that quantity for each quantity of transported carriers. This report is entered by meemoo in the database for registration and logistics, so that the content providers receive an initial feedback on the progress of the process in good time.

The creation of this report falls within the scope of this specification. The service provider, in consultation with meemoo, agrees on a format in which the data will be delivered. Typically these are XMLs.

3.7 Ingest and quality control of the files

Meemoo transfers the files to its own storage infrastructure and makes them accessible through its Media Asset Management (MAM) interface. The quality control procedures (such as file integrity) are performed using the XMLs and checksums created by the service provider. From this point on, the content partners can also download a copy of their own files and store these themselves, if desired. If meemoo detects errors in the digitization during the internal quality control (which cannot be attributed to the condition of the carrier itself), the carrier is returned to the service provider. If errors are found during this process in the structure of the delivered file packages or the formation of the associated XMLs, the service provider is required to redeliver the complete file package. The authority to approve the files always lies with meemoo and is done in consultation with the content partners and the VEB.

3.8 Addition of metadata

If the content partners themselves have additional metadata, these can be added in the MAM of meemoo. Further metadata can also be created directly in the MAM of meemoo. For example, the Abraham database (a catalog of Belgian newspapers) can be used. This metadata will serve as the basis for further dissemination of the material via new or existing platforms of meemoo, the content partners or other organizations.

meemoo vzw | Ham 175 | B-9000 Gent | BTW BE 0644.450.380 | RPR Gent, afdeling Gent
www.meemoo.be | info@meemoo.be | T: +32 9 298 05 01

4 General and technical project management

4.1 Overall project management

4.1.1 Insurance of the carriers

The tenderer must indicate **as minimum requirement ME01**: that the carriers are insured during these processes at the service provider against loss and/or damage and how this is done. The tenderer must demonstrate an insurance value of at least €15 per edition (even if there are several editions in one volume), or a total insurance value of €130,000 per batch that is simultaneously present at the digitization location.

To clarify VD01, more details can be given about the way in which the carriers are insured against loss and/or damage during the transportation to and from the content partner and during the stay with the service provider and information in particular about the insurance policy. A copy of this insurance policy must be submitted to meemoo at the time of award in a sealed envelope.

meemoo vzw | Ham 175 | B-9000 Gent | BTW BE 0644.450.380 | RPR Gent, afdeling Gent
www.meemoo.be | info@meemoo.be | T: +32 9 298 05 01

4.1.2 Project management and process

The tenderer must thoroughly explain in his tender how he wishes to approach the general project management of the digitization project and what quality guarantees he gives in this regard. Since the material may be protected by ethical, commercial and copyright laws, the tenderer must also exercise particular vigilance in this regard.

The applicant should indicate:

- **as minimum requirement ME02**: that meemoo can visit the digitization infrastructure during working hours, without prior notice.
- **as minimum requirement ME03**: that he has a procedure for urgent requests ('urgent request procedure'), whereby a digitized file of the requested carrier can be delivered via an FTP connection within two working days. Meemoo emphasizes that this procedure will only be invoked in exceptional cases, with a maximum average of 1 time per month, measured over the entire project duration.
- **as minimum requirement ME04:** that a one-to-one structure is set up for consultation between the project leaders on both sides. This means that a project manager is appointed on the part of meemoo and the applicant, through whom all communication takes place. Meemoo prefers as few changes between project managers as possible. In case of changes, meemoo must be informed at least two weeks in advance.
- **as minimum requirement ME05**: That the project manager at the service provider provides proof of at least three years of experience in project management via strict project methodology, by means of an attached curriculum. **To clarify VD05**, specifications can be given in the answer about how the communication between the project managers will be structured, in particular about the proposed communication channels, consultation rhythm, fixed agenda items and such.
- **as minimum requirement ME06:** that a test phase is set up prior to the pilot phase in which a small amount of newspapers successfully goes through each individual step of the process before moving on to the next step. An accompanying, non-binding description of the test and pilot phase can be found in attachment. During the test phase, further agreements between meemoo and the service provider are recorded in a separate document (PAD). These agreements shall detail, but not exceed, the terms of this RFQ. The mutual approval of this document is a precondition for starting the production phase.
- **as minimum requirement ME07**: that a pilot phase is set up prior to the production phase of the digitization in which a small number of test carriers successfully complete the process from start to finish.
- **as minimum requirement ME08**: that he ensures constant monitoring and professional processing of all carriers, with respect for their cultural-historical value. **To clarify VD08**, he can explain how he will do this, and with which (software) systems. Meemoo expresses its preference for a system that is automated and specially developed for this purpose, and where error sensitivity is reduced as much as possible. Monitoring through a simple spreadsheet system carries too many risks of human error and is therefore not allowed.
- **as minimum requirement ME09**: that he will prevent the commercial rights, ethical rights (privacy) and/or copyrights of the content from being violated by allowing the material to be digitized to become public through the digitization process or related processes on his side. **To clarify VD09**, he can indicate how he will prevent this.

