SPRINT Irish National Open Access Monitor

IVANA KONČIĆ
Ruđer Bošković Institute
Croatia
ikoncic@irb.hr

MARTINA ŽUGAJ
Ruđer Bošković Institute
Croatia
mzugaj@irb.hr

Meet the Curation Board

Ivana Končić

ikoncic@irb.hr

Head of the Department for Maintenance and Development of CroRIS

Ruđer Bošković Institute

Croatia

Bojan Macan

bmacan@irb.hr

Head of the Centre for Scientific Information

Ruđer Bošković Institute

Croatia

Martina Žugaj

mzugaj@irb.hr

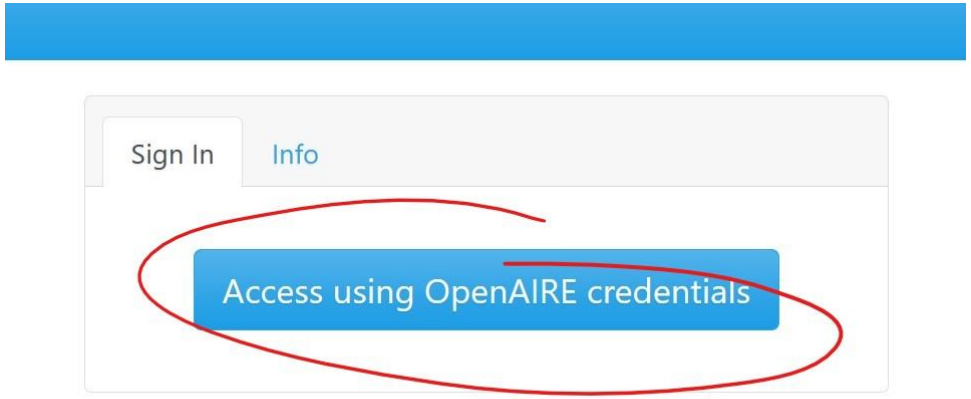
Graduate Librarian at the Centre for Scientific Information

