

OCLC Concise Batch Processing Guide



© 2002, 2005, 2008 OCLC, Inc.
6565 Kilgour Place
Dublin, OH 43017-3395 USA

The following OCLC product, service and business names are trademarks or service marks of OCLC, Inc.: OCLC, EDX, WorldCat and "The world's libraries. Connected".

In addition, the WorldCat symbol and OCLC symbol are service marks of OCLC. Third-party product and service names are trademarks or service marks of their respective owners.

OCLC grants permission to photocopy this publication as needed.

Table of Contents

1	About This Guide	4
2	What is Batch Processing?	5
	Maintain holdings in WorldCat	5
	OCLC-MARC output	5
3	Types of Batch Processing Projects	6
	General information	6
	Retrieving records and reports via OCLC Product Services Web	6
	Retrospective (retro) batchload	6
	Retrospective (retro) batchload followed by ongoing batchload	6
	Ongoing batchload	6
	Cancel holdings	7
	Batchload with output	7
	Local data retention	8
	Reclamation project	8
	Scan/delete processing	9
	Adding original records	10
	Group batchload	10
	PCC BIBCO participants and batchloading	11
	Batchloading of non-MARC records, or numeric search keys	12
	Local data record updating	13
4	How to Start a New Batch Processing Project	14
	Getting started	14
	Input media for Batch Processing	14
	Preparing files for electronic transmission via EDX	15
	Sample FTP session	16
	Setting expectations for new batchload projects	16
	Why do new batchload projects take up to 90 days to evaluate and process?	17
5	Unresolved Records	19

1 About This Guide

The *OCLC Concise Batch Processing Guide* is intended to provide high-level information about different types of Batch Processing projects and how they may be initiated.

For detailed information about OCLC® local data record updating, refer to the [Local Holdings Updating Service User Guide](#).

2 What is Batch Processing?

Maintain holdings in WorldCat®

Batch processing is a way for your institution to maintain your holdings and add records offline. You can send files of MARC or non-MARC bibliographic records or numeric search keys to OCLC, to set and cancel holdings. Bibliographic records may also be added to WorldCat (the OCLC Online Union Catalog). Local data records (LDRs) may be added and updated through batch processing.

No matter how incomplete your institution's cataloging is, files can be prepared for batch processing by OCLC if some basic access points are available. These basic access points include:

- Library of Congress Control Number (LCCN)
- Title
- Personal name-title
- International Standard Book Number (ISBN)
- International Standard Serial Number (ISSN)
- OCLC record control number

OCLC-MARC output

You may also choose to receive a copy of the matching OCLC bibliographic records found during Batch Processing.

3 Types of Batch Processing Projects

General information OCLC’s Batch Services staff assigns a unique project ID for all batchload projects. This unique project ID is used for tracking project file submittals and processing them through completion.

Records are billed under differing product codes, depending on the type of project and type of records being submitted for processing and record output. An evaluation fee for new projects (TAP 3921 or TAP 3845) may be charged to cover record analysis and setup.

Batch post-processing reports are available on [OCLC Product Services Web \(PSWeb\)](#) after every batch of records is processed. OCLC-MARC records from Batch Processing projects are also available on [OCLC Product Services Web \(PSWeb\)](#).

Retrieving records and reports via OCLC Product Services Web

To retrieve files using [OCLC Product Services Web \(PSWeb\)](#):

	Action
1	Click download records and reports .
2	Click OCLC Batch Services records and reports .
3	Type your OCLC authorization and password.
4	Click Enter .
5	Select the desired file to download.
6	Perform a Save As, as per your browser’s help instructions.

Retrospective (retro) batchload

Your institution can submit all of the records in your local system for Batch Processing. This is normally referred to as *retro load*. All records read are billed under one product code, TAP 3981 for single institutions, or TAP 3983 for groups.

After the retro load is complete, some institutions maintain their holdings online locally. Therefore this type of Batch Processing project is referred to as *one-time retro*. In this case, the institution does no further batchloading.