4.1.3 Subcontracting

Meemoo accepts that part of the service may be subcontracted, subject to these **minimum requirements**:

- **as minimum requirement ME10**: that the applicant states with the greatest transparency which part of the contract the applicant intends to subcontract to third parties, including the identity and details of these subcontractors.
- **as minimum requirement ME11:** that the main contractor remains solely responsible for guaranteeing the proper execution of the contract and guarantees that the subcontracting does not create any additional obstacles in project management, logistics, quality control or in any aspect of the general project approach.
- **as minimum requirement ME12:** that the main contractor always remains the responsible contact person in the communication with meemoo.

For further provisions regarding recourse to subcontracting for the execution of the contract, reference is made to the Overview document.

4.2 Logistics

4.2.1 Transport and storage of the carriers

The content partners will prepare the transport by packing the carriers either in collection boxes or wrappers and then placing them in the transport box. If necessary, meemoo will replace or reinforce the original boxes as present at the content partners. In this way, all boxes should be able to withstand the transport and normal digitization operations at the tenderer. If the box nevertheless suffers further damage and needs to be replaced, the tenderer must contact meemoo, with the goal of recovering the box itself and the metadata as stated on the original packaging.

4.2.1.1 Externa transport by the service provider

The service provider collects the carriers from the content partners involved. For transport, the packaged newspapers are placed in protective containers with a barcode. The delivery of the transport containers by the service provider is determined during the test phase.

Meemoo plans the transports in collaboration with the content partners and service provider(s), to ensure that the content partners know where the carriers are at all times and when they are expected to return. Meemoo plans the date at least two weeks in advance in consultation with the content partners and the service provider. The service provider is expected to contact the content partner at least two days before the pick-up to agree the exact time and place of pick-up. The service provider, under the supervision and possibly with cooperation of the content partner, places the material to be transported in the transport boxes supplied by the service provider at a time to be agreed.

After digitization, the carriers are returned to the original location where they were collected.

The applicant should indicate:

- **as minimum requirement ME13**: that the transport is organized in such a way that the integrity of the carriers is guaranteed. Transport must take place in moisture- and shock-resistant

containers according to the tenderer's proposal. In order **to clarify VD13,** the applicant must explain how and under which modified circumstances he will transport the carriers to the digitization infrastructure and back to where they were collected. In particular, he must provide information on at least the following aspects:
    - Which vehicle is used;
    - How the vehicle is air-conditioned;
    - What further measures he takes to secure the carriers.
- **as minimum requirement ME14**: that the applicant collects the carriers and delivers them to the content partners involved, usually located in the Flemish central cities or Brussels. **To clarify VD14,** the applicant must indicate how he will handle the transports and must provide a non-binding design of a collection and return calendar.

4.2.1.2 Internal storage at the service provider

Between the transport and the digitization itself, and later between the digitization and the transport back, the carriers must be stored by the applicant in an air-conditioned and secure location.