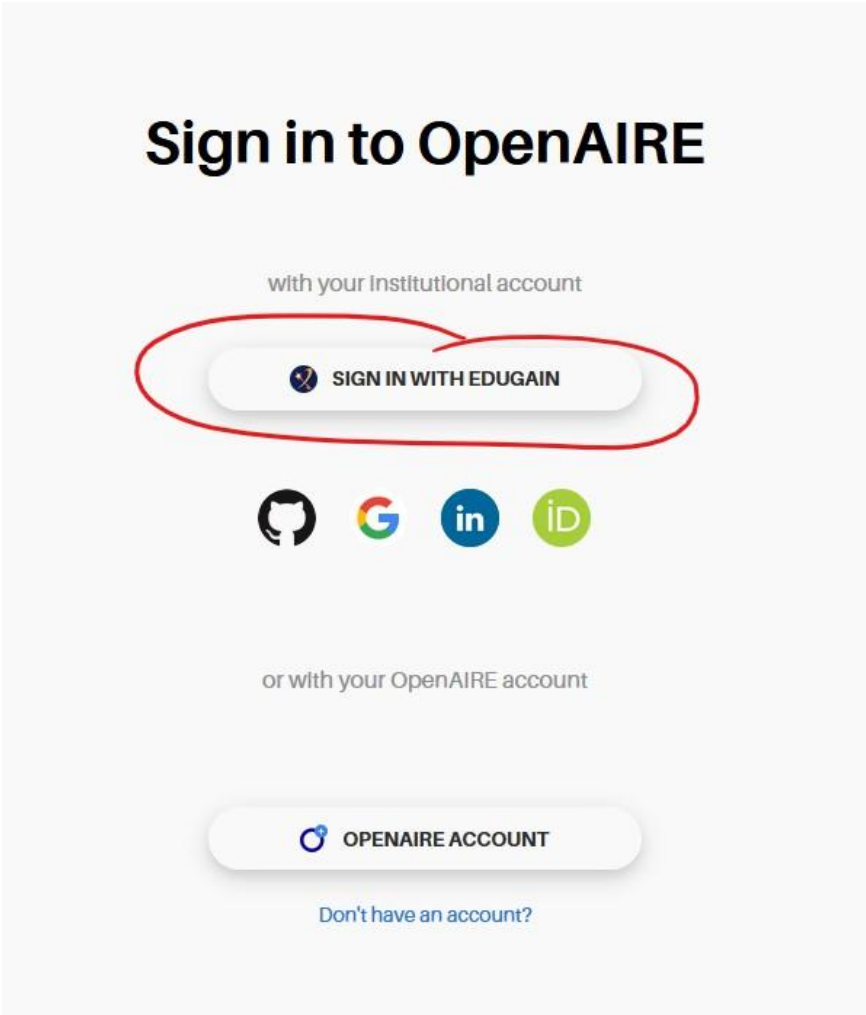
Ruđer Bošković Institute

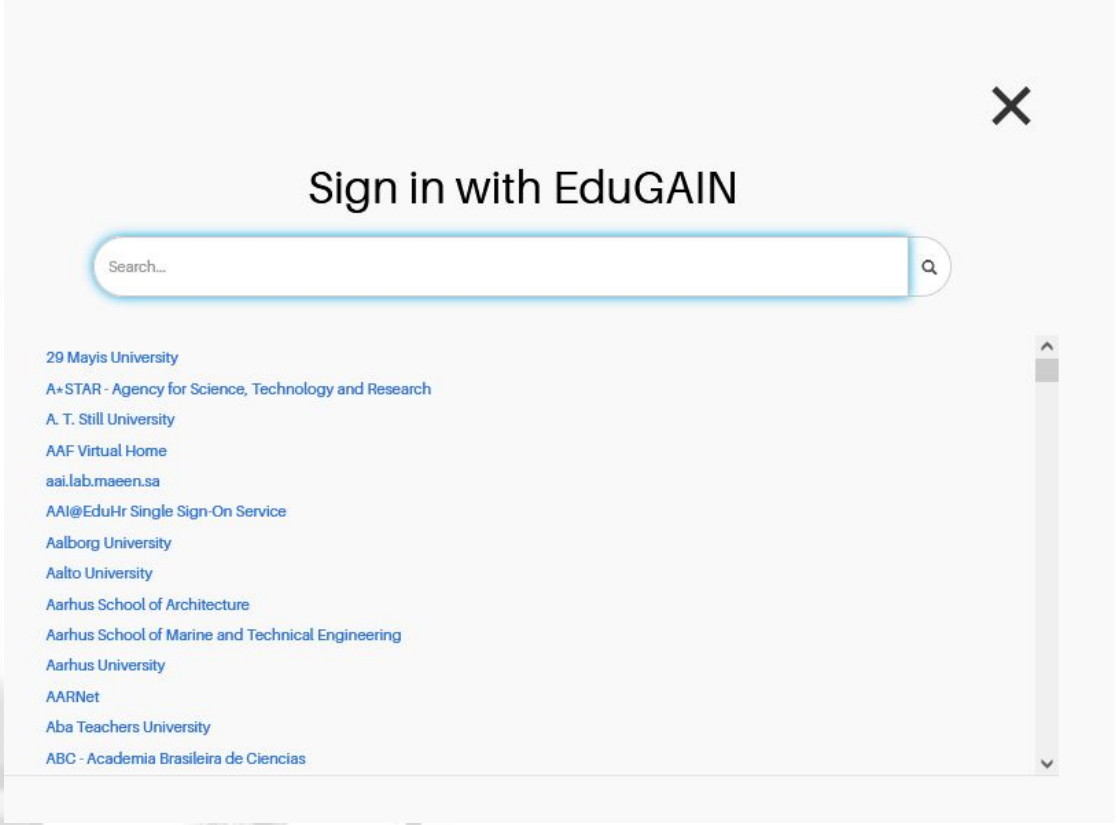
Croatia

Register as a curator

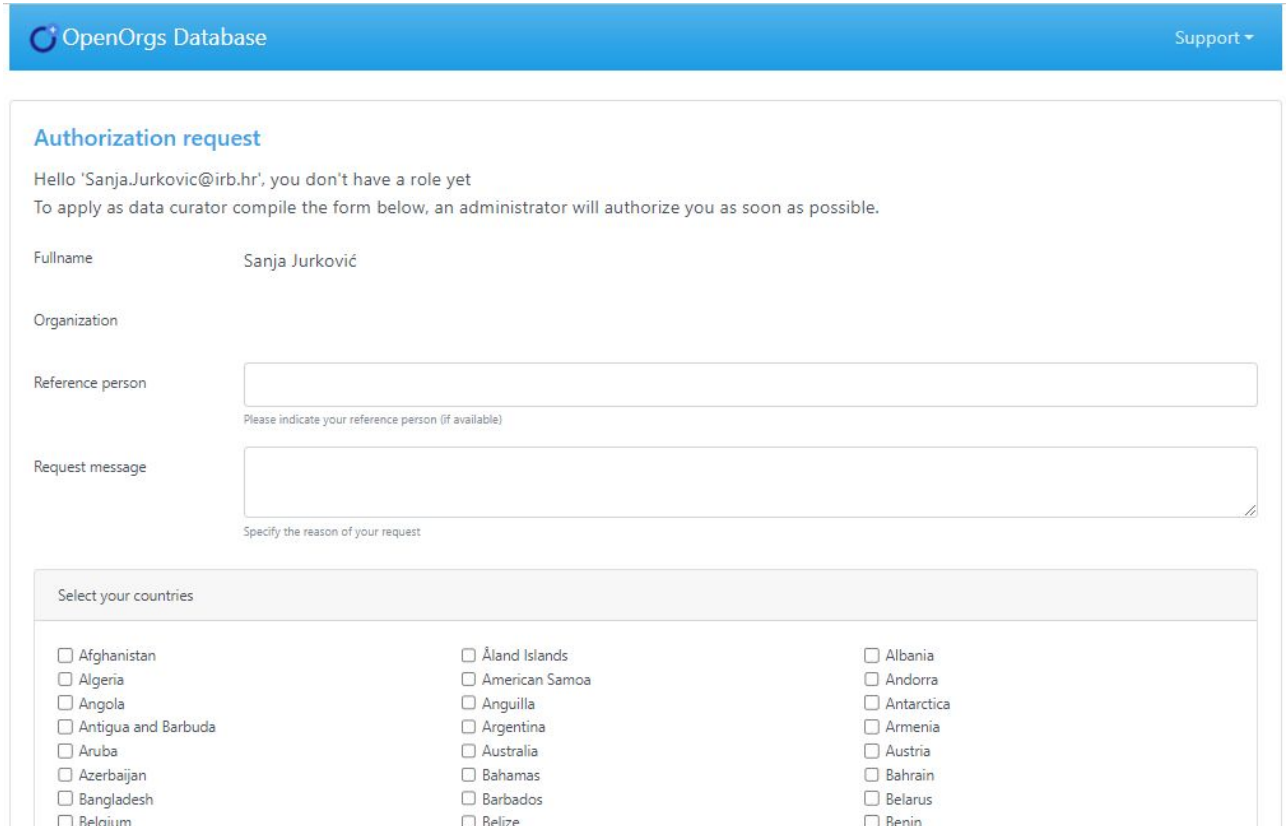
1. Sign in to the OpenOrgs platform at https://orgs.openaire.eu using your **institutional** account and request access. Choose “Ireland” in the form.
2. Administrators will grant you access to the platform (**openorgs-admin@openaire.eu**)







choose the service from the list and sign in



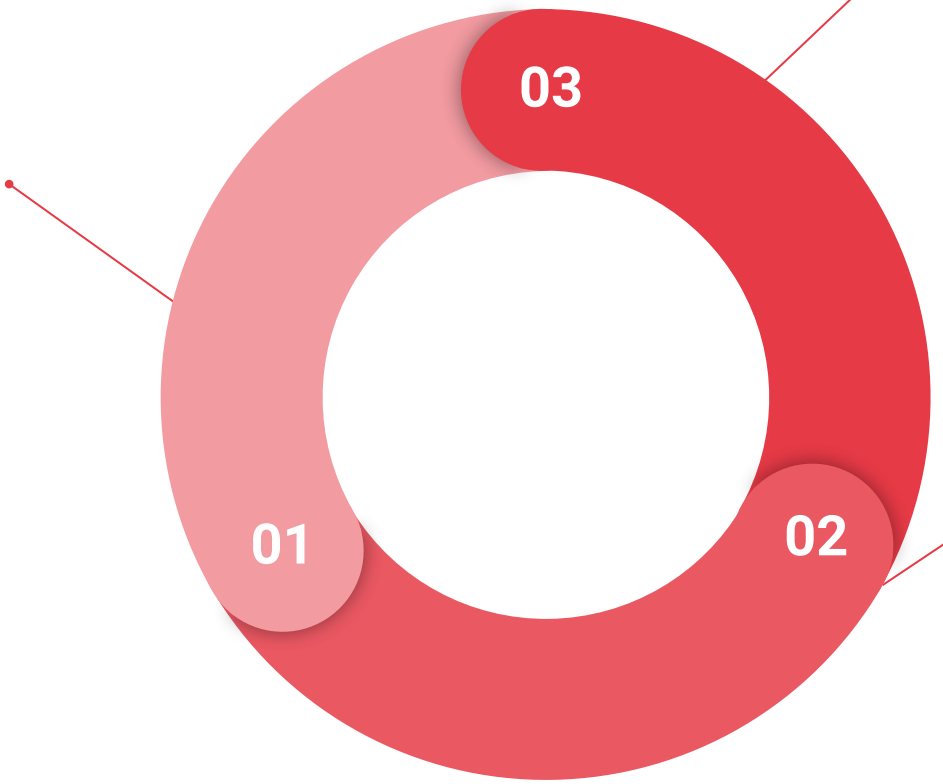
select "Ireland" and click "send request"

OpenOrgs curation tasks

OpenOrgs activity pillars

Automatic suggestion of duplicates

The process for the automatic identification of possible duplicates will periodically produce new suggestions for the data curators.



