**EBSCO**

EBSCO Information Services

# EDS DATABASE SUBMISSION GUIDELINES

The information below outlines preferred data formats and delivery methods for sending electronic data to EBSCO Information Services in support of the EBSCO Discovery Service.  Following these guidelines will ensure EBSCO processes your content as quickly as possible and will give customers the best experience with your data on EBSCO's platform.  We are happy to review sample files in advance.

## CONTENTS

## ONBOARDING PROCEDURES

Once the contract is signed and production resources are ready to begin a new project, a member of EBSCO's Publisher Content Management team will reach out to a data partner to begin the onboarding process.  The publisher content management specialist will facilitate the following onboarding workflow:

- Contacts data partner to initiate onboarding process
- Coordinates initial technical discussion, either by setting up a conference call or via email
- Liaises any questions to and from development team and data partner
- Sets up FTP site for data delivery
- Facilitates the metadata review process with EBSCO's development team
- Remains data partner's main contact throughout the database design and development process

## DATA FORMATS

Delivering data in the formats described below, and with the accompanying information, will ensure customers can begin accessing your content on EBSCO's platform as quickly as possible. Providing EBSCO with the preferred metadata elements listed will support the search tools customers use to access your content.

### *Preferred Metadata Formats*

1. MARC or MARC XML
2. Proprietary XML with separately fielded core metadata elements (see list below)
3. JSON

It may be possible to accept other data formats such as MODS or Dublin Core. Please contact our Databases Publisher Content Management Group at EISPCMdatabases@ebsco.com with questions.

*Unacceptable Metadata Formats:*

- MS Excel
- MS Word

### *Preferred Core Metadata Elements*

- Unique, persistent record identifiers (accession numbers)
- Title
- Author(s)
- Publication name
- Publisher location
- Publication date
- Document type

For periodical content (articles), the additional required fields are:
- Journal title
- Issue
- Volume
- Start page
- End page

To enhance the user experience, these are some additional fields we can display:
- ISBN/ISSN
- URL
- Title variant
- Series title
- Abstract
- Subjects or keywords
- Title source
- Type source
- Non-Author contributors
- Peer-Reviewed status indicator
- Language of item
- DOI (Digital Object Identifier)

## *Further Metadata Specifications*

For EDS, all records should:
- Have data that is well-formed and valid to its schema
- Meet UFT-8 encoding standard
- Be included in files no larger than 2GB

In addition, please provide a data dictionary, DTD, or schema if available. Let us know if you have MathML or LaTeX formatting tags, as EBSCO can accept rich markup (including scientific markup) but we currently do not display it.

## *Preferred Full Text Formats*

1. XML
2. PDF

*Unacceptable Database Full Text Formats:*
- MS Word

### *Further Full Text Specifications*

- Specify how metadata and full text are connected.
- UTF-8 encoding is preferred.

## LINKING

One of the major benefits of EBSCO Discovery Service is that it is constantly driving new users to publisher's websites. EBSCO does this through excellent linking tools.

- Please tell us how customers can link from an individual record to the publisher website.
- Let us know if the linking is already contained in the content or if EBSCO needs to create links.
- Specify whether links will lead customers to citation information or full text.
- If user authentication is required, please tell us if you authenticate users of your service via IP address or login.
- Please provide us with a test account.

## SUBSCRIPTION MODELS

We can accept multiple subscription models and sub-collections.

- Please tell us if your dataset contains multiple collections.  If so, please tell us how to distinguish between them.
- Please tell us if your customers purchase access to individual titles or to pre-defined collections.

## DELIVERY METHODS

### *Preferred Database Delivery Methods*

- FTP
  - The fastest and simplest way for delivery to EBSCO
  - Easy (drag and drop)
  - Allows publisher to control delivery of content
  - Reliable
- OAI-PMH harvesting
  - Preferred if FTP is not an option
  - Standard protocol
  - Can be slow as harvesters may time out, need to be restarted, or be subject to restrictions

### *EBSCO-Provided FTP Site*

- Upon request we will provide you with a link to an FTP directory, username and password

### *Publisher's FTP Site*

- Please provide the address (URL), username, and password.
  - Let us know if and when these credentials will expire.

### *OAI-PMH Harvesting*

- Please provide the address (URL), username and password.
  - Let us know if and when these credentials will expire.
- We can work with different harvesting protocols.

- Information needed:
  - Please tell us if the first harvest will include a complete dataset.
  - Tell us if there is a limit to the number of records we can pull per second.
  - Tell us how to harvest full text. Please specify, if applicable:
    - A URL that is different from the metadata.
    - Separate credentials for the full text.
    - Different restrictions for harvesting the full text.
    - Type of files; XML or PDF preferred.

## RELOADS AND UPDATES

EBSCO supports and encourages regular updating of datasets. We work with publishers to establish an update process that is as automated as possible for both parties.

Initial Delivery

- Please specify the number of records to be delivered.
- Specify if this will be a complete dataset.

Updates and Reloads

- If you will be sending regular updates, please describe:
  - Update frequency.
  - How many records will be included in an update.
  - If updates will include delete records.
  - Field(s) that contain the update or delete indicator, and the update date.
- If, instead of updates, you will be sending complete reloads of the data, please describe their frequency.