Retrospective (retro) batchload followed by ongoing batchload

You may also choose to continue with batchload on an ongoing basis. This would be referred to as *retro followed by ongoing*. For tracking purposes, a separate definition is created for both the retro and the ongoing. When going directly from retro batchload to ongoing batchload, there are no additional evaluation fees since the only differences between the definitions are project IDs and product codes. The setup remains the same.

Ongoing batchload

You can submit files to OCLC for batchloading on an ongoing basis. When a matching record is found in WorldCat, holdings are set or canceled according to the Record Status of each record read. The Record Status code (Leader byte 5) must be coded appropriately so that OCLC’s batchload software will know when to cancel or set holdings. Since all records read are billed during pre-processing, any holdings canceled receive a credit.

If your institution cannot identify “cancel holding” transactions by the record status code, these records must be submitted for batch processing separately. (See [Cancel holdings](#).)

Product codes for ongoing projects vary according to the type of records submitted:

- A project consisting of only OCLC-derived records is billed under product code TAP 3812 for single institutions or TAP 3814 for group.
- A project consisting of only non OCLC-derived records is billed under TAP 3911 for single institutions or TAP 3931 for group.

Commonly, ongoing batchload projects consist of both OCLC-derived and non-OCLC-derived records. These projects are referred to as *mixed-batch*. During batchload pre-processing, files are split and billed according to the presence of an OCLC control number. Records that contain an OCLC control number are billed under the OCLC-derived product codes; records without an OCLC control number are billed under the non OCLC-derived product codes.

Cancel holdings

There are two ways to perform cancel holding or deletes:

- A file containing cancel holdings records can be submitted for batchloading. There is no evaluation fee or per record read charge for this type of project. This type of project can be set up as ongoing or one-time. Records are processed under TAP 3919 for single institutions, or TAP 3939 for groups. Since records are not billed during pre-processing, there is no credit issued.
- If MARC records cannot be provided, a file of OCLC control numbers can be submitted for numeric search key processing.

If your institution's goal is to remove ALL holdings in WorldCat, a *scan/delete* should be requested. (See [Scan/delete processing](#).)

Batchload with output

When your institution sets its holding symbol in WorldCat through batchloading, you can also elect to receive a copy of the matching OCLC MARC bibliographic record. In addition to the per-record read and evaluation fee billed in pre-processing, a per-record matched (record returned) charge is incurred.

The "output" product code varies depending on the institution or group, and if it is billable. Batchload output is billed at one of the following product codes:

CTR	3376	Batch Single Record Copy
CTR	3349	Batch Single Record Copy No Bill
CTR	3338	Batch Group Record Copy
CTR	3339	Batch Group Record Copy No Bill

Local data retention

It is important to think about local data retention when requesting output from a batchload project. For example, you may need to retain your system control number in the output record so you can overlay the old record in your local system with the new OCLC MARC bibliographic records. For single institution batchload projects, any local data can be retained. If the data is not already contained in one of the fields listed below, it will need to be mapped to and retained in one of them:

035 049 059 09x 590 599 690 691 699 852 910 945-949

When local data is retained, the entire fields are retained. You cannot request partial fields or selected subfields.

Although most commonly used when libraries request batchload output, there are other reasons for retaining local data. Any single institution can request local data retention. This will ensure that your local data appears in the OCLC archive record should you ever need to request copies of these records through the Bibliographic Record Snapshot service. There are limitations on local data retention for group batchload projects (see [Group batchload](#)).

You can also choose between the following options to determine how you would like to receive your file of matching OCLC MARC records:

- Pick up records via an EDX™ account (FTP) or [OCLC Product Services Web \(PSWeb\)](#) under [Batch Services records and reports](#).
- Receive records on IBM 3480-compatible cartridge, 4mm (DAT tapes), 8mm, limited to 5.0 GB capacity cartridges or lower, 3-1/2" diskette or e-mail attachment.

Reclamation project

There are times when an institution's holdings have not been maintained in WorldCat for one reason or another. If this is the case, you may have no easy way of identifying the records for which you need to cancel or set holdings. When this occurs, a *reclamation project* may be the easiest way to get your holdings in WorldCat in sync with the holdings in your local system. This type of project is meant as a one-time solution.