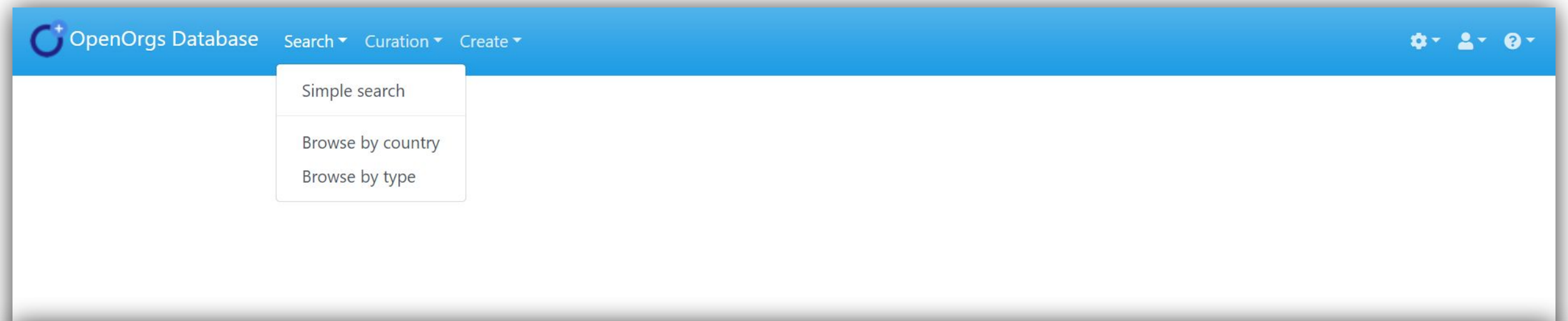
Metadata curation

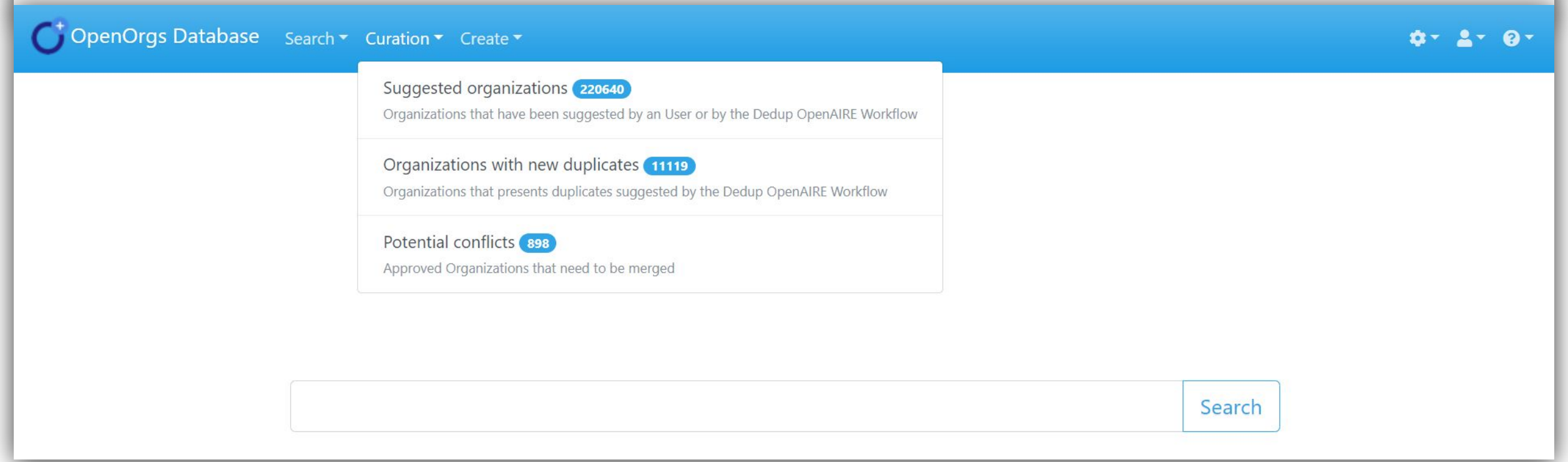
In OpenOrgs data curators can enrich the metadata description of the organization entities, compensating the lack of information available from the sources and improving the discoverability / completeness of the organization records.

