

YEWNO CONTENT INGESTION

Preferred Content and File Types

All content is best ingested with the accompanied metadata

We can accept all content types and standard formats including:

- text
- audio
- datasets
- images
- video

Preferred text content and metadata file formats:

- xml
- txt
- tex
- pdf
- json
- doc

Preferred metadata formats: ONIX for books and JATS for journals, other formats please contact us. If you have other content types available, please let us know. Content that is scanned and has not been OCR'd please reach out to us for ingestion options. Additional (Spanish, etc) language files are accepted, but will be ingested at a later date.

Historical data (archive/batch)

Preferred methods of delivery:

- FTP (ours or yours), s3 buckets, API (OAI-PMH or REST), or hard drive
- Folder structure should be consistent with bzip, zip, tar, or gzip compression

Questions and contact details

Questions about content ingestion? Please contact us.

Contact: Matthew Matey - matt@yewno.com