

# Delivery of Elsevier Bibliographic Databases in XML

Version: 1.4

Created: 16 Sep 2007

Changed: 24 Sep 2007

Changed: 4 Sep 2008

## XML

In 2008 specific XML Schemas are available for all Elsevier Bibliographic Database products. The next version of the AnI DTD, version 5.12, is available both as an XML schema and as a DTD.

The XML schema will be the primary document, and the DTD will be derived from it. New Schema versions can be released two times per year. A new Schema release will be communicated three months in advance of the change.

The new Schema and sample files will be placed on the website well in advance of the actual implementation.

The XML files will make use of namespaces. Therefore customers' XML software needs to be namespace aware. This requirement is satisfied by almost all current XML software. The schema location mentioned in the XML files will be a symbolic location. Customers, who want to validate the XML files against the XML schemas, must have XML software which is able to map the symbolic schema location to the real schema location using entity resolving software. The most common such software uses an XML Catalog, see

<http://xml.apache.org/commons/components/resolver/index.html>. A sample Catalog will be delivered with the XML schemas.

## Delivery

By default, files are posted on the FTP server: [FTP.elsevier.nl](http://FTP.elsevier.nl), login name and password will be provided if not already available.

The standard data delivery method is via an account on our FTP server where files can be picked up. Files are delivered to the FTP server on working days (schedule available).

Data delivery will consist of three data streams and zipped files are posted in sub-directories with their own number:

1. Directory # for "new" records
2. Directory # for "update" records
3. Directory # for "delete" records

### 1. New records

New items are delivered in a zip file.

### 2. Update records

Records that have been revised should replace the record that already exists in the database.

If the records are not present in the database, the records should be ignored. Updated records can be delivered daily in zipped files.

### 3. Delete records

Records that have status "delete" should be removed from the receiver's system. Delete records can be delivered daily in zipped files.

If the records are not present in the database, the records should be ignored.

In practice the frequency with which 'delete' and 'update' files are delivered will vary from weekly to daily according to the number of records that have a revised status.

## File naming

Each delivery consists of two files: a \*.zip file with the actual data and a \*.lbi file with information about the size etc. of the \*.zip file.

Files names can be customized but the default file naming conventions is as follows (the example uses EMBASE):

**New files** EMBASE\_<customercode>\_new\_<year>\_<sequencenumber>.zip

**Update files** EMBASE\_<customercode>\_update\_<year>\_<sequencenumber>.zip

**Delete files** EMBASE\_<customercode>\_delete\_<year>\_<sequencenumber>.zip

## Notification Mail

Every delivery will be accompanied by a standard notification mail containing information on the filename, number of records etc. If necessary the format and the information in contained in the notification mail can be customized.