In a reclamation project, an institution sends the records from their local system, for which they want holdings set. After batch processing is complete, any holdings not updated by this effort will be deleted from WorldCat. This is done using a Date Stamp (typically the date on which you extract your records to submit them to OCLC). These "untouched" holdings will be deleted in a separate process called a Scan/Delete. A scan is done on WorldCat, looking for your holdings symbol. When your symbol is found and the Set Holdings Date Stamp is earlier than the Date Stamp agreed upon, your holdings symbol is removed. Any holdings set on the Date Stamp date or after, will remain set on the record in WorldCat.

To initiate a reclamation project, first complete a Complete a [Batchload Order form](#).

After receiving a Batchload Project ID, submit a file (or files) containing ALL of your current holdings. Based on information supplied on the Batchload Order form, an OCLC database specialist will set up your project for processing. During

the setup process they will ask you what date to use for the Scan/Delete of "untouched" holdings.

OCLC will perform a Scan/Delete after processing your files, and remove holdings only from those records that have not been touched by the dataload. Any holdings set online after you've submitted your records for batchload (the Date Stamp) will NOT be removed by the Scan/Delete.

There is no reason to stop online cataloging, if that is part of your regular workflow. OCLC does recommend that you stop deleting holdings until the Reclamation is complete.

OCLC has offered members a one-time free reclamation project. Members are expected to keep their holdings in WorldCat current from that point forward, either online through Connexion or using ongoing Batchload.

Scan/delete processing

Scan/delete processing removes *all* occurrences of an OCLC institution symbol. All local data records (LDRs) can also be removed with this process.

OCLC does not require a file of records from your institution to perform this function. It is actually a scan process that looks for a given OCLC institution holding symbol and removes it from WorldCat. This scan process can only be done at the institution holding level, not the local holding library level. To remove a subset of your institution's holdings, a file of records must be submitted for a cancel holdings project. (See [Cancel holdings](#).)

To request a scan/delete, the following information must be submitted to OCLC Batch Services in writing (letter or e-mail):

- OCLC institution symbol
- Institution name
- Reason for the request (e.g., institution closed, merging with another institution, etc.)
- The type of scan/delete being requested:
 1. Remove all holdings from WorldCat for your institution, all institution holdings (IHBs) and delete all LDRs if any exist
 2. Delete LDRs and remove the holding of the related bib records (related IHB)
 3. Delete LDRs only (holding will remain on any related IHB)

OCLC does not have a specified turnaround time for completing scan/delete requests. If there is a need to have a scan/delete done within a specific time period, please include this information with the request.

Scan/delete is a *permanent process that cannot be reversed* once it has been performed. There is *no charge or related product code* for this service.

Adding original records

Originally cataloged records can be added to WorldCat through batchloading. Currently, this is only an option for single institution batchload projects. Adding this option for groups is currently not available.

Unresolved records that do not find a matching record in WorldCat after batchloading unique-key and extended matching can be examined to determine if they are 1) original records to WorldCat and 2) meet minimum cataloging requirements.

Records should adhere to national standards for cataloging, including those for descriptive cataloging and forms of headings. They should also meet OCLC K-level standards for completeness, as described in [Bibliographic Formats and Standards, Third Edition](#).

It should be noted that records can fail to find a match through batchloading for other reasons besides being original to WorldCat. Significant differences between the institution record and the WorldCat record can prevent a record from finding its match. Manual searching is needed to determine the true duplication rate.

If fewer than 20% of the non-matching records duplicate records in WorldCat, and if these records adhere to OCLC's cataloging standards, a new definition and setup will be created to add records that do not contain validation errors. This type of processing is called *conditional adds processing (CA)*.

The desirability of adding records is determined for the entire file, not record-by-record. Conditional adds analysis is a time consuming process that has to be "worked in" to the daily workload of Batch Support staff, while they continue to provide ongoing support for existing projects and define new projects.

Group batchload

Batchloading for groups has many similarities to single institution projects, but there are some distinct differences that need to be taken into account when planning a *group batchload* project.

A group of institutions can submit files of consolidated records for batchloading; that is, a single file of records can be processed to set or cancel multiple OCLC holding symbols in WorldCat. To do this, the records must contain library identifiers *within the bibliographic records themselves*. It does not matter where these codes (library identifiers) reside within the bibliographic record as long as they are in the *same field and subfield throughout the file*.

