Getting Started How-To

Video training is available on our YouTube channel

- Homepage Overview
- Communities & Collections
- Adding/Submitting Items
- Editing Items
- Managing Permissions (EPeople & Groups)
- Usage Statistics
- Bulk/Batch Metadata Editing
- Bulk/Batch Content Uploads
- DSpace Permissions Overview
- Individual Item Permission Changes
- Bulk/Batch Permissions Changes (Advanced)
- Checking Item Links

Homepage Overview

- For example: http://demo.dspacedirect.org/
- Provides a good overview of the basic features DSpace offers:
    - Side Menu
        - Search
            - Based on where you are in system, this also lets you search an individual Community or Collection
        - Browse By options
            - Based on where you are in the system, these also let you browse an individual Community or Collection
        - My Account
            - Options to login & register (more info below)
        - "Discover" works by authors/subject/dates
            - Based on where you are in the system, these lists will change to become Community/Collection specific
        - RSS Feeds
            - Based on where you are in the system, these links provide RSS/Atom Feeds specific to an individual Community or Collection
    - Content Area
        - Introductory Text (customizable)
        - "Top-level" Community listing (just displays the very top of the hierarchy)
        - Recently Added Items
            - This listing also appears on individual Community or Collection pages (at which point it is Community/Collection specific)

Communities & Collections

- For example: https://demo.dspacedirect.org/community-list
- Important to understand the difference between Community & Collection
    - Collections only contain Items (works or documents). In other words, you *always* add an Item to a Collection
    - Communities are used to "organize" Collections. Communities can contain other Communities (sub-communities) or Collections (but CANNOT contain Items).
        - Often Communities are used to "mirror" the organizational structure within an institution (colleges, departments, research units, etc).
        - Communities can also be used to assign rights to specific groups (can give a department control over their own Community)
- **How-To: Create a new Community**
    - Login as an Admin
    - Click "Create Community"
    - Community level "metadata" (only the "Name" is required) - Any of this info can be edited/changed at a later time
        - *Name*: the name of the Community (REQUIRED)
        - *Short Description*: A short "blurb" about this community (only displayed during search/browse as a very basic description)
        - *Introductory text*: The longer description about the Community displayed on its homepage (may include basic HTML if you want to add formatting, e.g. bold text, hyperlinks, etc)
        - *Copyright text:* If this Community needs a special note regarding copyright, one can be added (may include basic HTML). For example, if all works in this Community are copyrighted by a particular publisher/department it can be noted here.
        - *News:* Optional news section which will be displayed below the introductory text.
        - *Logo:* Optionally, a Community can have its own logo
    - Click "Create"
    - Once created, you are immediately moved to the "Edit Community" interface, so that you can assign roles if you wish.
        - For Communities, the only "Role" is Administrator. An "Administrator" has full rights (add/update/delete) on this Community and any Sub-Communities, Collections or Items that are contained within the hierarchy under this Community.
    - Click "Return"
- **How-To: Create a new Collection**
    - Browse to the newly created Community.
    - Notice the "Context" menu has different options (*HINT: The "Context" menu's options change based on where you are in the system.*)
        - Edit Community
        - Export Community (only for Admins - exports all content/metadata into a Zip file)
        - Export Metadata (used for bulk/batch metadata editing in a CSV)
        - Create Collection

        - Create Sub-Community (A Sub-Community is just a Community that happens to be within another Community...it's no different than a normal Community)
    - Click "Create Collection"
    - Collection level "metadata" (nearly identical to the Community metadata). Only a few (minor) additions:
        - *License:* If this Collection requires its own custom deposit license (i.e. it needs to be different from the site-wide deposit license), you can enter that license text here. It will be displayed during the deposit process **instead of** the normal deposit license. (This option is rarely used)
        - *Provenance:* An Administrator-only field which can be used to describe or add notes about the history/provenance of this particular collection. It is never visible to users.
    - Click "Create"
    - Again, once created, you are immediately moved to the "Edit Collection" interface, so that you can assign roles if you wish.
        - *Assign Roles:* Collections offer additional Roles (all are optional):
            - Administrators - people who have full rights (add/update/delete) on this Collection and any Items contained in this Collection
            - Submitters - people who can deposit new content to this Collection
            - Default read access - people who can *view/download* any new content added to this collection (not retroactive - changing this does NOT automatically change view/download rights on existing items in the Collection)
            - Reviewer roles: There are three roles which have to do with reviewing newly deposited content (before it becomes publicly available). All are optional. If you enable multiple steps, they will always occur in the order that they are listed (and people added to that step will receive an email whenever new content is added that needs review)
                - Accept / Reject step
                - Accept / Reject / Edit Metadata step
                - Edit Metadata step
        - *Content Source:*
            - Optionally, a Collection can be setup to "harvest" all of its content from an external location (via OAI-PMH and/or OAI-ORE). This is **rarely** used, unless your DSpace is aggregating content from other locations.
        - *Curate:*
            - This tab offers some basic "curation" / reporting scripts that can be run across your content. By default only a few reporting scripts are available (unless you have DuraCloud backups available)
                - Profile Bitstream Formats : Report what file formats are contained in Items within this Community/Collection
                - Check for Required Metadata : Double check all Item metadata, ensuring all required fields are filled out.
                - Check Links in Metadata : Double check all URLs in Item metadata, ensuring all links are still valid (and none throw a 404 Not Found error)
- **How-To: Hide a Community or Collection from Public View/Restrict Access to an Existing Community/Collection**
    - Unfortunately, Communities/Collections *always* show up in Browse by Community/Collection. However, you can access restrict a Collection to Administrators, so no one else can access the Collection homepage, etc. We have an example at https://demo.dspacedirect.org/handle/10673/337
    - To create a Collection only visible to Administrators, you'll edit the Policies on the Collection.
        - Click "Edit Collection"
        - Click "Assign Roles"
        - Click "Edit authorization policies directly" (link at bottom)
        - Change all policies to list "Administrator" group instead of "Anonymous"
        - Click "Save"
    - This change is NOT retroactive to existing items in a Community or Collection. to restrict access to an existing Community or Collection, follow the steps above and then follow the steps described here: Batch Permission Change

