ORCID Integration

You are reading the documentation for DSpace-CRIS 5.x and 6.x as available in the maintenance branches.

For DSpace-CRIS 7 please check the PDF technical documentation provided on GitHub as noted here a specific section is devoted to the ORCID Integration

- Introduction
- Authentication
  - Profile claiming
  - Creation of a local researcher profile
- Submission lookup
- Synchronization
  - from ORCID to DSpace-CRIS
  - from DSpace-CRIS to ORCID
    - Mapping local properties to ORCID
- ORCID as a local profile URL

Introduction

ORCID API Version 3

Currently the last line of development DSpace-CRIS 5.10.x use ORCID API v3, see https://members.orcid.org/api/news/xsd-30-update

ORCID API version

Since DSpace-CRIS 5.8.0 we use the ORCID API v2, see https://members.orcid.org/api/news/xsd-20-update

ORCID Metadata mandatory definition

Note that "orcid" metadata in the Researcher Profile entity MUST be a "text" property definition (check https://github.com/4Science/DSpace/blob/dspace-5_x_x-cris/dspace/etc/orcid_from_link_to_text.sql to perform the migration)

About ORCID

more info about ORCID here https://orcid.org/

DSpace-CRIS provides the most advanced and complete integration between a CRIS / Repository system and ORCID, achieving most of the ORCID integration use cases.

The available functionalities depend on the kind of API Key that you have: Public or Member. You must configure your client-id, client secret and specify the endpoint appropriate for your API level in the

[dspace-installDir]/config/modules/authentication-oauth.cfg (via build.properties, that contains configuration like below)

```
#####################
# ORCID
#####################
#Production Registry
#Step        |     Member API                                |    Public API
#-------------------------------------------------------------------------------------------------------
#Authorize   |    https://orcid.org/oauth/authorize          |    https://orcid.org/oauth/authorize
#Exchange    |    https://orcid.org/oauth/token              |    https://orcid.org/oauth/token
#Use         |    https://api.orcid.org/v3.0                 |    https://pub.orcid.org/v3.0
#-------------------------------------------------------------------------------------------------------
#
#Sandbox
#Step        |     Member API                                |    Public API
#-------------------------------------------------------------------------------------------------------
#Authorize   |     https://sandbox.orcid.org/oauth/authorize |    https://sandbox.orcid.org/oauth/authorize
#Exchange    |     https://sandbox.orcid.org/oauth/token     |    https://sandbox.orcid.org/oauth/token
#Use         |     https://api.sandbox.orcid.org/v3.o        |    https://pub.sandbox.orcid.org/v3.0
#-------------------------------------------------------------------------------------------------------
authentication-oauth.orcid-api-url = https://pub.orcid.org/v3.0
authentication-oauth.application-token-url = https://orcid.org/oauth/token
authentication-oauth.application-authorize-url = https://orcid.org/oauth/authorize

# register for free on ORCID to use an institutional Public API
# IMPORTANT!! Please fill authentication-oauth.application-client-name with name of client registered into
orcid registries (need by the putcode flow retrieve)
authentication-oauth.application-client-name =
authentication-oauth.application-client-id =
authentication-oauth.application-client-secret =
authentication-oauth.application-client-redirect = ${dspace.baseUrl}/oauth-login

# If you have only PUBLIC API the scope need to be; /authenticate scope now includes /read-public scope
authentication-oauth.application-client-scope =/authenticate
# if you have MEMBER API the suggested scopes are as following; /authenticate scope now includes /read-public
scope
#authentication-oauth.application-client-scope =/authenticate /read-limited /person/update /activities/update
```

Please note that the properties in this file are used also for the other ORCID integration functionalities. So you need to make sure that they are accurate, even if you decide to disable the ORCID authentication.

Authentication

The ORCID authentication is enabled by default in DSpace-CRIS, nevertheless you need to configure the key properties listed in the above introduction.

The following properties are specific of the authentication integration:

```
autoregister=true
choice-page=false
orcid-embedded-login=true
```

The *autoregister* flag is self explanitory. *choice-page*, if true, lists the ORCID authentication explicitly as an authentication method, so that, if you have multiple authentication methods configured, such as LDAP, Shibboleth, etc., ORCID is also listed in that choice page. The *orcid-embedded-login* is the preferred mode, a Login with ORCID button is displayed side-by-side with password or ldap authentication.