A *translation table* is required for group processing. This table lists the library identifiers (in the bibliographic records) and the corresponding OCLC institution symbols that are to be set or canceled during batchloading.

A *"unique" group batchload symbol* will be assigned for batchload projects when the group symbol also represents one of the members of the group for whom holdings will be set. The OCLC database specialist assigned to the group will also request that the unique symbol be profiled at OCLC, as it is for internal batchloading purposes only. No holdings will ever be set for that unique group symbol; it is used primarily for tracking files and billing.

Batchload projects are very limited in the area of *local data retention*. A group project can be set up to retain, for example, a system control number in an 035 field or a 9XX field.

When MARC record output is requested with a group batchload project, it should be noted that a copy of the matching OCLC MARC record would be created for each institution setting holdings on a record. For example, one record containing three library identifiers will set three holdings, and three copies of that same matching OCLC record will be found in the resulting output file. The group will also be billed for three OCLC MARC records returned (CTR 3338).

Group post-processing reports contain a list of holdings set or cancelled for each member of the group (symbol on translation table). That report gives the group members statistics of that specific batch processed and the accumulated set or cancelled stats from the beginning of the project. On [OCLC Product Services Web](#) (PSWeb), group post-processing reports end with a .GROUP extension.

PCC BIBCO participants and batchloading

Batchloading has been enhanced to allow PCC BIBCO participants to:

- Submit files of mixed PCC and non-PCC records.
- Set holdings in WorldCat for both PCC and non-PCC records.
- Replace WorldCat records of lesser quality with PCC records.
- Perform maintenance on their own PCC records by resubmitting those records in subsequent files.
Note: A PCC record will only replace an existing PCC record in WorldCat if it is from the same inputting institution.
- Add non-matching PCC records to WorldCat.

PCC BIBCO participants can include their PCC records in the same file as non-PCC records. Using the criteria below, OCLC will split the PCC records out of your incoming batchload file. This process creates two files, one containing PCC records and another containing non-PCC records.

Criteria used to identify PCC records:

1. The encoding level (Leader 17) is blank or 4.
2. The bibliographic level (Leader 07) is not "s" or "b".
3. Field 042 is present and contains a single subfield (subfield a) with contents "pcc" or "PCC".
Note: The uppercase form is not valid; however, the pre-processing software will correct it.

Two separate batchload projects and setups will be created for your institution. One is for the processing of the PCC records and another for the non-PCC records. The non-PCC records are processed through the normal batchload job stream. PCC records are processed through a separate job stream created just for PCC records. This job stream allows for setting holdings, replacing records

and adding records as described above. There is no additional evaluation fee for the separate PCC definition and setup.

Of course, you may separate PCC from non-PCC locally prior to submission to OCLC for separate Batch Processing.

Batchloading of non-MARC records, or numeric search keys

In response to member library requests, OCLC has expanded the scope of batchloading by developing custom processes that help "build" usable MARC records from "patterned data" received in non-MARC formats, and from numeric search keys.

Patterned data

"Patterned data" received in non-MARC formats must meet the following guidelines:

- *Labeled data-field* descriptions immediately precede their content (e.g. AUTHOR: Smith, John TITLE: My Life Story) or labeled columns of data in a spreadsheet.
- *Delimited data-fields* and records are separated by a unique identifiable character or series of characters (e.g. tab delimited files).
- *Positional data-field* descriptions and their content are related by record position (e.g. fixed length formatted records).

None of these data forms are mutually exclusive. Sometimes records can be a composite of all three of these basic data types. Examples of commonly used non-MARC data formats are Microsoft[®]Excel spreadsheets, Access[®] databases, and Microsoft[®]Word documents.

Once the MARC records are built, the file is analyzed like any other file submitted for batchloading.

There is no additional cost to have non-MARC data converted to MARC for batchloading, although it should be noted this pre-processing could extend the 90-day turnaround time.