Adding/Submitting Items

- Items are what you will be working with the MOST (after you get your Communities/Collections setup). They contain metadata and, optionally, bitstreams (files).
- Two ways to submit an Item to a Collection
    - From your "Submissions" page (in the "My Account" box).
    - OR, browse to a Collection that you have rights to deposit into, and click the "Submit a new item to this collection" link
- **How-To: Submit an Item**
    - It is a multi-stage process.
        - You can "Save & Exit" at any time (your changes are also auto-saved each time you transition to a new page). You can restart any unfinished submissions from your My Account "Submissions" page
        - You can also move backwards if you realized you forgot something in a previous step
        - *NOTE: These deposit steps can be tweaked/changed/rearranged, but any changes must be performed by DuraSpace (for an additional fee, depending on the extent of the changes)*
    - *Step 1: Initial Questions*
        - Captures some basic info to determine what metadata to ask the user for
    - *Step 2-3: Describe (x 2)*
        - Captures the basic metadata about this new item on two pages. Only Title & Date are required (Date is auto-captured though, unless you say the item was previously published)
        - Behind the scenes this metadata is all stored as Qualified Dublin Core
    - *Step 4: Upload*
        - Optionally, upload one (or more) files to this item.
            - File Description can just be used to optionally describe the file contents (e.g. "Presentation slides" or "Video of talk")
        - Optionally, add an embargo date. If an embargo date is added, then the file will not be downloadable/viewable until **after** that embargo date has passed. (Administrators can still access the file)
    - *Step 5: Review* - just review everything previously entered (with an option to modify anything)
    - *Step 6: License*
        - This is the deposit license which all users must agree to before they can deposit their item.
        - NOTE: An electronically-signed copy of the deposit license is actually stored within the deposited item. (The copy is "signed" with the name of the user who agreed to the license & the date)
    - *Step 7: Complete*

        - Once complete, one of two things will happen:
            - IF the Collection has one or more "Review steps" enabled, then the Item will go into an "approval workflow". It will not be publicly available until the review is complete. If the item is rejected the submitter will be notified
                - The submitter can check the status of the review process from their My Account "Submissions" page
            - If no "Review steps" are enabled, then the Item is available immediately.

Editing Items

- **How-To: Editing an Item**
    - Browse to an Item (while logged in as an Admin)
    - Click "Edit Item" from the "Context" menu
    - "Item Status" tab
        - Basic info about the item in question
        - "Authorizations" - Edits the permissions on this Item. (Not recommended to tweak unless you know what you are doing.)
        - "Withdraw" - Immediately withdraws the Item. Withdrawing essentially hides the item and temporarily removes it from the DSpace archive. However, the item still exists, and can be restored by "reinstating" it.
        - "Move" - Moves the Item to a different Collection
        - "Make it Private" - Marks the item as Private. It is still in the archive but it is only accessible to Administrators until it is made public again.
        - "Permanently Delete" - Immediately deletes the Item. As noted, this is a permanent action and cannot be undone. The item is fully removed from the system.
    - "Item Bitstreams" tab
        - Allows you to add/remove Bitstreams (files) to/from the Item
        - You can also reorder Bitstreams, if multiple exist. This lets you determine which bitstream is listed first on the Item page.
        - **Bundles**: In DSpace, Bitstreams (files) are kept in "Bundles" (essentially just groups of files). There are three main Bundles which DSpace handles automatically:
        - *ORIGINAL* : These are files which were uploaded when the Item was created/deposited. These are also the files that are available for download/viewing within DSpace.
        - *LICENSE*: This is a "hidden" bundle which stores an electronically signed copy of the deposit license (which was signed when the Item was deposited). It is only viewable to Administrators.
        - *THUMBNAIL* : If one (or more) of the files in the "ORIGINAL" bundle were images (BMP, GIF, JPEG, PNG), then DSpace will automatically generate a Thumbnail version for display. The auto-generated thumbnail is stored in this Bundle. Note that thumbnails are generated via a service that runs overnight, so thumbnails for ideas added will not appear until the following day (i.e. allow for up to a 24-hour time period to elapse before thumbnails may be generated and available).
        - *TEXT* : If one (or more) of the files in the "ORIGINAL" bundle were common textual formats (HTML, Word, PowerPoint, PDF, Plain Text), then DSpace will automatically generate a plain text version of the document (for its search within document feature). The auto-generated plain text file is stored in this Bundle.
    - "Item Metadata" tab
        - Allows you to directly edit the Qualified Dublin Core metadata associated with this Item.
        - *BE CAREFUL*. It is assumed you know what you are doing. Metadata changes here are not validated in any way. So, anything you save will be accepted as-is.
        - You'll also notice here that there are several hidden metadata fields that are automatically generated/updated by DSpace (namely "dc.description.provenance" and "dc.date.accessioned")
    - "View Item" tab
        - If you've made any changes in the above tabs, this tab lets you "preview" what the new Item page looks like. That way you can quickly fix problems if you notice anything.
    - "Curate" tab
        - Similar to Communities & Collections, you can also run basic "curation" / reporting scripts on individual items.