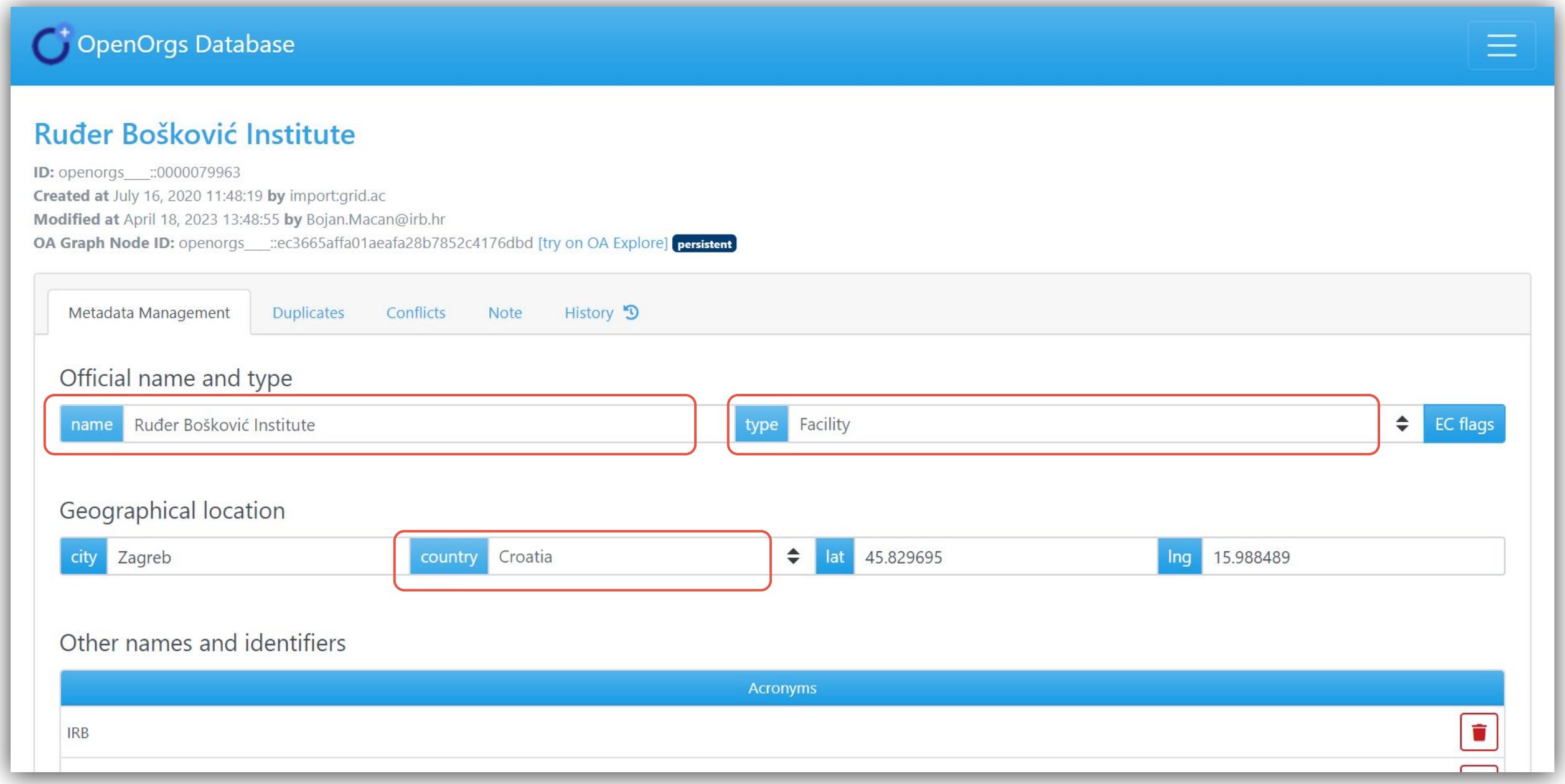
Manual management of duplicates

Deciding whether two or more digital representations are duplicates of the same object can be a task that only humans can carry out precisely. The OpenOrgs tool allows data curator to manage the ambiguities in the data lack of the organization mentions aggregated by OpenAIRE.

- **Curating** and **enriching metadata** for organisations.
- Identifying and **resolving duplicates**.
- Adding **parent-child relations** (if any).
- **Approving organisations** and providing them with stable **OpenOrgs IDs**.
- **Adding new** organisations to the OpenOrgs database.
- Documenting **insights** in the notes section.
- **Resolving conflicts**; merging and distinguishing.


OpenOrgs Database
Search
Curation
Create
Simple search
Browse by country
Browse by type


OpenOrgs Database
Search
Curation
Create
Suggested organizations 220640
Organizations that have been suggested by an User or by the Dedup OpenAIRE Workflow
Organizations with new duplicates 11119
Organizations that presents duplicates suggested by the Dedup OpenAIRE Workflow
Potential conflicts 898
Approved Organizations that need to be merged
Search


OpenOrgs Database
Ruđer Bošković Institute
ID: openorgs___::0000079963
Created at July 16, 2020 11:48:19 by import:grid.ac
Modified at April 18, 2023 13:48:55 by Bojan.Macan@irb.hr
OA Graph Node ID: openorgs___::ec3665affa01aeafa28b7852c4176dbd [try on OA Explore] persistent
Metadata Management
Duplicates
Conflicts
Note
History
Official name and type
name
Ruđer Bošković Institute
type
Facility
EC flags
Geographical location
city
Zagreb
country
Croatia
lat
45.829695
lng
15.988489
Other names and identifiers
Acronyms
IRB

Other names and identifiers

Acronyms

IRB
RBI
new acronym...

Aliases

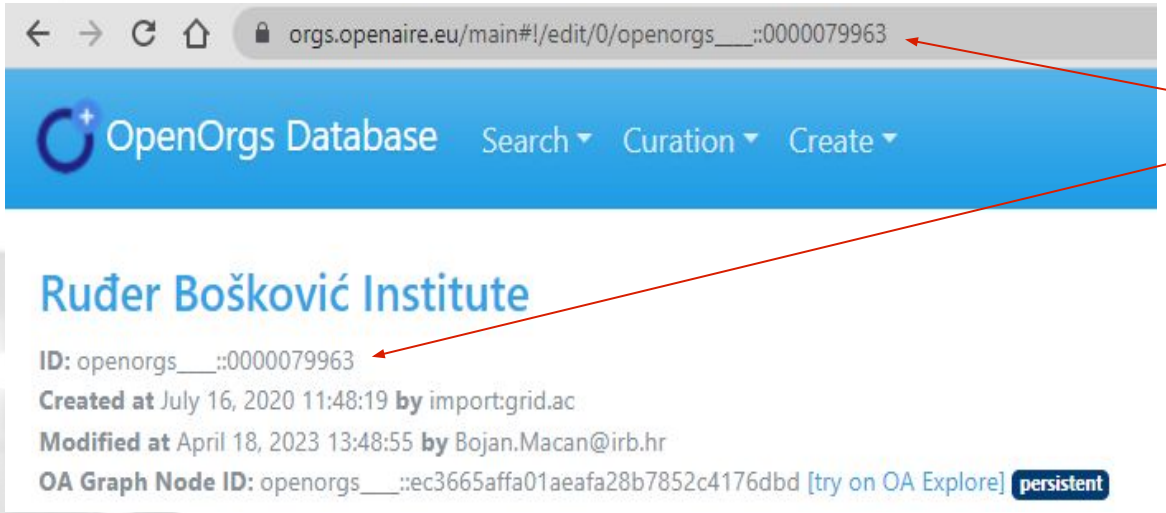
name	language
Institut Ruđer Bošković	hr
Rudjer Boskovic Institute	en
new alias...	language...