Numeric search key

By working in a text editor or spreadsheet file, a user can create a list of numeric search keys that can be used to retrieve matching records, set holdings, or cancel holdings in WorldCat. Only the following numeric search keys may be used for a numeric search key processing project.

- Library of Congress Control Number (LCCN)
- International Standard Book Number (ISBN)
- International Standard Serial Number (ISSN)
- OCLC record control number

Send numeric search key files via e-mail attachment to: batchload@oclc.org.

Local data record updating

OCLC local data record updating (previously known as LDRUS), provides automated batch processing that allows OCLC Union List or Group Access/Union List group participants to add and update copy-specific-holdings data in WorldCat.

The service takes library serials holdings records, batch processes them, and creates local data records (LDRs) that are then added to WorldCat for use in OCLC Union List, WorldCat Resource Sharing, and OCLC FirstSearch.

Local data record updating is intended to be an ongoing service. Regular updates are suggested, but the schedule varies according to individual needs. The service can also be used as a one-time project, with holdings maintained manually thereafter.

There is currently no charge for batch processing of serials holdings data via local data record updating. Additionally, the batchload evaluation fee (TAP 3921) is currently waived for libraries whose holdings data is compliant with the [MARC 21 Format for Holdings Data](#). The batchload evaluation fee (TAP 3921) will be charged for libraries submitting non-MARC files. Additional charges may apply if the data does not meet the local data record updating specifications and requires custom programming.

Post-processing reports are made available on the OCLC Product Services Web after each batch of records is processed. There is no record output available from this product.

General Guidelines

1. All holdings records, regardless of format, **must** include the OCLC control number of the bibliographic record to which they are attached.
2. The institution must be able to supply all holdings records associated with a specific bibliographic record.
3. There must be only one holdings record per copy.
4. The institution must participate in an OCLC Union List or Group Access/Union List group.

Please consult the [OCLC Batch Services: Local Data Record Updating Guide](#) for detailed information about record requirements, procedures for initiating projects and procedures for submitting files.

4 How to Start a New Batch Processing Project

Getting started

To initiate any new Batch Processing project, two items are necessary:

- A completed [Batchload Order form](#). Your institution or your regional service provider can complete the order form.
- An initial file of records to evaluate.

Note: An Application and Agreement for Tapeloading Membership Agreement and [EDX Assessment form](#) may also need to be completed, depending on the type of project and delivery method selected.

Once OCLC Batch Services receives both a completed order form and a file of records, the project is considered "in-house." Many times these two items are sent to OCLC on different dates, which is perfectly acceptable. It should be noted that the date the last item is received is used as the "in-house" date for that new project.

Input media for Batch Processing

Physical media

Records can be submitted to OCLC for Batch Processing on IBM 3480-compatible cartridges, 4mm (DAT) tapes, 8mm tapes (limited to 5.0 GB capacity cartridges or lower), 3-1/2" diskettes or CD-ROMs.

When putting data onto a physical media for Batch Processing the following should be considered:

- Only place one file (MARC records) on the tape (no volume or header files).
- Do not compress or write the records in an archiving format.
- Use the extended ASCII character set.

Always label the tape, cartridge or diskette with the **OCLC institution symbol**, **project ID** and the **record count**. A sticker label that will adhere to the flat portion of the tape or cartridge should be used.

Electronic transmission

Your institution can electronically submit data to OCLC for Batch Processing via FTP (EDX). Institutions that submit data on an ongoing basis should establish an EDX account. To do so, complete an [EDX Assessment form](#) if you are not already subscribing to another OCLC product via FTP. The EDX subdirectory (EBSB) will be created for that institution based on receipt of the [Batchload Order form](#) or [Batch Search Key Processing Order form](#).

An institution that is submitting data via FTP for a one-time Batch Processing project does not need to establish an EDX account. For one-time transfers, please contact OCLC at batchservices@oclc.org. Files submitted electronically via e-mail attachment should be sent to batchload@oclc.org.

Preparing files for electronic transmission via EDX

When submitting records to OCLC for Batch Processing via EDX, a DATA file containing MARC records, and a LABEL file containing characteristics of the MARC record file must be created for each file submitted.

