

# **Transfer requirements for digital archives**

**Version 1.2  
16-03-2022**

**Regionaal Archief Nijmegen**

## Foreword

---

For a correct ingest of digital archives in our digital repository a number of technical requirements must be met. These requirements concern metadata, digital objects and the way of delivery. A digital archive that meets the requirements can be transferred to our digital repository. The requirements serve as a guideline for depositors when preparing for a transfer, but also as a guideline for organizing their digital information management in a sustainable way.

Besides the technical requirements for the ingest of digital archives, there are also general requirements for archives in case of an official transfer. These are independent of the (digital) form of the archives and concern, for example, the quality of the descriptions of the content, appraisal and selection and public accessibility. Both the general and technical requirements are used to ensure that the archives are in good, orderly and accessible condition.

The requirements are not static and are managed on the basis of the development of (inter)national standards and new insights. Regionaal Archief Nijmegen aligns its transfer requirements as much as possible with (inter)national standards, guidelines and best practices. The developments in the field of sustainable digital information management are rapid and the development of standards often follows at a distance. For example the metadata standard intended for Dutch government organizations is currently in further development. This means that the requirements are likely subject to (minor) changes as developments and insights advance.

### Transition phase

There is a high probability that metadata and digital object are not (yet) in line with all of the technical requirements. After all, in many cases these requirements have only become known to the depositors after the creation of their digital objects. Therefore a transition phase will be necessary for the digital objects that were created in an earlier stage. We expect that the depositors use this transition phase to make sure that the current requirements will be implemented within their own local (archival) systems.

The requirements listed in the left column below serve as a starting point for the current and new processes and systems in which digital archives are created and managed. The right column contains requirements that apply to older digital archives when the relevant requirement in the left column cannot reasonably be met. If there are no requirements in the right column, the requirements in the left column apply without question.

Prior to each official transfer of digital archives, Regionaal Archief Nijmegen will discuss with the depositor whether the requirements are met or not. An intake document will be used for this purpose.

General	Requirements	Transition phase
	The information package <u>only</u> contains files for permanent preservation. Files that should be destroyed must have been removed from the information package in accordance with the selection list before the transfer.	
	The information package is properly arranged.	
	The information package <u>only</u> contains closed dossiers (a set of files).	
	The information package <u>only</u> contains complete dossiers (all relevant information from the source systems of the depositor is included and relations with other systems are included in the metadata when applicable).	
	It is clear which digital objects are part of the transfer, in which source systems these objects are being managed and which file formats are present.	
	The information package <u>only</u> contains digital objects with unambiguous descriptions that correspond with the actual content.	
Metadata	Requirements	Transition phase
	The metadata is based on a documented metadata schema that complies with TMLO or MDTO.	The metadata is based on a documented metadata schema and when applicable supplemented with process specific metadata. The documented metadata schema must include a mapping to TMLO or MDTO. The metadata must include at least context information, integrity information, provenance information, representation information, and access information (source: ED3, Archiefregeling; Articles 17 and 24, TMLO, MDTO).
	The metadata is delivered in the ToPX format (preferably the latest version, ToPX 2.3) or the MDTO-XML format based on the sidecar structure. <sup>1</sup>	The metadata is delivered in a CSV, XLS, XLSX or XML file format.
	The MD5 or SHA-256 checksum algorithm is used to fill the TMLO metadata fields 21.7.1 (integrity algorithm) and 21.7.2 (integrity value).	

<sup>1</sup>For more information on ToPX and the sidecarstructure see: <https://www.nationaalarchief.nl/archiveren/kennisbank/voorwaarden-export-naar-e-depot>

	The information package contains <u>no</u> metadata without a related digital object.	
	The information package <u>only</u> contains digital objects for which the conditions surrounding the processing of personal data are known and described.	
	If a digital signature is used, at least the following is recorded: the holder of the digital signature, the moment of validation and the result thereof, the agent responsible for validation, and the identification of the certificate (source: Archiefregeling; Article 24).	
<b>Digital objects</b>	<b>Requirements</b>	<b>Transition phase</b>
	The information package is free of viruses.	
	The information package contains <u>no</u> corrupt files (all digital objects are openable, readable and perceivable).	
	The information package contains <u>no</u> zero byte files (no digital objects with a file size of 0 bytes).	
	The information package <u>only</u> contains digital objects with an intact, unaltered bitstream.	
	The information package contains <u>no</u> digital objects without related metadata.	
	The information package contains <u>no</u> encrypted digital objects.	If the information package contains encrypted digital objects then the associated decryption key is provided no later than at the time of transfer (source: Archiefregeling; Article 26).
	Use of compression technology is only permitted in so far as it does not result in such a loss of information that the requirements for the orderly and accessible condition of the digital objects can no longer be met (source: Archiefregeling; Article 26).	
	The information package <u>only</u> contains file names and directory names with the following characters: (ASCII) letters, (ASCII) numbers, underscores (_) or hyphens (-). Punctuation marks (? " ! ' ( ) ; : , ), special characters (\ / * < >   % & = # { } \$ @ +) and diacritical mark (é ç è à ù ë ö ï) are not allowed.	