Identifiers

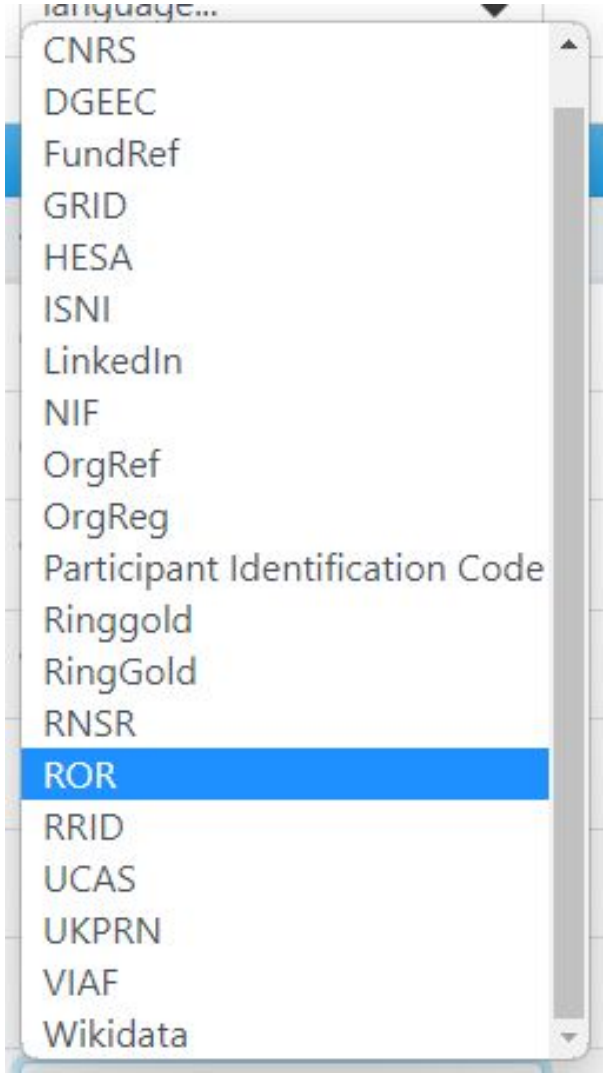
id	type
4763329	OrgRef
grid.4905.8	GRID
143735567	VIAF
Q7383690	Wikidata
501100007259	FundRef
999875031	PIC
0000 0004 0635 7705	ISNI
new id...	type...

Urls

http://www.irb.hr/
http://...



OpenOrgs ID

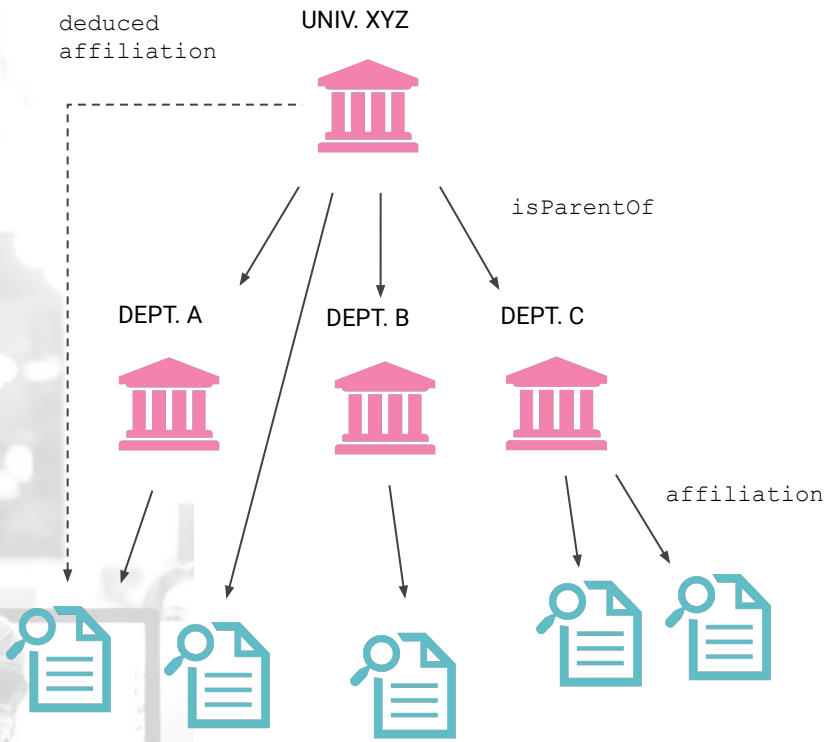


Where to find org's IDs?

- orgs official websites, documents, etc.
- other databases
  - **ISNI** https://isni.org/page/search-database/
  - **ROR** https://ror.org/
  - **PIC** https://bit.ly/41kE71Y
  - **VIAF** https://viaf.org/

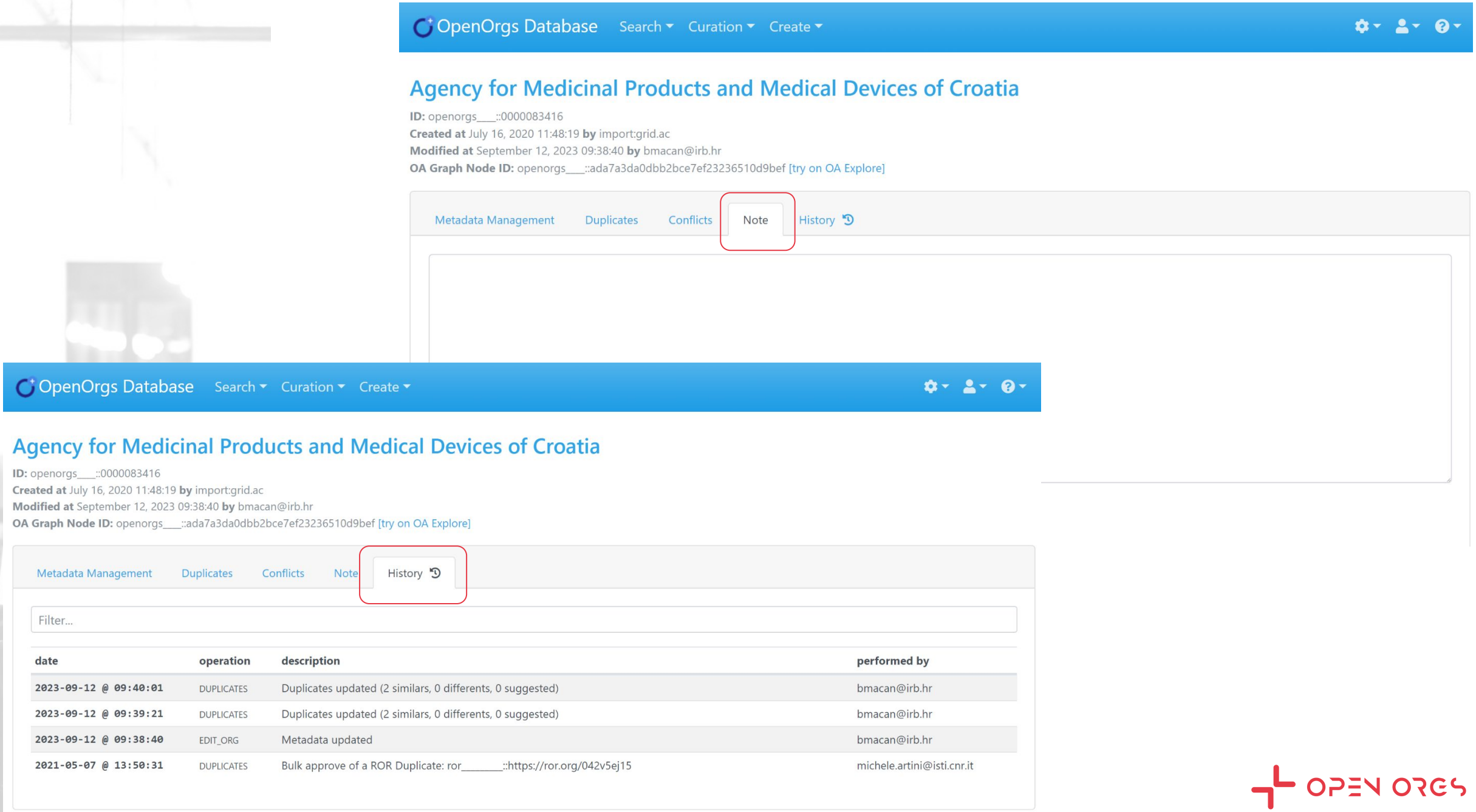
Establishing relationships between organisations