Managing Permissions (EPeople & Groups)

- Under the "Administrative" menu, there are tools to add individual EPeople & Groups (which can then be used in Community or Collection "Roles").
- **How-To: Add a New EPerson (as an Administrator)**
    - Click on "People" -> "Click here to add a new E-Person."
        - Email (Both the user's email address and also their username)
        - Name (First & Last)
        - Telephone (optional, only available to Administrators)
        - (Other fields not necessary to fill out)
    - Find the newly created user & Click on it
        - Press the "Reset Password" button. An email will now be sent to the user's email address, which lets them setup a password in DSpace.
- **How-To: Register a New Account**
    - Logout of system
    - Click on "Login" link
    - There's a link to "Register" as a new user. This lets anyone setup an account with your DSpace. However, new accounts will not have **a ny** special permissions until you give their account special permissions. (So, even if a user sets up an account, they won't be able to do anything in your system until you allow them to.)
    - *NOTE: If you are using LDAP or Shibboleth with DSpace, new user accounts will be automatically created the first time a user logs into DSpace via LDAP/Shibboleth.* So, once they login, a DSpace E-Person will be automatically created which is associated with their LDAP /Shibboleth account.
- **How-To: Add a New Group (as an Administrator)**
    - *INFO: Groups can be used to manage permissions across several individuals.* You can choose to create as many (or as few) groups as you wish to help you manage DSpace permissions.

        - *Synchronizing Groups with LDAP:* If you are using LDAP, you can ask LYRASIS to setup a mapping between LDAP Groups (i. e. LDAP Organization Units or "OU") and internal DSpace Groups. This provides an automated way to "sync" group membership between an external system (LDAP) and DSpace's internal Groups. DSpace does NOT do this mapping automatically. It needs to be configured (by LYRASIS) for specific Groups.
    - Login as an Administrator
    - Click on "Groups" -> "Click here to add a new Group."
        - Name (Each group needs to have a name. Names can include spaces, so name it something that describes the group. E.g. "Mathematics Department" or "Staff" or similar
        - Add members to the group. Groups can contain individual EPeople or other Groups.
    - Click Save.
    - *HINT:* You'll notice in the Group listing a lot of groups named "COLLECTION_" or "COMMUNITY_". These groups are internal groups that DSpace creates for different Collection or Community Roles. They are essentially "special" groups which are directly associated with a particular role in a particular Collection or Community. (These special groups are also accessible when editing roles on Communities & Collections – see above).
- Once you have created user accounts (E-People) and Groups, you can use those to assign permissions within specific DSpace Communities or Collections (see the "Assign Roles" tab when editing a Community/Collection).
    - In addition, by adding EPeople or Groups to the DSpace "Administrator" group you can give users Site-wide Administrator permissions (add/edit/delete anything in the system).