The applicant should indicate:

- **as minimum requirement ME15**: that the handling of the material is done with great care, avoiding any contact with moisture, large temperature fluctuations and other harmful factors. The storage must take place in an air-conditioned room protected by an alarm. **To clarify VD15:** the applicant can indicate under which circumstances the storage takes place and how the climate and security in the storage space(s) are regulated. Meemoo expresses a preference for the following climate control:
    - 17-20 °C, with fluctuations of max. 2°C over 24 hours.
    - RH of 45-55% with fluctuations of 3% over 24 hours.
    - Maximum exposure to 50 lux when the carriers are unpacked.
- **as minimum requirement ME16:** that the internal logistics flow and timing is transparent. **To clarify VD16,** the applicant can add a clear diagram of:
    - the different locations where the carriers are located and the phases that the carriers go through during the entire process;
    - the maximum time it takes for one carrier to complete the entire digitization workflow from collection to delivery back to the collection location, according to the tenderer's schedule; in other words: how long the content provider must give up its carrier for digitization.

4.2.2 Identifying and arranging the carriers

Meemoo will provide each newspaper issue with a barcode. This is not applied to the carrier itself, but:

- in the case of bound editions: on at least one flyleaf at the front of the bundle, with all editions listed with date and, if applicable, edition number and the corresponding barcode.
- in the case of individual (loose) editions: on the packaging of each individual edition.

The service provider will be able to identify the carrier via this barcode. Meemoo will also provide the service provider with a list (extracted from the registration database) of all editions and a number of features that may be relevant for digitization. Where possible, the characteristics in this list will provide an indication on the level of editions about support technical aspects:

- Administrative: e.g. name of the content partner, PID, barcode of the carrier, barcode of the box, batch ID, ...
- Form: eg format, carrier unit (loose or bound, can be opened 180° or not), condition of paper and binding.

EFRO EUROPEES FONDS VOOR REGIONALE ONTWIKKELING

meemoo vzw | Ham 175 | B-9000 Gent | BTW BE 0644.450.380 | RPR Gent, afdeling Gent
www.meemoo.be | info@meemoo.be | T: +32 9 298 05 01

- Content: eg publication date as an indication of the age of the carrier.

This data list can be updated until the transport of the carriers to the service provider takes place. Meemoo supplies one METS XML file per edition to the service provider. The naming of the METS XML is based on the PID (unique identifier) of the carrier; the service provider must adopt and retain that naming (see chapter 4.4).

Meemo expects:

- **as minimum requirement ME17**: that, subject to exceptions agreed in writing, the order of the batches and the internal order of the carriers within their boxes as delivered before the digitization (ascending by number) is respected when they return after the digitization.
- **as minimum requirement ME18**: that the applicant can work according to the workflow mentioned above (METS XML, identification by means of barcode, naming, etc.).

4.2.3 Delivery of the files

After the digitization, the files must be delivered on LTO6 to meemoo's data center, located in Oostkamp, Belgium. Meemoo will ingest the files in its own storage infrastructure and store them sustainably via the Media Asset Management system via its management system.

The candidate should indicate **as a minimum requirement ME19**: that the applicant can supply the digital files on LTO6 tapes, in LTFS format. All files must be written directly to the root directory according to the attached SIP specification. **To clarify VD19**, he can indicate which schedule he will adhere to for the delivery of the files, by adding a concept calendar to his offer based on the speed of delivery of the files on the LTO tapes. The required quantities of LTO tapes must be collected by the service provider at meemoo's data center before the start of the pilot phase.

4.2.4 Temporary safety copy of the files

For security reasons, meemoo asks the service provider to keep a backup copy of all files for a certain period of time, starting from the delivery of the files to meemoo.

After digitization, the applicant must:

- **as a minimum requirement ME20:** to store a security copy of all files for another six months, counting from the delivery of the files to meemoo. After these six months, the tenderer will destroy all backup copies of the files.
- **as a minimum requirement ME21:** that this security copy can be delivered within four working days after request from meemoo, without costs for meemoo, on the same carrier and in the same format as the delivery of the original files. Any necessary changes to the metadata cannot give rise to additional costs for meemoo.

4.3 Digitization

In principle, meemoo requires the digitization of the entire carrier, including any blank pages, but excluding - in the case of bundles - the front and back covers. Memoo requires text recognition via OCR from each of the digitized pages. For digitization, each carrier must be manipulated correctly.