- Only parent-child relationships ('**is parent of**' or '**is child of**').
- **Child production** is included in **parent production**.
- Important for monitoring production, particularly between universities, departments, and institutes.
- To link departments or institutes within an org, they must first be **approved as an organisation** and then linked as a child to the main org.
- Consider naming children orgs with a specific rule (e.g. University of Zagreb, Faculty of Law).
- One org can have **more than one** parent (for "hybrid" cases).



Relations

Relations
This organizazion **is parent of** University of Split, Arts Academy
This organizazion **is parent of** University of Split, Catholic Faculty of Theology
This organizazion **is parent of** University of Split, Faculty of Chemistry and Technology
This organizazion **is parent of** University of Split, Faculty of Civil Engineering, Architecture and Geodesy
This organizazion **is parent of** University of Split, Faculty of Economics, Business and Tourism
This organizazion **is parent of** University of Split, Faculty of Electrical Engineering, Mechanical Engineering an
This organizazion **is parent of** University of Split, Faculty of Humanities and Social Sciences
This organizazion **is parent of** University of Split, Faculty of Kinesiology
This organizazion **is parent of** University of Split, Faculty of Maritime studies
This organizazion **is parent of** University of Split, Faculty of Science
This organizazion **is parent of** University of Split, School of Medicine
This organization rel type... related organization...


OpenOrgs Database
Search
Curation
Create
Agency for Medicinal Products and Medical Devices of Croatia
ID: openorgs___::0000083416
Created at July 16, 2020 11:48:19 by import:grid.ac
Modified at September 12, 2023 09:38:40 by bmacan@irb.hr
OA Graph Node ID: openorgs___::ada7a3da0dbb2bce7ef23236510d9bef [try on OA Explore]
Metadata Management
Duplicates
Conflicts
Note
History
OpenOrgs Database
Search
Curation
Create
Agency for Medicinal Products and Medical Devices of Croatia
ID: openorgs___::0000083416
Created at July 16, 2020 11:48:19 by import:grid.ac
Modified at September 12, 2023 09:38:40 by bmacan@irb.hr
OA Graph Node ID: openorgs___::ada7a3da0dbb2bce7ef23236510d9bef [try on OA Explore]
Metadata Management
Duplicates
Conflicts
Note
History
Filter...
date
operation
description
performed by
2023-09-12 @ 09:40:01
DUPLICATES
Duplicates updated (2 similars, 0 differents, 0 suggested)
bmacan@irb.hr
2023-09-12 @ 09:39:21
DUPLICATES
Duplicates updated (2 similars, 0 differents, 0 suggested)
bmacan@irb.hr
2023-09-12 @ 09:38:40
EDIT_ORG
Metadata updated
bmacan@irb.hr
2021-05-07 @ 13:50:31
DUPLICATES
Bulk approve of a ROR Duplicate: ror_____::https://ror.org/042v5ej15
michele.artini@isti.cnr.it

Adding new orgs (if it is not present already)

- Add it by clicking on **create → a new org**.

- There are three mandatory metadata: **name**, **country** and **type**.

Recommendations:

- For the org's name, it is recommended to use the English version as the main name and include the name in the national language as an alias.
- Always try to include the organization's URL or check that it is correct and use UTF-8 encoding.

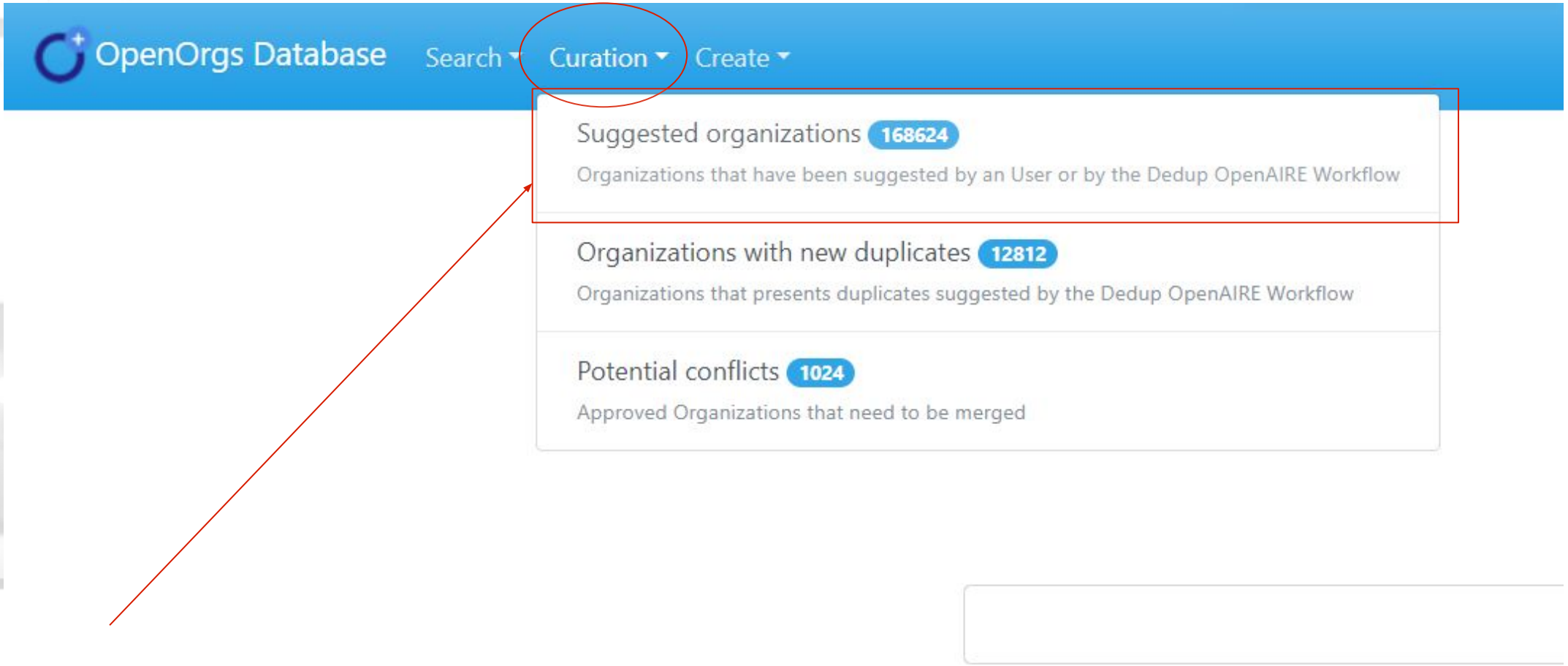
Tips:

- To determine the organization's type, check the EC flags.
- Notes can be added within an orgs page to keep track of choices made.

Approval of new orgs

Main concepts:

- **Suggested** org: is an organization to be approved by a data curator; the approval produces an OpenOrgs ID.
- **Approved** org: is an organization that was confirmed by a curator. It is persisted with a stable identifier (OpenOrgs ID) and its metadata fields can be enriched/curated.


OpenOrgs Database
Search
Curation
Create
Suggested organizations 168624
Organizations that have been suggested by an User or by the Dedup OpenAIRE Workflow
Organizations with new duplicates 12812
Organizations that presents duplicates suggested by the Dedup OpenAIRE Workflow
Potential conflicts 1024
Approved Organizations that need to be merged

OpenOrgs Database Search Curation Create

Suggested Organizations

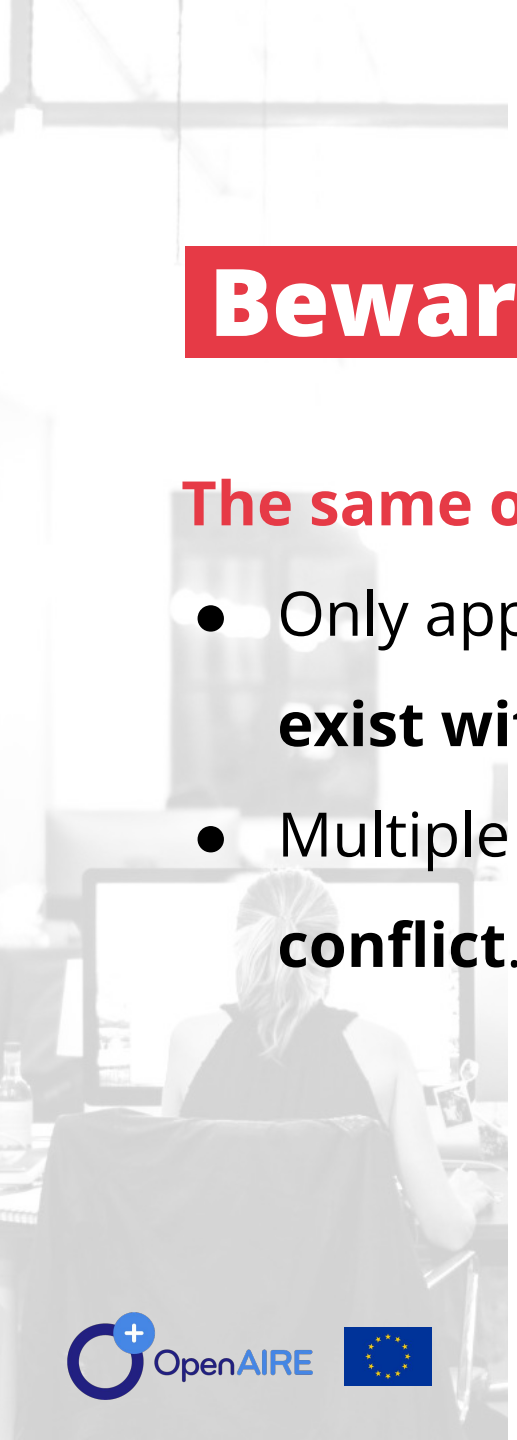
Current country: HR

download as CSV

den

Organization name	Place	Acronyms	Type
Dentitio	-, HR	Dentitio	UNKNOWN
HEP - ESCO DOO ZA VODENJE I FINANCIRANJE PROJEKATA ENERGETSKE UCINKOVITOSTI	-, HR	HEP ESCO	UNKNOWN
HRVATSKE AUTOCESTE DOO ZA UPRAVLJENJE, GRADENJE I ODRZAVANJE AUTOCESTA*CROATIAN MOTORWAYS	-, HR	HRVATSKE AUTOCESTE DOO	UNKNOWN
MINISTARSTVO PROSTORNOGA UREDENJA, GRADITELJSTVA I DRZAVNE IMOVINE	-, HR	MINISTRY OF PHYSICAL PLANNING, CONSTRUCTION AND STATE ASSETS	UNKNOWN
SMART ISLAND KRK DRUSTVO S OGRANICENOM ODGOVORNOSCU ZA PROJEKTIRANJE,GRADENJE I USLUGE	-, HR		UNKNOWN
University of Rijeka, Faculty of Dental Medicine	-, HR		UNKNOWN
ZAVOD ZA PROSTORNO UREDENJE GRADA ZAGREBA	-, HR	BUREAU FOR PHYSICAL PLANNINGCITY OF ZAGREB	UNKNOWN
ZAVOD ZA PROSTORNO UREDENJE KOPRIVNICKO-KRIZEVACKE ZUPANIJE	-, HR		UNKNOWN