Usage Statistics

- DSpaceDirect supports two primary types of usage statistics:
    - Google Analytics (if enabled) - DSpace can use Google Analytics to track page views/accesses with the system. This is similar to using Google Analytics to tracking any other website. The use of Google Analytics is recommended. Google Analytics more thoroughly weeds out spider and bot activity, resulting in more accurate statistics.
    - DSpace Internal Statistics - DSpace also tracks its own basic Usage, Workflow & Search statistics.
- DSpace Statistics are available site-wide, or on any specific Community, Collection or Item page. Based on where you are in the system, you will receive a slightly different statistical report.
- *NOTE: By default, DSpace Internal Statistics are ONLY available to Administrators. If you desire, DuraSpace can instead make them publicly available for your site*. Unfortunately, at this time, there are only two options: private (Admin only) or public.
- **How-To: View Site-Wide DSpace Statistics**
    - Browse to the homepage (logged in as an Admin)
    - Notice the "Statistics" box has three options:
        - *DSpace Usage Statistics*: At the homepage level, this gives a general count of "hits" for individual Items in the system. The report is extremely basic at this time and only lists the title of the Item.
        - *DSpace Search Statistics*: At the homepage level, this provides a summary of recent searches performed via the DSpace search box. There are also some historical reporting options.
        - *DSpace Workflow Statistics*: At the homepage level, this provides a general count of any approval workflows/reviews that have taken place. There are also some historical reporting options. *NOTE: If you do not use Workflow Reviewer Roles on any Collections, this report will always be empty.)*
- **How-To: View DSpace Community/Collection Statistics**
    - Browse to a specific Community or Collection
    - The "Statistics" box has the same three options, but now the results will be **specific to this Community/Collection** (and provide extra details)
        - *DSpace Usage Statistics*: At the Community/Collection level, this report provides extra detail about accesses of this Community /Collection page. You'll now get a summary of total hits to the Community/Collection homepage, along with monthly trends, and top countries/cities (where the hits are coming from)
        - *DSpace Search Statistics*: At the Community/Collection level, this provides a summary of recent searches performed via the DSpace search box (specific to this Community or Collection).
        - *DSpace Workflow Statistics*: At the Community/Collection level, this provides a general count of any approval workflows /reviews that have taken place (specific to this Community or Collection) *(NOTE: If you do not use Workflow Reviewer Roles on any Collections, this report will always be empty.)*
- **How-To: View DSpace Item Statistics**
    - Browse to a specific Item
    - The "Statistics" box has just one option, but now the results will be **specific to this Item** (and provide extra details)
        - *DSpace Usage Statistics*: At the Item level, this report provides extra detail about usage of the Item. You'll now get a summary of total hits to the Community/Collection homepage, along with monthly trends, number of file downloads and top countries /cities (where the hits are coming from)
        - *NOTE: Items do not have Search/Workflow statistics as both of those statistical reports are only applicable to Communities & Collections.*
- **How-To: View/Analyze Page Visits (hits) in Google Analytics**
    - In Google Analytics, this info is under the "Behavior -> Site Content -> All Pages" section. It should allow you to easily see the top pages visited, and filter that list based on a page name or path.
    - For example, to analyze the number of visits/hits on a single item's homepage (i.e the number of people who viewed the item metadata) ..
        - Either filter this list by putting in the URL of the item page into the searchbox
        - OR, click on "Page Title" as the "Primary Dimension" and put the title of the Item into the searchbox
- **How-To: View/Analyze File Downloads in Google Analytics**
    - File downloads are recorded as "events" in Google Analytics. So, this info is under the "Behavior -> Events -> Pages" section. By default it'll show the top URLs used to download Items, but it also provides ways to filter that information based on page name or URL.
    - For example, to analyze the number of file downloads for a single item (even if it has multiple files)
        - Either filter this list by putting in the URL of the item into the searchbox,
        - OR, click on "Page Title" as the "Primary Dimension" and put the title of the Item into the searchbox
        - In the results, clicking on the title will bring you to a page that lists downloads per file (if the item has multiple files)
- **How-To: View Total Number of Items in a Repository**
    - The total number of items in the repository is most easily found by navigating to Browse by Title, as title is a required field.
- **How-To: View the Number of Items Added in a Year**

- Unfortunately, the built-in DSpace statistics and Google Analytics do not list this data. The easiest way to determine the number of items added in a given time frame is to export the repository metadata to CSV, and look at the dc.date.accessioned in that export. The dc.date.accessioned field is not included by default in metadata exports from DSpaceDirect repositories; please contact dspacedirect@lyrasis.org if you would like to be able to export this field.