	The information package <u>only</u> contains valid and standardized file formats (open, well-documented, stable and software- and manufacturer-independent).	The information package <u>only</u> contains file formats that are listed in our overview of preferred and acceptable file formats (see attachment).
	The digital objects are <u>not</u> compressed in a container format like RAR, TAR or ZIP.	
	The digital objects do not contain technical links with other digital objects that are not transferred.	
<b>Delivery</b>	<b>Requirements</b>	<b>Transition phase</b>
	The information package is delivered through a secure FTP server or another secure file transfer service.	
	The information package is delivered including a current, complete and logically consistent overview of the digital archive.	
	The information package is delivered including a completed intake document.	

## Attachment: file formats

The table below shows for each file type which file formats are preferred when it comes to sustainable accessibility and which file formats are acceptable. Acceptable means: there are better file formats, but this file format is not in danger of becoming obsolete and its readability can be guaranteed for the foreseeable future.

File formats		
Filetype	Preference	Acceptable
<i>Images (raster)</i>	<ul style="list-style-type: none"> <li>• Portable Network Graphics (.png)</li> <li>• Tagged Image File Format: Baseline TIFF (.tif/.tiff)</li> </ul>	<ul style="list-style-type: none"> <li>• Joint Photographic Experts Group (.jpg/.jpeg)</li> <li>• JPEG 2000 (.jp2)</li> </ul>
<i>Images (vector)</i>	<ul style="list-style-type: none"> <li>• Scalable Vector Graphics File (.svg)</li> </ul>	<ul style="list-style-type: none"> <li>• Adobe Illustrator (.ai)</li> <li>• Encapsulated PostScript File (.eps)</li> </ul>
<i>Audio</i>	<ul style="list-style-type: none"> <li>• Free Lossless Audio Codec File (.flac)</li> <li>• Waveform Audio File Format (.wav)</li> </ul>	<ul style="list-style-type: none"> <li>• Advanced Audio Coding (.aac/.m4a)</li> <li>• MPEG-1/MPEG-2 Audio Layer III (.mp3)</li> <li>• Windows Media Audio (.wma)</li> </ul>
<i>Databases</i>	<ul style="list-style-type: none"> <li>• Comma Separated Values (.csv)</li> <li>• Extensible Markup Language (.xml)</li> <li>• Open Document Database (.odb)</li> </ul>	<ul style="list-style-type: none"> <li>• Microsoft Access Database (.accdb/.mdb)</li> <li>• Software Independent Archiving of Relational Databases (.siard)</li> <li>• Structured Query Language (.sql)</li> </ul>
<i>E-mail</i>	<ul style="list-style-type: none"> <li>• E-mail Message (.eml)</li> <li>• Extensible Markup Language (.xml)</li> </ul>	<ul style="list-style-type: none"> <li>• Outlook Mail Message (.msg)</li> <li>• PDF/A-1 of PDF/A-2 (.pdf)</li> </ul>
<i>Presentation</i>	<ul style="list-style-type: none"> <li>• Open Document Presentation (.odp)</li> <li>• PDF/A-1 of PDF/A-2 (.pdf)</li> </ul>	<ul style="list-style-type: none"> <li>• Microsoft Powerpoint (.ppt/.pptx)</li> </ul>
<i>Spreadsheets</i>	<ul style="list-style-type: none"> <li>• Comma Separated Values (.csv)</li> <li>• Open Document Spreadsheet (.ods)</li> </ul>	<ul style="list-style-type: none"> <li>• Microsoft Excel (.xls/.xlsx)</li> <li>• PDF/A-1 of PDF/A-2 (.pdf)</li> </ul>
<i>Text</i>	<ul style="list-style-type: none"> <li>• PDF/A-1 of PDF/A-2 (.pdf)</li> <li>• Open Document Text (.odt)</li> <li>• Plain Text File (.txt)</li> </ul>	<ul style="list-style-type: none"> <li>• Microsoft Word (.doc/.docx)</li> <li>• Portable Document Format (.pdf)</li> <li>• Rich Text Format (.rtf)</li> </ul>
<i>Video</i>	<ul style="list-style-type: none"> <li>• Audio Video Interleave (.avi)</li> <li>• Matroska Video (.mkv)</li> <li>• Material Exchange Format (.mxf)</li> </ul>	<ul style="list-style-type: none"> <li>• Apple Quicktime Movie (.mov)</li> <li>• Moving Picture Experts Group (.mpeg/.mp4)</li> </ul>
<i>Websites</i>	<ul style="list-style-type: none"> <li>• Web Archive (.warc)</li> </ul>	<ul style="list-style-type: none"> <li>• Archive (.arc)</li> <li>• Hypertext Markup Language (.html)</li> </ul>