

Criterios establecidos para la ingesta de datos de BDU

Protocolo de cosecha

El protocolo utilizado es OAI/PMH tanto para repositorios, revistas como para registros bibliográficos.

Se capturan individualmente para identificar los registros y poder armar las colecciones y jerarquías.

Los registros de revistas y repositorios se cosechan en forma individual (un archivo xml por registro)

Los registros de bibliotecas se cosechan y almacenan en forma combinada (múltiples registros en un solo archivo) debido a razones de rendimiento. Los tiempos de cosecha se han ampliado hasta 5 minutos al advertirse problemas en los motores OAI/PMH.

Los registros de bibliotecas se cosechan solo en formato marcxml. Se utiliza una versión modificada de las definiciones para soportar las distintas situaciones:

- ID (001 o 999\$c)
- building, collection, library_str, institution_str y otras que se generan automáticamente desde el sistema BDU.

Jerarquías

Se gestiona todo el sistema mediante la definición de un árbol de jerarquías:

- Nivel 1 - Institución
- Nivel 2 - Repositorios, revistas y bibliotecas.
- Nivel 3 - Sets (grupos de registros definidos en el motor OAI/PMH)

Cada registro mantiene la información necesaria para ser identificado: datos de institución, colección e identificadores numéricos de biblioteca, institución, etc.

Datos de origen del registro

La información obtenida mediante la cosecha no se modifica ni adapta. Los datos se incorporan tal cual son generados por los proveedores de contenido. Se ha modificado el sistema Vufind para mostrar el origen del registro, lo que permite llegar a la institución y el repositorio o biblioteca desde donde se obtiene el registro.

Sistemas capturados

Si bien el protocolo OAI/PMH debería ser suficiente para definir qué es lo que se puede y lo que no se puede capturar se hacen adaptaciones en función del software utilizado.

Actualmente se han capturado registros de los siguientes sistemas:

- KOHA: diversas versiones.
- Dspace: versiones 4.5 a 7
- PMB: (solo registros DC. Actualmente se están solucionando problemas en la versión convert_pmb2xmlmarc)
- Vufind: (registros marcxml y DC, usando sets)
- Greenstone
- Otros softwares.

Exportación para usuarios registrados

Se requiere registro de usuarios para poder exportar y guardar registros .

La exportación se realiza en forma unitaria mediante el enlace que aparece en la ficha de registros marcxml y por la lista desde el perfil de usuario (en formato masivo).

Nombres de las instituciones, bibliotecas y revistas

Se establece:

1- Nombrar a la *institución mayor* en forma completa.

2- Bibliotecas:

- Una sola unidad de información: "Biblioteca" en *singular* + la *sigla de la institución entre paréntesis*.
- Un sistema de bibliotecas: "Bibliotecas" en *plural* + la *sigla de la institución entre paréntesis*.
- Con nombre específico: abreviar para una visualización más acorde.

Ej:

- Universidad de Buenos Aires → Institución mayor
- Bibliotecas (UNCuyo) → Sistema de bibliotecas de la UNCuyo.
- Biblioteca BIBHUMA (FAHCE-UNLP) → Biblioteca de la Fac. de Humanidades y Ciencias de la Educación, UNLP
- Biblioteca H. González (UNGS) → Biblioteca Horacio González
- Biblioteca Cnl. M.J. Olascoaga (IGN) → Biblioteca Coronel José María Olascoaga