Bulk/Batch Metadata Editing

- DSpace's "Batch Metadata Editing" tool allows you to export sets of DSpace Item Metadata (all Items or just those in specific Communities /Collections) into a CSV file. The metadata values/fields in the CSV file can then be edited using Microsoft Excel (or OpenOffice Calc or LibreOffice Calc). Once editing is complete, you can re-import the modified CSV to apply the metadata changes into DSpace.
- Step-by-step Tutorials:
  - Detailed "Batch Metadata Editing" Tutorial (Maureen Walsh from Ohio State University). (Note that DSpaceDirect uses the XMLUI, so those are the instructions to follow in the tutorial)
    - https://kb.osu.edu/dspace/bitstream/handle/1811/47279/WalshMaureenP_SPARC-DSUG_2010-slides.pdf?sequence=7
  - Recorded "Batch Metadata Editing" tutorial (Leonie Hayes and Stuart Lewis from the University of Auckland). (Note that the recording is an interactive tutorial and it is hard to see the screen)
    - http://or2010.fecyt.es/Publico/videos/index.aspx?video=5
  - Also see the official DSpace Documentation at Batch Metadata Editing
- **How-To: Batch Metadata Editing basics**
  - Batch Metadata Editing can be performed at a Community or Collection level
  - Browse to a specific Community or Collection. The "Context" menu will display an option for "Export Metadata".
  - Click "Export Metadata". This will generate a CSV file that contains all the metadata for every Item within that Community or Collection hierarchy.
    - *WARNING:* For extremely large communities or collections the export (and import) processes may take a long time (or cause your site to significantly slow down). Therefore, DSpaceDirect currently only allows you to modify 500 Items (i.e. lines in CSV) at a single time. This 500 item limitation can be increased as needed, but it is not recommended (as it can cause performance issues with your site when using these tools).
  - Edit the CSV using either Microsoft Excel or OpenOffice Calc
    - More information on the CSV / Spreadsheet format is available in the DSpace Documentation section on Batch Metadata Editing
    - *EXCEL WARNING:* By default, Excel will not open a CSV in Unicode/UTF-8 encoding. This means that special characters may be improperly displayed and also can be "corrupted" during re-import of the CSV.
      - You need to tell Excel this CSV is Unicode, by importing it as follows:
        - Open Excel (and create an empty sheet, if one doesn't open by default)
        - Select "Data" tab
        - Click "From Text" button (in the "External Data" section)
        - Select your CSV file
        - Wizard Step 1
          - Choose "Delimited" option
          - In the "File origin" selectbox, select "65001 : Unicode (UTF-8)"
            - NOTE: these encoding options are sorted alphabetically, so "Unicode (UTF-8)" appears near the bottom of the list.
          - Click Next
        - Wizard Step 2
          - Select "Comma" as the only delimiter
          - Click Next
        - Wizard Step 3
          - Select "Text" as the "Column data format" (*Unfortunately, this must be done for each column individually in Excel*)
            - At a minimum, you MUST ensure all date columns (e.g. dc.date.issued) are treated as "Text" so that Excel doesn't autoconvert DSpace's YYYY-MM-DD format into MM/DD/YYYY
            - To avoid such autoconversion, it is safest to ensure each column is treated as "Text". Unfortunately, this means selecting each column one-by-one and choosing "Text" as the "Column data format".
          - Click Finish
  - Perform your edits. Once finished, re-upload the changes to DSpace.
    - You can remove entire columns from the spreadsheet to make it easier to concentrate on editing just a few metadata fields. But, the 'id' (first) column MUST be kept.
      - Removing an entire column with **not** delete that metadata (rather DSpace will just ignore it). However, please be careful to remove the ENTIRE column (including the column header). Metadata values are only deleted if you leave the column header in place but clear out one or more values (rows in a column)
    - Some metadata fields may appear duplicated with ISO language tags within the spreadsheet (e.g. "dc.subject" and "dc.subject[en_US]" columns). This is nothing to be concerned about, it simply means that some of your metadata fields specify a specific language and others do not.
      - For example, a "dc.subject" column would include subjects with *no language* specified; whereas, a "dc.subject[en_US]" column would include subjects with USA English specified as the language, and a "dc.subject[es]" column would include subjects with Spanish specified as the language.
      - You are welcome to move values between these columns. Moving a value from "dc.subject" to "dc.subject[en_US]" and saving would update that value to include a language specifier of USA English. Similarly, moving a value from "dc.subject[en_US]" to "dc.subject" would update that value to include *no language specifier*.
    - Many more editing tips are available via the tutorials linked above.
  - Click "Import Metadata" (under "Administrative" menu)
  - Select the CSV
  - Review the changes and save the changes.
    - WARNING: Please make certain that the changes displayed on the Review screen look correct. Once you save, you will be unable to "undo" the changes without either re-editing to metadata (or if you deleted something entirely it may need to
  - *NOTE: A more detailed walkthrough with screenshots of this entire process (and additional hints) is available in the Batch Metadata Editing tutorials linked above*

Bulk/Batch Content Uploads

- DSpace allows you to batch upload content + metadata in a specific Zip package format that DSpace calls the Simple Archive Format.
- How-to: Create an upload package
    - Details about the Simple Archive Format (SAF) are provided in the DSpace documentation at Importing and Exporting Items via Simple Archive Format
    - An SAFBuilder is also available which can transform spreadsheets (Excel or CSV) into valid Simple Archive Format (SAF) packages: https://github.com/DSpace-Labs/SAFBuilder
    - SAFBuilder is a commandline tool, and as such, it requires a few prerequisites to run. These are all listed in the SAFBuilder instructions at: https://github.com/DSpace-Labs/SAFBuilder/blob/master/README.md

      1) Install Java JDK (version 7 or 8). Here's the Java 8 JDK (download a Windows version and install): http://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html
      2) Maven should also be installed. https://maven.apache.org/download.cgi Good guide for installing on Windows at https://www.mkyong.com/maven/how-to-install-maven-in-windows/
      3) Git is optional. You can use it to download SAFBuilder (or just download the Zip of it, as I show below)