Profile claiming

When the ORCID Authentication is enabled and the autoregister flag is true. it is possible for a researcher to claim an existent researcher profile that has an ORICD in DSpace-CRIS just logging in. Indeed, when the user logs in, the system will create a user account using the information from ORCID and, if his ORCID match one already assigned to an orphan researcher profile, this new account become the owner of the profile. In such way, external collaborators can login in the system and gain limited privileges to improve their profile without any explicit activity by the repository staff.

Creation of a local researcher profile

When the user is logged in via ORCID and proceeds to create his local researcher profile, the system automatically captures from ORCID several information:

Submission lookup

A special authority is build-in DSpace-CRIS to use the ORCID registry togheter with the internal directory of researchers as authority list for metadata editing

More information on the authority framework can be found in the official DSpace documentation, the relevant configuration are in the dspace.cfg file as listed below

```
plugin.named.org.dspace.content.authority.ChoiceAuthority = \
 org.dspace.app.cris.integration.ORCIDAuthority = RPAuthority,\
 ...

choices.plugin.dc.contributor.author = RPAuthority
choices.presentation.dc.contributor.author = lookup
authority.controlled.dc.contributor.author = true
...
```

For this functionality the use of a member API key or a public API key impact only on the performance and acceptable rate of request to the ORCID service. As you are looking for ORCID profile of unknown researchers only the public information about such researcher are returned

The org.dspace.app.cris.integration.ORCIDAuthority combines the result retrieved from the internal researchers database with the ones found in ORCID giving precedence to the internal matches. When an author name is selected from the ORCID registry returned list a temporary authority key tracking the ORCID is assigned to the metadata. This information is used when the publication is approved to generate a local researcher profile retrieving further information about the researcher from the ORCID registry such as his biography, email address, name variants, etc more information on this process are available here

Synchronization

DSpace-CRIS is able to capture (Public or Member API) and send information (Member API only) from and to ORCID.

from ORCID to DSpace-CRIS

The dataflow from ORCID to DSpace-CRIS happen in two scenarios:

1. when an user that has been authenticated via ORCID create his researcher profile
2. when an author of a publication is looked up in the ORCID registry during the submission

The scenario #1 has been discussed in the Authentication > Creating a local researcher profile section

The scenario #2 has been discussed in the Submission lookup section.

Using the code available in the dspace-5_x_x -cris branch is also possible to import publication from an ORCID profile in DSpace-CRIS. An ORCIDDataloaderProvider for BTE is included that allow to input an ORCID in the first submission screen (Search by Identifiers) and importing one or all the publications. The dataloader will use the structural information available in the ORCID Work plus the citation field if supplied in a format understandable by another BTE provider (such as BibTex, EndNote, etc.)

The ORCID Online dataloader provider need to be enabled in the bte.xml file

Currently it is not possible to import from ORCID information related to other entities like:

- projects
- organisations

from DSpace-CRIS to ORCID

Previously, v1.2 allowed members to update a researcher's given name, family name, or biography. This had the potential to overwrite the researcher's preferred name, or how they chose to describe themselves in the system, or even lock them out from editing that data.The v2.0 no longer has edit access to any of these fields. The logic behind this is that the user has chosen to add these fields themselves, and no party can edit data added by another party.

If the researcher grants permission to DSpace-CRIS to update his ORCID profile, this is done during the authentication via ORCID or directly from the ORCID tab in the researcher profile, DSpace-CRIS is able to send these local information to ORCID (below the shortname of the property as in the default configuration):

- affiliation
- authorid
- education
- country
- keywords

- personalsite
- scopusid
- variants

the researcher can set his own preferences about the synchronization both about which information needs to be send than how/when this happen (automatically or manually). All the preferences are stored as property in the researcher profile.

To configure which information are sent to ORCID the researcher needs to edit his profile, in the default configuration a tab named ORCID (shortname:"orcid") with a box named ORCID Synchronization settings (shortname:"*orcidsyncsettings*") is included that allow to

- define which publications to send: all, only publications selected, only publications not hidden (see), none
- define which projects to send: all, only projects selected, only projects not hidden (see), none
- define which biographic information will be send

The synchronization procedures use a convention to discover which preference are available, see Mapping local properties to ORCID below.