Requirements for a LABEL file:

Listed below are the five mandatory fields for the LABEL file. Use the following guidelines to create the LABEL file in Notepad (preferred) or any text editor software.

- The LABEL file must contain the five mandatory fields listed below:
 - DAT (date compiled, including zeros)
 - RBF (number of records in DATA file)
 - DSN (exact name given to DATA file)
 - ORS (institution symbol)
 - FDI (project ID assigned to your batchload project)
- Separate each field with two spaces, for example, DAT <space><space> 20020520000000.0.
- Use carriage return or press <Enter> after each field.
- Assign a file name beginning with the word "LABEL" followed by a period and a specific file name extension, for example, LABEL.D020520.

Example of how the LABEL file should look:

```
DAT 20020520000000.0
RBF 4027
DSN DATA.D020520
ORS VJA
FDI P000806
```

Note: If you are using an Innovative/III system, do not follow the LABEL file instructions above. Refer to your systems manual or contact the Batch Administration Section for further details.

Sample FTP session

Internet FTP Sample Session

The following instructions describe using the Internet FTP with OCLC.

Note: Your session may look different. If you have trouble, contact OCLC Batch Administration (ext. 6026 or e-mail batchservices@oclc.org) or OCLC User & Network Support at (800) 848-5800 or (614) 793-8682, or e-mail support@oclc.org.

Step	Action	Example
1.	Enter FTP address.	edx.oclc.org
2.	Type logon.	t[xxx]1 (xxx = OCLC institution symbol)
3.	Type password.	
4.	Change directory.	cd 'EDX.EBSB.[xxx].FTP' (Enter single quote before and after directory name. xxx = OCLC institution symbol.)
5.	Verify that your transfer type is set to <i>binary</i> .	bin
6.	Complete the action to "put" the DATA file to OCLC.	put DATA.D000301 (The DATA file should be named DATA.D followed by the date in the YYMMDD format.)
7.	Complete the action to "put" the LABEL file to OCLC.	put LABEL.D000301 (The LABEL file should be named LABEL.D followed by the date in the YYMMDD format.)
8.	The system responds with a message when the transfer is complete.	
9.	End your FTP session.	

E-mail notification

Your institution can receive an e-mail confirmation when files are successfully picked up. Prior to sending files, you must provide at least one e-mail address to OCLC Batch Administration to set up e-mail notification.

Setting expectations for new batchload projects

A new batchload project can take anywhere from 30 to 90 days to set up (see **Diagram 1**). The 90-day turnaround time starts when the project is "in-house." On average, new projects are turned around in 50 days.

If there is some type of funding deadline or other time period that needs to be considered, OCLC must be made aware of this upfront. This can be achieved by including this information in the Special Instructions area of the [Batchload Order form](#), or by having your regional service provider send an e-mail message to OCLC Batch Services. Special circumstances will be considered whenever possible. However, the best course of action is to get the project to OCLC at least 90 days prior to any needed completion date.

Many times an institution will ask if sending a smaller file of records—a subset of the entire file—will speed up the initiation process. The answer to this question is no. Smaller files are less likely to be representative of the records an institution sends to OCLC for the project. As a result, the processing of subsequent files may be delayed or be in error. OCLC recommends that you send all the records you have ready for batch processing when you initiate the new project.