      After the above are installed, just download SAFBuilder to any directory, this "installs" it.
- How-to: Upload the SAF package to DSpace
    - Keep in mind that large packages (over several GB in size) may prove difficult to upload via the web. They may timeout during upload or processing. Therefore, with DSpaceDirect, you may wish to consider creating several separate upload packages (and upload them individually) if you have a larger set of content to upload
    - Step by step upload instructions are available in the DSpace documentation at UI Batch Import (XMLUI). (*Note: All DSpaceDirect sites use the DSpace XMLUI user interface*)

- Tips and Common Questions:
    - For any columns that require multiple values (including the filename column), separate the multiple values using a double pipe ||
    - DSpace generates thumbnails for the following formats only: BMP, GIF, JPG, JPEG, JPEG 2000, PNG, TIFF, and PDF. Thumbnails are auto-generated via an overnight process. So, they will not appear until the next day.

DSpace Permissions Overview

DSpace has several permission types that are defined in the system:

- ADD: This permission is available only on Collections or Communities. It gives someone the ability to submit/add Items within a Collection, or create/add Collections under a Community.
- DELETE: This permission gives someone the ability to permanently remove/delete an object. Usually only Administrators should be given this permission.
- READ: This permission gives someone the ability to view an object. For example, READ access on an Item lets someone view the Item's metadata. READ access on a file (bitstream) lets someone view/download that file.
- WRITE: This permission gives someone the ability to edit/change an object. For example, WRITE on a Collection lets someone edit/change that Collection's name or description.
    - Keep in mind, this is different from ADD permission. Someone who has ADD permission on a Collection but does **not** have WRITE permission would only be allowed to submit Items to the Collection. They would not be allowed to edit the Collection name or description.

Individual Item Permission Changes

- DSpace allows you to modify the permissions on any Item at any time. This may be useful if an item needs to be access restricted at a later date, or was accidentally deposited as restricted but needs to be made openly available.
    - NOTE: You can also perform batch/bulk permissions changes (for all Items in a single Collection). More information on that can be found below in the section "Bulk/Batch Permissions Changes (Advanced)"
- **How-to: Change Permissions on a single Item**
    - Login to your site as an Administrator
    - Locate the Item whose permissions you wish to change
    - Click "Edit this Item" in the "Context" side menu
    - On the "Item Status" tab, click "Authorizations". This will bring you to a "Policies" page which details the current permissions on this Item.
    - Overview of the "Policies" page:
        - You'll see a listing of *all* current policies/permissions on the Item, along with all its associated files (called "bitstreams")
            - To modify a policy, click on its "ID".
            - To add a new policy, click the appropriate "Add" link
            - To remove an existing policy, check the checkbox next to the policy, and click the "Delete Selected" button.
        - Types of Policies:
            - *Item Policies* - These policies define the permissions on the Item's **metadata**. In other words, they define who can locate the Item via the DSpace browse & search tools. If the "Item Policies" are limited to a specific Group, then ONLY users who belong to that Group will be able to search/browse to locate that Item.
                - Please note: restricting access to an Item's metadata also means that search engines like Google or Google Scholar will NOT be able to index the Item. That means the Item will NEVER appear in Google or Google Scholar search results. If you want an Item to appear in Google search results, but do not want everyone to be able to download its file(s) then you should leave the Item policies as unrestricted, while restricting access to individual files (Bitstreams). See below for more info.
            - *Bundle Policies* ("Policies for Bundle ___") - These policies define permissions across a "Bundle" (loose grouping) of Files. Bundles in DSpace are mostly "hidden" from users, but they include "ORIGINAL" (the originally uploaded files), "LICENSE" (a copy of any deposit licenses), "TEXT" (optionally, the extracted plain text from originally uploaded files). Generally speaking, the Bundle permissions here should really just be *identical* to the permissions you wish to set on the various Files (Bitstreams).