EFRO EUROPEES FONDS VOOR REGIONALE ONTWIKKELING
meemoo vzw | Ham 175 | B-9000 Gent | BTW BE 0644.450.380 | RPR Gent, afdeling Gent
www.meemoo.be | info@meemoo.be | T: +32 9 298 05 01

4.3.1 Manipulation, digitization, equipment and software

With regard to manipulation and digitization, meemoo asks:

- **as a minimum requirement ME22:** that the digitization takes place without damage to the original and without loss of information.
- **as a minimum requirement ME23**: that the tenderer further applies the Metamorfoze Light 1.0 guidelines, with eciRGBv2 as the color profile. These are included in Appendix 6 to these documents. During the test phase, further agreements will be made between meemoo and the service provider about a possible update of the test targets provided in Metamorfoze 1.0 and the quality control software to be used. **For clarification, VD23** asks meemoo whether the tenderer can also comply with ISO/TR 19263-1:201.7.
- **as a minimum requirement ME24:** that no digital processing takes place after the digitization for the production of the master files, unless with explicit permission from meemoo.
- **as a minimum requirement ME25:** that the bindings are under no circumstances released of bound copies. A maximum of one page may be captured per recording (unifolio). It should be taken into account that some newspapers cannot be unfolded 180°. Data on this will be provided between the end of registration of the newspapers concerned and the start of digitization.
- **as a minimum requirement ME26:** insofar as not contrary to the requirement stated in ME22: that the digitized page can be seen straight, completely and optimally legible on the image.
- **as a minimum requirement ME27:** insofar as not contrary to the requirement stated in ME22: that the background is black, unless for thin, translucent paper, then a white or light background should be used.

With regard to the equipment and software used, meemoo asks:

- **As a minimum requirement ME28**: that the digitization must be done with equipment and software that can achieve the intended quality standards and that the calibration of the equipment for this is carried out before the start of each shift. The tenderer must hereby indicate, for the purpose of **clarification VD28:**
  - what equipment and software he uses to capture the newspapers.
  - which equipment and software he uses to process the files.
  - how he foresees the calibration of the digitization chain before the start of the digitization.
- **as a minimum requirement ME29:** that he only uses a book scanner with a book rocker, and in no case a flatbed or transit scanner.

The tenderer may indicate as a **free option VO01:** whether he proposes additional options to optimize the quality of the digitization and the searchability, and which ones.

4.3.2 File formats and codecs

Meemoo asks **as a minimum requirement ME30:**

- for the creation of an archive master file: the file format and codec to be used is Uncompressed Baseline IBM TIFF v6.0, at 300 ppi. See Appendix 7 for the TIFF specifications.

4.3.3 Digitization flow of regional editions

The project includes titles with regional editions. These are newspapers that appeared in different regions, with a national section on the one hand - containing the same content for all regions - and on the other a regional section, with different content per region.

These titles are tied together in several ways:

- some bundles contain only the regional pages
- other bundles contain both the national and regional pages for each edition, so the national pages are always the same.

Meemoo has not yet determined its own preference with regard to the digitization, archiving and dissemination of these types of titles, but will let this partly depend on the answers received to the following mandatory options:

- **Mandatory option OP01:** everything of two regional editions is digitized (national pages and regional pages). Of the rest, only the regional pages are digitized. In total, approximately 72,000 pages are digitized (approx. 22,000 national pages + approximately 50,000 regional pages). During registration, it will be clearly indicated which editions it concerns and which pages should be skipped. Tenderers are asked to clarify in their tender how they will tackle this situation during digitization.
- **Mandatory option OP02:** all existing pages (regional and national) are digitized (in total approx. 100,000 pages to be digitized = approx. 50,000 national pages + approx. 50,000 local pages). This means that the content of approximately 30,000 pages is digitized twice.

Meemoo will weigh up the proposals and quoted prices of the applicants for OP01 and OP02 and make a choice for one of the two. The internal consequences for the archiving and dissemination of the files will also be taken into account.

