



Preserve and enhance your existing press material

The European leader in asset digitisation, Numen provides standard digitisation services that perfectly comply with all technical and quality requirements asked by the market.

No time to lose

Once printed for short-term use and with a limited life span, newspapers are extremely fragile. The more they are handled by researchers and the general public, the more their condition deteriorates. Digitising them is therefore vital for their long term preservation.

A pressing opportunity

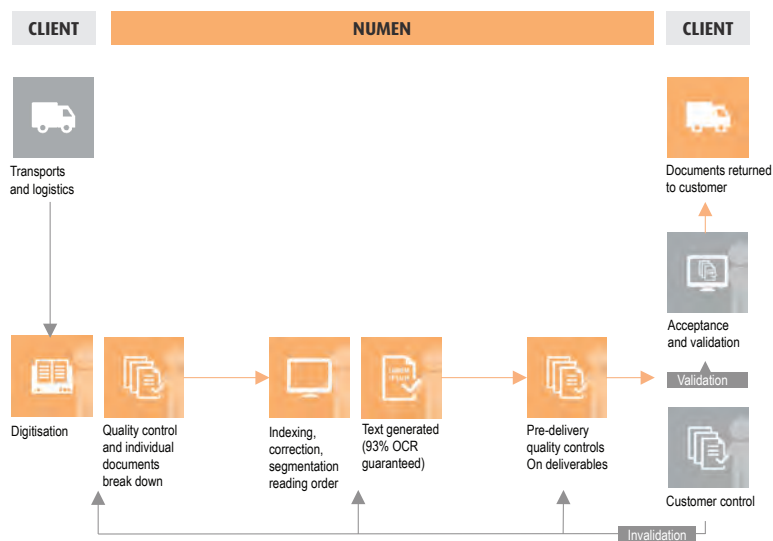
The value of press material can be significantly enhanced by converting it into high quality text and storing it online, widening opportunities for genealogical, scientific and economic information datamining or targeted queries.

A tailor-made package

Numen offers a straightforward, onestop package designed:

- › to process any type of press material, regardless of the format or binding method: (sheet paper, bound, flatbed scanning, reduced size scanning, from A5 up to A0)
- › to capture perfect images on large format scanners and output reliable metadata
- › to optimise the quality of the text generated by OCR, ensuring a baseline 93% word recognition rate
- › to deliver ready to use files matching core library standards: JPEG 2000 image files, metadata in METS and/or Refnum format, PDF file with hidden text layer

SYSTEM DIAGRAM



Scan by batches rather than requested by demand in order to get the best value for money.



How it works

NUMEN AVAILABLE OPTIONS

Image capture

- › Image files in different formats (TIFF, PNG, JPEG, low resolution, etc)
- › Spine and edges digitisation for valuable works

Indexing

- › Down to the article cataloguing
- › Human operators keyed-capture by article titles or key elements

Text conversion

- › High quality OCR (quality rate per word between 95% and 97%, possible higher)
- › Alto files supplied

Dissemination methods

- › Digital content dissemination and internet intranet hosting platform provision

Transport and logistics

Newspapers are transported and stored:

- › by experienced personnel
- › in secure conditions
- › covered by a high end "nail to nail" insurance policy

Image digitisation and processing

- › Wide array of scanning equipment for all situations (up to A0 format)
- › Images captured as 400 dpi uncompressed TIFF files (colour or greyscale)
- › Hand cropped images
- › Image numbering (sorting into individual documents)
- › Comprehensive quality control for each page
- › Documents returned in their original state
- › Image file delivered

Indexing

- › Metadata is entered for each document issue (page number, date of publication and number)
- › All metadata (bibliographies, production, techniques, etc.) is computer processed and checked
- › Deliverable: METS and/or Refnum (for OAI-PMH harvesting) metadata XML files

Text conversion

- › Manual corrected OCR segmentation to ensure that all text zones are captured
- › Reading order controls (excluding tables and adverts)
- › Automatic correction designed to achieve a minimum quality guarantee of 93% per word for texts in French (excluding tables, adverts and documents with poor legibility)
- › Deliverable: PDF with hidden text layer

Secure

Long-term preservation of information that would otherwise deteriorate in its current state

Your files are transported, stored and handled by one trusted service provider

Optimise

Your dematerialisation process with one single service provider for all types of press material

Access all your digital files with our fully library standard compliant services

The best quality-price ratio

Dynamise

Providing access to your newspapers has beneficial effects on both research and culture

Boost your economic value by disseminating your press material