        - *Bitstream Policies* ("Bitstream __") - These policies define permissions on individual Files (called Bitstreams). The "READ" permission lets users view/download the specific files in question.
    - Action types:
        - "READ" permissions gives a single group the ability to view or download the Item
        - "WRITE" permissions gives a single group the ability to edit or modify the Item
    - Start and End Dates:
        - Optionally, individual policies can have a start or end dates. This allows you to set temporary (or expiring) permissions on an Item (e.g. for embargoes).
        - Start Date: Policies will not take affect until the *start date* has passed
        - End Date: After this date, the specific policy will no longer be in effect.
        - If there is no Start or End date, the policy is in effect *at all times*
    - Some Sample Policies:
        - *An Item Policy that provides "READ" access for "Anonymous" group.* This policy will allow anyone in the world to access the metadata for this specific Item. It means that anyone will be able to search/browse for this Item in DSpace (and search engines like Google or Google Scholar will be able to index this Item, so anyone will also be able to find it within Google)
        - *An Item Policy that provides "READ" access for only the "On Campus Users" group.* This policy will only allow users who belong to the "On Campus Users" group to browse or search for this Item within DSpace. This means that all other users will NOT be able to locate this Item within DSpace. It also means that search engines like Google or Google Scholar will also NOT be able to locate the item (as they will not have access to the Item's metadata).
        - *A Bitstream Policy that provides "READ" access for "Anonymous" group.* This policy will allow anyone in the world to view or download that particular file (bitstream). (Please note however that users will be UNABLE to search/browse for the link to the file unless the Item Policy also provides READ access for the Anonymous group.)
        - *A Bitstream Policy that provides "READ" access for only the "On Campus Users" group.* This policy will only allow users who belong to the "On Campus Users" group to view or download that particular file (bitstream). (Please note however that users will be UNABLE to search/browse for the link to the file unless the Item Policy also provides READ access for either the "On Campus Users" group OR the "Anonymous" group.)
        - POLICY NOTES:
            - Administrators can ALWAYS access/edit all Items in the system, no matter what policies are set on the Item or its individual Bitstreams.
            - Policies are always "additive". This means that if you have two Policies, then *both* permissions are in effect. In other words, a more restrictive policy is *always* overridden by a less restrictive one (e.g. "READ for Anonymous group" will mean that anyone in the world can view/download the Item, even if there is also a secondary policy stating "READ for 'On Campus Users' group")
    - Some Sample Use Cases:
        - You want an Item to be searchable/browseable publicly and its file(s) to be viewable/downloadable to anyone in the world
            - Set ALL policies to "READ" access for "Anonymous" group
        - You want an Item to be searchable/browseable publicly, but its file(s) should ONLY be viewable/downloadable by members of the "On Campus Users" group
            - Set the Item Policy to be "READ" access for "Anonymous" group
            - Set all Bundle/Bitstream Policies to be "READ" access ONLY for the "On Campus Users" group
        - You want an Item to ONLY be searchable/browseable/downloadable by members of the "On Campus Users" group (WARNING: In this scenario, the Item will also NEVER appear in Google or Google Scholar as that particular Item is essentially "invisible" to all users except for those in the "On Campus Users" group)
            - Set ALL policies to "READ" access for ONLY the "On Campus Users" group

Bulk/Batch Permissions Changes (Advanced)

- DSpace provides an "Advanced Policy Manager" (also known as the "item wildcard policy admin tool"), which allows an Administrative user to perform bulk permissions changes to all of the Items or Files (bitstreams) within a specified Collection.
    - For more information on permissions / policy settings in general, please also refer to the section "Individual Item Permissions Changes" above.
- WARNING: This Advanced Policy Manager is a bit of a "beta-level" tool. It works, but it's not the most user friendly page in DSpace. It's also not the smartest tool, so you need to sometimes take several steps to make the changes you want to make.
- **How-To: Batch Permissions Changes basics**
    - Sample Use Case: The easiest way to explain how this tool works is via a common use-case. Suppose that you have a Collection of open access (viewable/readable to anyone in the world) Items which you now want to restrict to only be viewable to a group of users called "On Campus Users". Here's the steps you would take to perform that change:
        - Login to your site as an Administrator
        - Under the "Administrative" side menu, click on "Authorizations" (under Access Control submenu)
        - Just under the box at the top of the page, click the link that says "Click here to go to the item wildcard policy admin tool"
        - *Step 1: Remove existing metadata access rights for all Items in the specified Collection.* To do so, fill out the form as follows:
            - Description: (optional, usually is left blank as it's only really useful for bulk changes to embargo)
            - Group: (leave blank in this case as you will remove any existing permissions)
            - Action: READ (you want to remove "READ" access)
            - Content Type: Item (you want to remove READ access on an Item level – this controls metadata access)
            - Collection: [select the collection]
            - Start Date: (leave blank - this is only useful for bulk-changes to embargo dates)
            - End Date: (leave blank - this is only useful for bulk-changes to embargo dates)
            - CLICK the "Clear Policies" button
            - (NOTE: Even though you get no confirmation screen, the changes will be immediately applied)
        - In Step #1, essentially all we've done is remove access to the Item *metadata*. The metadata is now only visible (readable) by Administrators. However, the content files within those Items are unfortunately still accessible (if someone had bookmarked the URL)
        - *Step 2: Remove existing content file access rights for all Items in the specified Collection.* To do so, fill out the form as follows:

        - Description: (optional, usually is left blank as it's only really useful for bulk changes to embargo)
        - Group: (leave blank in this case as you will remove any existing permissions)
        - Action: READ (you want to remove "READ" access)
        - Content Type: bitstream (you want to remove READ access on the files, or bitstreams)
        - Collection: [select the collection]
        - Start Date: (leave blank - this is only useful for bulk-changes to embargo dates)
        - End Date: (leave blank - this is only useful for bulk-changes to embargo dates)
        - CLICK the "Clear Policies" button
        - (NOTE: Even though you get no confirmation screen, the changes will be immediately applied)
    - In Step #2, we've also removed access to the Item *files*. This means that only Administrators can now access/download any files associated with the Items. Now, we need to assign NEW permissions for our "On Campus Users" group in the following two steps.
    - *Step 3: Give the "On Campus Users" group access to all metadata for all Items in the specified Collection.* To do so, fill out the form as follows:
        - Description: (optional, usually is left blank as it's only really useful for bulk changes to embargo)
        - Group: Select the "On Campus Users" group
        - Action: READ (you want to add "READ" access to the selected group)
        - Content Type: item (you want to add READ access on Items)
        - Collection: [select the collection]
        - Start Date: (leave blank - this is only useful for bulk-changes to embargo dates)
        - End Date: (leave blank - this is only useful for bulk-changes to embargo dates)
        - CLICK the "Add Policies" button
        - (NOTE: Even though you get no confirmation screen, the changes will be immediately applied)
    - In Step #3, we've now given the "On Campus Users" group the ability to read the metadata for all Items in this collection. So, the final step is to also give them the ability to read/download files associated with these Items.
    - *Step 4: Finally, give the "On Campus Users" group access to all files for all Items in the specified Collection. To do so, fill out the form as follows:*
        - Description: (optional, usually is left blank as it's only really useful for bulk changes to embargo)
        - Group: Select the "On Campus Users" group
        - Action: READ (you want to add "READ" access to the selected group)
        - Content Type: bitstream (you want to add READ access on all files, or bitstreams)
        - Collection: [select the collection]
        - Start Date: (leave blank - this is only useful for bulk-changes to embargo dates)
        - End Date: (leave blank - this is only useful for bulk-changes to embargo dates)
        - CLICK the "Add Policies" button
        - (NOTE: Even though you get no confirmation screen, the changes will be immediately applied)
    - At the end of this process, all the Items (and their Files) in the selected Collection will now only be accessible to users who belong to your "On Campus Users" group. Other non-Administrative users will be presented with an Access Restricted message.

Checking Item Links

BasicLinkChecker

DSpace provides a Basic Link Checker as part of the system Curation Tasks. This can be used for small collections as follows:

- Login as a DSpace Administrator
- Under the Administrative menu, select Curation Tasks
- Enter the Handle of the Community, Collection, or Item to check
- Select "Check Links in Metadata" in the Task menu and choose Perform.

As noted above, this will work fine for individual items or small collections. Larger collections will likely take too long to run and will result in an error.

Checking with Exported Metadata

Links can be checked by exporting the site metadata and using an external process to verify links. An example of an external process using Google Sheets follows

- First, export the metadata to be checked
    - Login as a DSpace Administrator
    - Select the Community or Collection to be checked
    - In the Context menu select "Export Metadata" and save the resulting CSV file
- Links can be checked using Google Sheets and a simple script
    - Open Google Drive and drag the CSV file into an appropriate folder (this uploads the file)
    - Right-click on the file and select "Open with > Google Sheets"
    - Find the metadata column with links to be checked (often this is dc.identifier.uri[]). You may choose to hide other columns to simplify the view.
    - Select Tools Script Editor
    - Replace the default script with this code:

```
function getStatusCodes(urlset){
  if('' == urlset) {
    return '';
  }
```

```
  var urls = urlset.split("||");
  var responseCodes = [];

  for (var i=0; i<urls.length; i++){
    var responseCode = getStatusCode(urls[i]);
    responseCodes.push(responseCode);
  }

  return responseCodes.join();
}

function getStatusCode(url) {
  var options = {
    'muteHttpExceptions': true,
  };

  var response = UrlFetchApp.fetch(url.trim(), options);
  return response.getResponseCode();
}
```

        - This code includes two functions. The getStatusCodes function expects an array (list) of URLs to check. The getStatusCode function expects only a single URL. Which of these you use depend on whether the metadata column you need to check has one URL or multiple URLs in each row. If in doubt, use getStatusCodes, as it will work for one or more URLs.
    - Save the script with File Save
        - You may be asked by Google to provide permissions to access your spreadsheet at this point. You will need to grant these permissions.
    - Back on your Google Sheets file, select an empty column that will be used for script results. On the first row with data (usually row 2) add this to the cell, replacing "Y2" with the cell ID where the URL to be checked can be found, then hitting enter. (This is also where you can choose to use the getStatusCode function rather than getStatusCodes.)

```
=getStatusCodes(Y2)
```

    - The result posted in the cell should provide an HTTP response code. The success code is 200. A code of 404 means the page cannot be found. Other response codes are listed here: https://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html. If something else goes wrong with the request you may see an error listed here.
    - Assuming this works properly for the first row, you can apply the function to all rows by either:
        - Selecting the cell where you placed the function, selecting the small box in the bottom right corner of the cell and dragging it down to all other cells
        - Or, selecting the cell where you placed the function, copying it (Edit Copy), then selecting all rows in the column and pasting (Edit Paste). This method works better if the number of rows is large.
    - Once you have a response code (or multiple response codes if there are multiple URLs) in each row you will be able to review the results looking for non-200 codes that may need further investigation.