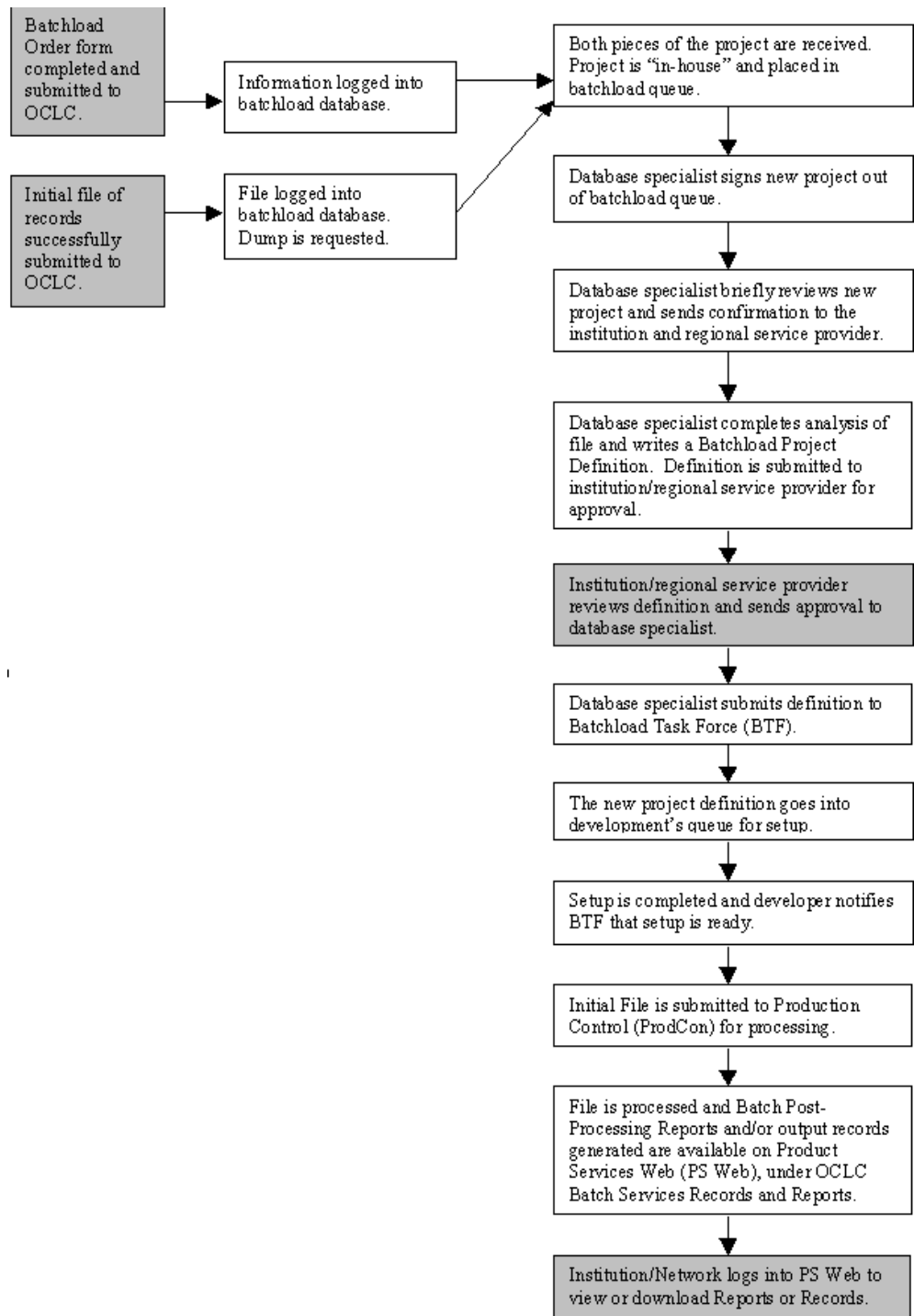
It is not necessary for you to send a "sample file" of records prior to sending live data for analysis and subsequent processing, but you do have this option.

Why do new batchload projects take up to 90 days to evaluate and process?

It is often asked, "*Why do new batchload projects take up to 90 days to evaluate and process?*" The diagram below will help answer this question. All new batchload projects go through the following steps:

Diagram 1: Workflow for new batchload projects

Shaded boxes represent institution action needed. Other boxes are OCLC action needed.



5 Unresolved Records

Unresolved or non-matching records from batchload can be handled in the following ways:

- Disregard the unresolved records.
- Unresolved batchload records are available on the [OCLC Product Services Web \(PSWeb\)](#). These unresolved records have been processed through batchload and DDR processing. Under OCLC Batch Services Records and Reports, files ending with .UNRES contain the non-matching records from the batch referenced in the file name (e.g., D020905.B4068.UNRES). Users can save the file of MARC records to their local work station and resolve any unresolved records as they choose, either by importing them into the local bib file of Connexion, or importing and editing them in their local system for future batch submittals.
- For one-time or annual batchload projects only, a text version of the unresolved records can be returned to your institution via e-mail attachment.