The system monitors changes to the researcher profile and linked entities (publications and projects) putting in a queue (the orcid_queue table on the database) the list of objects that it is necessary to synchronize. This is done for the DSpace items via a consumer enabled by default in the dspace.cfg

```
event.dispatcher.default.consumers = versioning, discovery, eperson, harvester, orcidpush, crisorcid,
itemauthority, dedup
...
# orcid push consumer
event.consumer.orcidpush.class = org.dspace.app.cris.integration.authority.CrisOrcidQueueConsumer
event.consumer.orcidpush.filters = Item+INSTALL|MODIFY_METADATA|MODIFY
```

and for the CRIS entites with a model listner org.dspace.app.cris.model.listener.OrcidQueueListener enabled via spring

```
<bean id="ORCIDListener" class="org.dspace.app.cris.model.listener.OrcidQueueListener">
        <property name="orcidPreferencesUtils" ref="orcidPreferencesUtils" />
</bean>
```

The synchronization can happen automatically via a batch script that need to be configured as a CRON Job without parameters

```
./dspace dsrun org.dspace.app.cris.batch.ScriptPushOrcid
```

or triggered manually by the researcher. A custom box ORCID Registry Queue ("*orcidsyncqueue"*) is included in the default configuration to show the queue of object to synchronize and allow manual processing.

If a researcher set the synchronization mode to "manual" the normal execution of the synchronization script cause to send a reminder email to the researcher with pending synchronization actions (email template: orcid-alerts)

The script can be also used from the command line to force a first push of information to ORCID indipendently by the content of the ORCID queue table

```
./dspace dsrun org.dspace.app.cris.batch.ScriptPushOrcid [-a | -s <crisID>] [-p]
```

-a process all the researcher

-s process a single researcher, whom with the specified crisID

-p force the script to use POST call to ORCID instead of PUT (put is more efficient but replace all the existent publications, projects, organisations in the ORCID profile created from DSpace-CRIS with the new value. POST will just append the new values). Please note that information input manually by the researcher in his ORCID profile or created in his profile by other system (like crossref, scopus, etc.) cannot be never overridden by DSpace-CRIS

Note that DSpace-CRIS will perform an operation on the Orcid Queue when

1. add/remove/change metadata at first level
2. add/remove/save metadata at nested level (like affiliation and education)

Point 1. and 2. will perform a request of push for the Researcher Profile on Orcid Queue. In particular the second point means that modification on value at metadata "nested" object (like modification on "Role" of the "affiliation" nested object) not perform any operation on Orcid Queue but the "Save" event will do it. So the operation of open a nested object without change it perform an ADD operation on the Orcid Queue. By the way on the metadata first level (like authorid, country, keywords, personalsite, scopusid, variants) any ADD/REMOVE/CHANGE events occuring the system will perform a push of the Researcher Profile that need to perform a synchronization.

Mapping local properties to ORCID

The shortname of all the ORCID preference flags start with orcid-profile-pref- and end with the shortname of the property that need to be syncronized with ORCID. The label of the preference property must match the name of the target Element in the ORCID XML message in this way you can easily set a custom mapping for your extra or renamed properties. The shortname start with "system-" prefix is used to build request document to send to ORCID but due to the v2 limitation is not possible synchronized its values.

The default configuration is

Preference shortname	DSpace-CRIS property	ORCID property
orcid-profile-pref-affiliation	affiliation	affiliations-employment
orcid-profile-pref-authorid	authorid	external-identifier-Researcher ID
system-orcid-profile-pref-biography	biography	biography
orcid-profile-pref-education	education	affiliations-education
system-orcid-profile-pref-email	email	primary-email
system-orcid-profile-pref-fullName	fullName	name
orcid-profile-pref-iso-country	iso-country	iso-3166-country
orcid-profile-pref-keywords	keywords	keywords
system-orcid-profile-pref-otheremails	otheremails	other-emails
orcid-profile-pref-personalsite	personalsite	researcher-urls
system-orcid-profile-pref-preferredName	preferredName	credit-name
orcid-profile-pref-scopusid	scopusid	external-identifier-Scopus Author ID
orcid-profile-pref-variants	variants	other-names

ORCID as a local profile URL

Each property containing unique values can be used to address an researcher profile in DSpace-CRIS with a more meaningful URL than using the internal stable CRIS ID. This functionality is very useful in conjuction with ORCID so that

```
[dspace.url]/cris/rp/details.htm?lt=orcid&lv=[ORCiD ID]
```

is automatically resolved to the researcher profile that old this specific ORCiD ID. With such functionality it is easy to link existent systems to DSpace-CRIS without that them need to know anything about DSpace-CRIS, its internal database structure and IDs. Also staff number or other identifier can be used in the same way changing the lt parameter (lookup term) to match the shortname of the property holding the external ID