4.4 OCR

With the goal of further processing of the files and their dissemination, meemoo also wishes, **as minimum requirement ME31**, a conversion to machine-readable text of the printed and typed texts. The tenderer must also indicate **for clarification VD31**:

- what equipment and software he uses for OCR text recognition.
- what maximum percentage of errors in the text recognition he can guarantee (expressed in percentages CER and WER per article, which can be checked by meemoo on a randomly composed sample of 10 articles), taking into account that these are historical newspapers from the nineteenth and twentieth century.
- which (semi-)automatic OCR and/or image correction it provides, i.e. which techniques are used to optimize the OCR.
- whether the OCR system used, allows the recognition of separate articles with the following metadata: title, date, edition and page of the carrier, as well as the beginning and end of the article, possibly also the length of the article.

In addition, the following minimum requirements apply:

- **as a minimum requirement ME32**: that the METS standard is used for drawing up the XML with the metadata.
- **as a minimum requirement ME33:** that the OCR files are supplied in XML format according to the ALTO standard. For clarification purposes, the tenderer must indicate **as clarification VD33** whether the dimensions as stated in the ALTO XML can be expressed in centimetres, pixels or both, whereby meemoo expresses a preference for pixels.

The tenderer may indicate as **a free option VO02:** whether he provides for a manual correction of the titles of the articles in the result of the OCR and what maximum margin of error he guarantees in this regard.

4.5 Quality control

With the goal of the further processing of the files and the dissemination, meemoo will perform a quality check on the technical characteristics and the file integrity of the supplied files.

To enable file integrity checks and proper XML creation, meemoo asks:

- **as minimum requirement ME34:** that UTF-8 encoding is used for all XML files.
- **as minimum requirement ME35**: that an MD5 checksum is generated for the tiff and ALTO files, which is included in the reporting XML per edition. For more details, please refer to section 4.7. The applicant must also indicate **as clarification VD35** which quality control mechanisms it applies internally to the digitized files and the XML schemas.
- **as minimum requirement ME36:** that the date, time and camera (brand, type and serial number) are included in the embedded metadata of the essence files and the recordings of the test targets.

The following procedure will be used for the technical quality control of the recording of the files themselves:

- At the start of each part of the day, before the start of the recordings, the service provider makes recordings of targets. The measurements are applied to these targets with a quality control software (see ME22) in accordance with the guidelines of Metamorfoze Light 1.0.
- The recorded targets and the measurements are directly forwarded to meemoo.
- Meemoo repeats the measurements on the forwarded targets using the same quality control software as the service provider and validates whether any deviations fall within the agreed margins. If not, the service provider will be notified and the digitization infrastructure will have to be reset.
- The same procedure as described above is then followed until the required values are reached and the actual production can start for that part of the day. Any recordings that had already been made with the infrastructure for which the targets and measurements for that part of the day had not been approved by meemoo, must be retaken at the expense of the service provider.

Meemoo asks **as minimum requirement ME37:** that he can work according to the above procedure.

4.6 Secondary files

In addition to the digitized files, meemoo also requires machine-readable files that provide insight into the progress of the digitization and OCR files to make the contents of the newspapers searchable.

Meemoo wants clear and detailed information about the logistics and digitization to be kept at edition level, among other things to give the content partners insight into this, also with regard to carriers that are not or only partially digitized.

This information is collected as follows:

- Before the transport of a batch, meemoo supplies an XML file from the registration database to the service provider. This contains data about the editions themselves, created during registration. There are also empty fields that must be filled in by the service provider during the digitization.
- The service provider uploads the data from this file to its own workflow system.
- During digitization, the service provider adds data about the process steps and the results of logistics and digitization. This is done in the fields and terminology suggested by meemoo. The PREMIS standard is used as a guideline here.

If the digitization is only partially successful, the essence and the METS XMLs must be delivered. Even in the event that the attempt at digitization essentially did not yield essence files, this attempt must be reported in the METS XML.

As a leading, **non-binding** example for these documents, the file 'METS XML after digitization' is attached with:

- Fields filled with sample data from the registration database.
- Fields filled with reporting data created by the service provider.

