D. Processing IR workflow records into the Live IR

NOTE: Go to the C. Creating Records in your IR Workflow Space page to create records in your IR workflow.

Processing IR workflow records into the Live IR

Once you have uploaded records or manually created a record, the title for each record will be displayed in your workflow space. You can now process the record through each step in the workflow. Once processed, you can upload the record into the live IR or keep the record in your workflow for further modification. Don’t worry if you upload it. You can withdraw and modify any content in the record at any time.

NOTE: Throughout the repository, each field has a question mark icon upper right that can be clicked to provide information about that field.

NOTE: Fields marked with a gold bordered star must be populated.

1. Sign into the IR.
2. Click on Manage deposits in the upper task bar. This will bring you to your personal workflow space.
3. At the top of the page are your options for viewing all your records. If you want to display only the records you have imported and have not yet processed, un-click all the boxes except User Workarea.

4. The menu to the right of each uploaded record is shown below:

Click on the Edit Item icon to start processing your uploaded record through your workflow.

- The five basic steps in the workflow process are shown at the top and bottom of each workflow window. They are Type, Upload, Details, Subjects, Custom, and Deposit. You can go back and forth through the workflow by clicking on these buttons at top or bottom of each window.

1. Type page

The first step is Type. Clicking on this button will bring you to the Type window. Here you can select the metadata menu for the digital objects you want to add to the IR. Each Item Type has its own custom metadata “menu”. For example, if you select Thesis for your item type, the Details...
Once you have selected the **type** click on **Upload** in the top banner to go to the upload page

### 2. Upload page

Use this page to upload the documents, datasets, images and other digital objects for which you are creating the record. You can upload and attach as many digital objects in as many formats you want to the record.

- You can **Browse** to select a digital object or you can paste in a URL to the object. Click on **Upload** to upload the digital object.

**NOTE:**

- It is the responsibility of the depositor to determine copyright restrictions and use policies on any items they deposit in this repository. As with any publication you should exercise reasonable care to ensure that you item does not infringe upon anyone else’s copyright.
- Many publishers place restrictions on access to published material. The number of such publishers is diminishing, these restrictions are constantly changing, and there are a number of complexities in the models offered by publishers. The Sherpa/Romeo database is a great tool for determining upload restrictions for different publications. Click the site link and enter your publication title or ISSN to check current restrictions.

You may wish to use the SHERPA RoMEO tool to verify publisher policies before depositing.

**NOTE:**

You can utilize the **Library Databases** and **Journals** to retrieve full text of articles and other items for upload into a record.

- Once you have uploaded an item you can assign attributes to the item such as **item Visibility, Embargo period**, and other attributes.
- Click on **Show Options** to open up the attribute options window

- Click on **Content** and select the content type
3. Details page

This is the main metadata page for the record. Each item type has a record with its own custom metadata (see Type page above). The details described below are for the item type called Paper.

- **Format** should automatically be the format of the item you uploaded. If not, select the proper format from the drop down window.

- **Description**: add a short description of the item, e.g. Paper for a journal article.

- **Visible to**: click on the level of visibility you would like for this item.

- **Licence**: the choice for licenses is presently limited to Creative Commons licenses. Usually the default UNSPECIFIED is used.

- **Embargo expiry date**:

  If you would like to limit access until a particular date enter it here. Then click Update Metadata.
a. Title field

should be populated with name of the article or digital item from the source record.

NOTE: ISI database and others export titles in all uppercase. These should be replaced with the Original title. For example:

Original imported Title:
THE USE OF CHITOSAN AS A FLOCCULANT IN MAMMALIAN CELL CULTURE

Replacement Title:
The use of chitosan as a flocculant in mammalian cell culture

b. Abstract field.

If not populated you can look for the abstract in the document once you bring it up to determine the CSHL author(s)

For some article formats there may be a summary or first paragraph that may be used as an abstract.

c. Authors fields

Use a standard name form (called an authority name) for each author's name. Do not use multiple name forms for the same name or this will complicate retrieval. An author's name is the primary metadata for document retrieval. Insure that you use only one form of a name and that the name you use is not a duplicate of another author's name in the repository. See the Custom page below for our CSHL Authors ontology which contains authority names for CSHL authors. The standard name form is Smith, A. Y. As you type in names the IR automatically checks the name you type in against all the authors names previously deposited and presents a list of these names. You can click on a name in the list to select that name and automatically tag the record to that name.

d. Affiliations field

The correct format is from largest to smallest institutions, and then address from smallest area to largest area. An example would be

Cold Spring Harbor Laboratory, Watson School, One Bungtown Road, Cold Spring Harbor, NY

e. Communities field

Select the appropriate CSHL Community/Communities for your item (e.g. one of the labs for a publication, Watson School for a thesis, Meetings and Courses for a course program, etc.) You can select multiple communities by holding the control key while selecting.

f. Publication Details field

Click on Refereed or non-refereed. Most CSHL publications are refereed.

g. Status

The default is Published. Choices are:

<table>
<thead>
<tr>
<th>Status</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Published</td>
<td>Final Published version</td>
</tr>
<tr>
<td>In Press</td>
<td>version of paper after Peer Review and with all revisions complete but not published</td>
</tr>
<tr>
<td>Submitted</td>
<td>version of paper submitted to publisher before peer review</td>
</tr>
<tr>
<td>Unpublished</td>
<td>version of paper before it is submitted to a publisher</td>
</tr>
</tbody>
</table>

h. Journal or Publication Title

The title of the journal, publication etc. Use the full title.

i. ISSN field

The ISSN number is a unique identifier for a publication. This metadata is handy for determining current copyright information. The standard format is 3971-7036

j. Publisher

Enter the full name of the Publisher of the journal or publication where available

k. Official URL
The best official URL for a journal article/book/book section etc is the URL able to accessed by a majority of potential users.

You can also check the URL in the record itself in the Deposit Page section below.

- **URL choice**
  
  i. **The best URL is directly to the record from PubMed** e.g the URL should look like this http://www.ncbi.nlm.nih.gov/pubmed/14625589 and not http://www.ncbi.nlm.nih.gov/pubmed/?term=The+crystal+structure+of+the+Argonaute2+PAZ+domain+reveals+an+RNA+binding+motif+in+RNAi+effector+complexes
  
  ii. **The second best is to the document within the publisher or journal site. This URL should be to page with as much freely accessible content as possible i.e.** http://nar.oxfordjournals.org/content/19/3/579.abstract not http://nar.oxfordjournals.org/content/19/3/579.long
  
  iii. **The third best is to a database record (i.e. Scopus).**

- **URL format:** The format must be http://xxxxxxx

  Replace the URL in the URL field with this URL.

- **Check URL.**

  **You must verify that the URL is live.** The best way to check the URL is to copy the URL from the field, open a browser window, paste the URL into the window and confirm that it is live.

- **Only one URL should be entered in this field. Enter all other URL’s in the Related URL’s field (see below).**

  i. **Volume**

  The volume number of the journal in which the article appeared

  j. **Number**

  The issue number of the journal in which the article appeared

  k. **Page Range**

  Numerals only. The sequence of pages of the item. Do not enter pp. (e.g. 21 to 34). For electronic publications using article numbers as page ranges (e.g. article number X23697) enter the article number in the Additional Information field below

  l. **Date**

  The date this article was completed, submitted to a publisher, published or submitted for a Ph.D.

  m. **Date Type**

  The event to which the date applies

  n. **DOI**

  DOI numbers are automatically linked to full text by the IR so this field should be populated whenever possible. **DOI’s can be the best way to link to full text because they are persistant vs. URL’s**

  Format is 10.1158/1078-0432.CCR-07-5030

  o. **Related URL’s**

  Enter all related URL’s not entered in the Official URL field if appropriate

  The Related URL location is used to add the URL for the article/book/book section at the publisher/journal official site. This URL should be to page with as much freely accessible content as possible i.e. http://nar.oxfordjournals.org/content/19/3/579.abstract not http://nar.oxfordjournals.org/content/19/3/579.long

  The Related URL location is also used to add the URL to locations that might contain data related to the paper, be they links to outside sites or to an entry in the IR.

  All added Related URLs should have their URL type specified to its appropriate type.

  p. **Funders**

  enter funding information such as "Non-US Government Funded", grant numbers and other funding information

  q. **Projects**
enter project codes for this item

r. Contact Email Address

enter contact Email address where appropriate (e.g. corresponding author, reprint contract)

s. References

not presently mapped to import formats. Enter article references here

t. Uncontrolled Keywords

Along with Author Name and the item title, Keywords and subject terms are critical metadata for retrieving articles in the repository. It is critical to format and/or add the keywords and formats that users will be inputting to retrieve the item.

**Punctuation**

Remove all non-standard punctuation. Many databases add punctuation such as commas, hyphens etc during processing. These make the keywords useless for searching once the keyword are uploaded into the IR because the punctuation must then be added by a searcher. For example the comma in the phrase DNA, host should be removed or an IR searcher will need to enter the search term in that form, instead of host DNA which is the standard form. When in doubt, add any version of a keyword a researcher might use to retrieve an item.

**Examples:**

PubMed adds a number of slashes and asterisks as part of their manual curation process. These should all be removed to facilitate search and retrieval in the IR. For example \*therapeutic\cancer should be changed to therapeutic cancer

**Upper Case Terms**

ISI Web of Science often generates all keywords as upper case. Standard case versions need to be added under each term or a searcher will need to enter all upper case words. For example I show a list of upper case keywords from ISI before and after I modified them

**Before Modification**

IDENTIFICATION
ACTIVATION
ONCOGENE
INVITRO

**After Modification**

Identification
Activation
Oncogene
In-vitro

Note that:

- The first letter of the lower case versions are all upper-case. This is because the IR automatically recovers both upper and lower-case versions of every term starting with an upper case letter, so by making the first letter upper-case you automatically create both an upper and lower case version of the term.