The implementation of these export files must be taken into account in the quotation. Meemoo points to a landmark development time of 10 work days to enable and fine-tune this XML-based approach.

4.7 SIP composition

Meemoo requires **as a minimum requirement ME38** that the requested files per newspaper edition are supplied in a so-called SIP, which is structured as follows:

Complex	Location	Filename	Info
Pid.complex			folder = SIP container
	root	pid_mets.xml	METS file (incl. metadata o. the SIP)
directory	/alto		Folder with output of the OCR
		pid_0001_alto.xml	XML with output of the OCR per page
		pid_0002_alto.xml	
		pid_0003_alto.xml	
		...	
directory	/tiff		Folder with output of the digitization per page - masters

		pid_0001_tif.tif	
		pid_0002_tif.tif	
		pid_0003_tif.tif	
		…	

The applicant must take into account the implementation of this SIP composition in its offer. The exact contents of the expected files are discussed below:

- pid.complex: root folder of the SIP, named after the PID, with extension .complex, packaged as a zip file.
- pid.xml: METS XML with metadata from the registration process, supplemented by the service provider with metadata from the digitization process, the latter being structured according to the PREMIS standard. As a non-binding example, a pid.xml has been added to this file to provide insight into the structure of the supplied metadata fields from the registration and to further supplement metadata fields that document the digitization by the service provider.
- pid_0001_alto.xml: xml containing the results of the OCR process, structured according to the ALTO standard.
- tiff: folder in which the archive master files are collected per page in tiff format.
- pid_0001_tif.tif: Archive master file per page in tiff format, where the sequence '0001' is filled in according to the order of the pages.

The pid.xml files (according to METS standard) must be supplied in two ways:

- Within the SIP, on the LTO, in the meemoo data center, unless digitization has yielded no substantive results.
- by e-mail to astrid.theerens@meemoo.be: 1 zip file per supplied batch containing 1 pid.xml per newspaper edition (incl. the one for which the digitization has not produced any essence files) for an update of the meemoo registration database (see 4.8 Reporting).

4.8 Reporting

Meemoo wishes to stay informed about the progress of digitization and also to provide the content partners with a transparent insight into this. **As a minimum requirement ME39, therefore**, the applicant must set up a reporting routine, e.g. via an online accessible spreadsheet system, in which the progress of the digitization is reported every two weeks. At a minimum, the applicant will be asked for the following:

- The number of digitized carriers (and the delta since the previous report)
- Number of non-digitizable carriers (and the delta since the previous report)

4.9 Timeline of the digitization project

As an indication, the planned timing that meemoo wishes to adhere to for digitization is given below:

- March 25, 2022: Announcement of the award decision.
- April 9, 2022: End of the legal standstill – effective award.
- Second week of April 2022: First consultation with the service provider.
- Last week of April 2022: Start of the test phase.
- First week of June 2022: Start of the pilot phase.
- First week of July 2022: Start of the production phase.
- End of July 2023: Expected end of the production phase.
- September - November 2023: Planned production phase for a possible additional digitization Nieuwe Tijdingen- see also mandatory option OP03[1]

The tenderer must indicate: **as minimum requirement ME40**: that he can carry out the assignment within the specified periods, subject to manpower and delays at the expense of meemoo.

Meemoo points out that the registration of the newspapers must be done with the content partners and that this is only possible if the corona measures allow it. While the registration process will be set up in such a way that it will have a certain head start of the transport data, delays in the registration process due to the corona measures will be considered a case of force majeure. These delays might in turn may delay the supply of newspapers for the digitization process,

The tenderer must indicate as **mandatory OP03 option**: the additional costs for the digitization of 50,000 additional pages in the context of the Nieuwe Tijdingen project, according to the same specifications and project layout as the Primeur newspapers.

[1] Timing to be determined in consultation, but in any case after completion of the contract in the context of the Primeur project and before the end of 2023.

meemoo vzw | Ham 175 | B-9000 Gent | BTW BE 0644.450.380 | RPR Gent, afdeling Gent
www.meemoo.be | info@meemoo.be | T: +32 9 298 05 01