Remember that you should add any version of the term or other terms a researcher may use to locate the item.

u. Additional Information

Delete any text that appears in the Additional Information box

4. Subjects page
NOTE: To make additions and changes to the Subject ontology you must have administrator privileges. Please email Institutional Repository Administrator with any additions or changes.

NOTE: To modify the Subject ontology please go to the IR Subject Ontology page

This page contains an ontology of CSHL specific subject terms.

There are two ways you can select subject terms as tags for your record:

1. Search for subject terms and select them
2. Browse the subject ontology and select subject terms

1. Search for Subject Terms and select them

- Enter the term into the Search for subject window and click Search

- Select the subject term(s) from the list by clicking on Add. You can select multiple terms by clicking on each Add button.
- You can remove selected terms by clicking on the Remove button next to the selected term

2. Browse the subject ontology and select subject terms.

- Click on the Subjects Button, if you are on the Subjects page click on Clear to bring up the subject ontology
Click on the buttons to drill down through the ontology. Select the subject term(s) by clicking on Add. You can select multiple terms by clicking on each Add button. You can remove selected terms by clicking on the Remove button next to the selected term.
This page was specifically created for custom CSHL metadata.

1. **PMCID field**
   1. Enter the PMCID number if it was not auto-imported from PubMed Central. PMCID is the PubMed Central ID number for publications made Open Access by NIH Open Access policy. This metadata is valuable because a PMCID number entered into this field will link to full text on PubMed Central once the record goes live. The format is PMCXXXXX, e.g. PMC3166838
   2. **NOTE:** The **PMCID** is **not the same as the PMID**. The PMID is the PubMed number assigned to every article in its open access and non-open access database. PMID numbers, while useful as metadata, do not link to full text in this IR.

2. **Dataset Accession ID fields**
   1. Databases and their data sets accession identifiers associated with this paper.

   **NOTE:** GeneBank and GEO accession numbers entered into these fields using the proper format become live links to the dataset(s) in each database when the record goes live.

   b. Geo and GeneBank accession number formats are shown below

<table>
<thead>
<tr>
<th>Database</th>
<th>Dataset accession number format</th>
</tr>
</thead>
</table>

Custom
Accession numbers can be entered one to a row after the database identifier (i.e. GEO). You can click on More input rows to enter more numbers. You can also paste up to 20 accession numbers after the database identifier in each row. Pasting multiple accession numbers in one row means that when the record goes live, clicking on any accession number in that row will bring up all the datasets for all the accession numbers in that row.

Image 1 shows dataset accession numbers entered one to a row. Image 2 shows multiple dataset accession numbers entered after a database identifier in one row.

3. CSHL Authors

Determine the CSHL Authors from the affiliations listed in the database record or from the publication itself. You can also confirm full names of current CSHL employees by searching the FACES database.

You must use the most complete form of the authors name e.g. Ware, D. H. would be Ware, Doreen H.

Select the Full Name of every CSHL Author from the CSHL Authors list. You can select multiple CSHL Authors by holding the control key while selecting.

If the name is not listed please determine the most complete form of both the abbreviated and full name for the author.

To determine the most complete form of the name:

- You can confirm authors full names by performing an author search in Scopus database from the Library homepage by clicking on the Author tab, entering the author's last name and "Cold Spring Harbor" for the affiliation, clicking on search and then by clicking on the author's name from the retrieved names to retrieve the author's publication record.
You can also confirm full names of current CSHL employees by searching the FACES database. The Library of Congress Authorities database contains authority names but Scopus is actually more up-to-date for scientist's names.

**NOTE:** To make additions and changes to the CSHL Authors ontology you must have administrator privileges. Please email Institutional Repository Administrator with any Author name additions or changes.

**NOTE:** To modify the CSHL Authors ontology please go to the IR CSHL Authors Ontology page.

### 6. Deposit page

Once you have completed the previous pages this page allows you to deposit the item. Where the item goes from here depends on your user status.

1. **Users:** when you deposit an item it automatically goes into review by an editor and/or administrator
2. **Editor:** when you deposit an item you can put it into your own review queue or upload it to the live IR
3. **Administrator** when you deposit an item you can put it into your own review queue or upload it to the live IR. You can also review all the items created by users and editors who have moved items into your review queue

#### 1. Users

1. Click on **Deposit Item Now**

Clicking on the deposit button indicates your agreement to these terms.

b. You may need to correct any remaining issues like non-working URL's before being allowed to deposit. Migrate back to the appropriate page in the record to correct these issues if necessary. If the record is deposited you should see **Item Has been deposited** and **Your item will not appear**... messages below

**NOTE:** The **Details** tab allows you to access the record to check the URL and other links, modify the record and re-deposit it.

#### 2. Editors and Administrators

1. Click on **Deposit Item Now**
Reviewing Records.

To review records go to the E. Review and Removing Records in the Live IR page

Back to Table of Contents