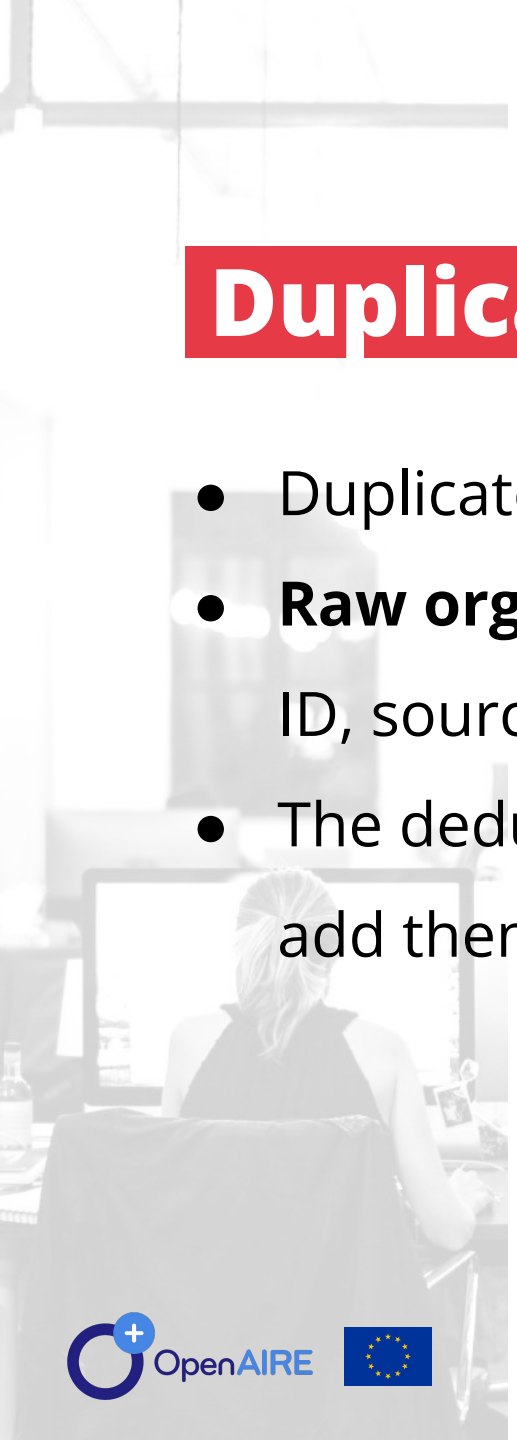
In this case also a relation must be established!



Beware!

The same org must not have more than one OpenOrgs ID!

- Only approve the organization after checking that it **does not already exist with the OpenOrg ID**.
- Multiple OpenOrg IDs for the same organization will **result in a conflict**.



Duplicates curation

- Duplicates can occur between **approved** and **raw** organisations.
- **Raw organisations** are external organizations without an OpenOrgs ID, sourced from platforms such as Corda, Re3data, and ROR.
- The deduplication **algorithm** may suggest duplicates, or users may add them **manually** within an organization's page.

Duplicates curation

Complete the automated part

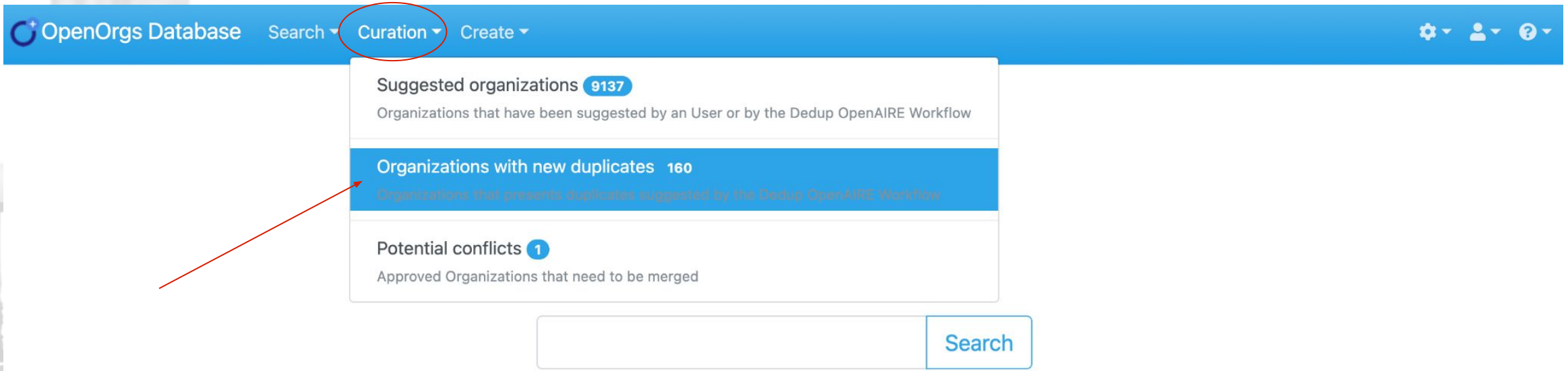
- **Duplicates section** → lists organisations suggested by an algorithm to be duplicates.
- It is an automated process that requires **confirmation** or **rejection** by the data curator.

Manual entry

- **Duplicates addition in the single org page** → The data curator manually adds duplicates in the main org page.

Potential new duplicates section

Curation menu → Organizations with new duplicates



University of Zadar

ID: openorgs____::0000016041
Created at July 16, 2020 11:48:19 **by** import:grid.ac
Modified at July 16, 2020 11:48:19 **by** import:grid.ac
OA Graph Node ID: openorgs____::5e91b35f34a6d4ead6432ef2f77d24bf [try on OA Explore]

Metadata Management | Duplicates | Conflicts | Note | History

Current organization

Name	University of Zadar
Type	Education
Place	Zadar, HR
Acronyms	
Also known as	Sveučilište u Zadru Sveučilište u Zadru University of Zadar Universitas Studiorum Iadertina Свеучилиште у Задру
Urls	http://www.unizd.hr/
Other identifiers	3070362 **(OrgRef)** grid.424739.f **(GRID)** Q2496233 **(Wikidata)** 0000 0001 2159 1688 **(ISNI)** https://ror.org/00t89vb53 **(ROR)**

Duplicates

Related organization	Acronym	Country	Source
University of Zadar **PID (PIC):** 991684060 **URL:** http://www.unizd.hr legal body · legal person · non profit · research organization · higher education	UNIZD	HR	**Original Id:** corda_____::991684060 **OA Graph Node ID:** corda_____::4cec14029b16076cc6534f34d5e52d0c [try] **Provenance:** CORDA - COmmon Research DAta Warehouse
SVEUCILISTE U ZADRU **PID (PIC):** 991684060 **URL:** http://www.unizd.hr legal body · legal person · non profit · research organization · higher education	UNIZD	HR	**Original Id:** corda__h2020::991684060 **OA Graph Node ID:** corda__h2020::4cec14029b16076cc6534f34d5e52d0c [try] **Provenance:** CORDA - COmmon Research DAta Warehouse - Horizon 2020

An org with wrong suggested duplicates

University of Zagreb, Faculty of Agriculture

ID: openorgs____::0000101403
Created at October 6, 2022 11:30:46 **by** Martina.Zugaj@irb.hr
Modified at October 6, 2022 11:30:46 **by** Martina.Zugaj@irb.hr
OA Graph Node ID: openorgs____::c55be4e1c4433a491bcce8e71df78ed4 [try on OA Explore]

Metadata Management | Duplicates | Conflicts | Note | History

Current organization

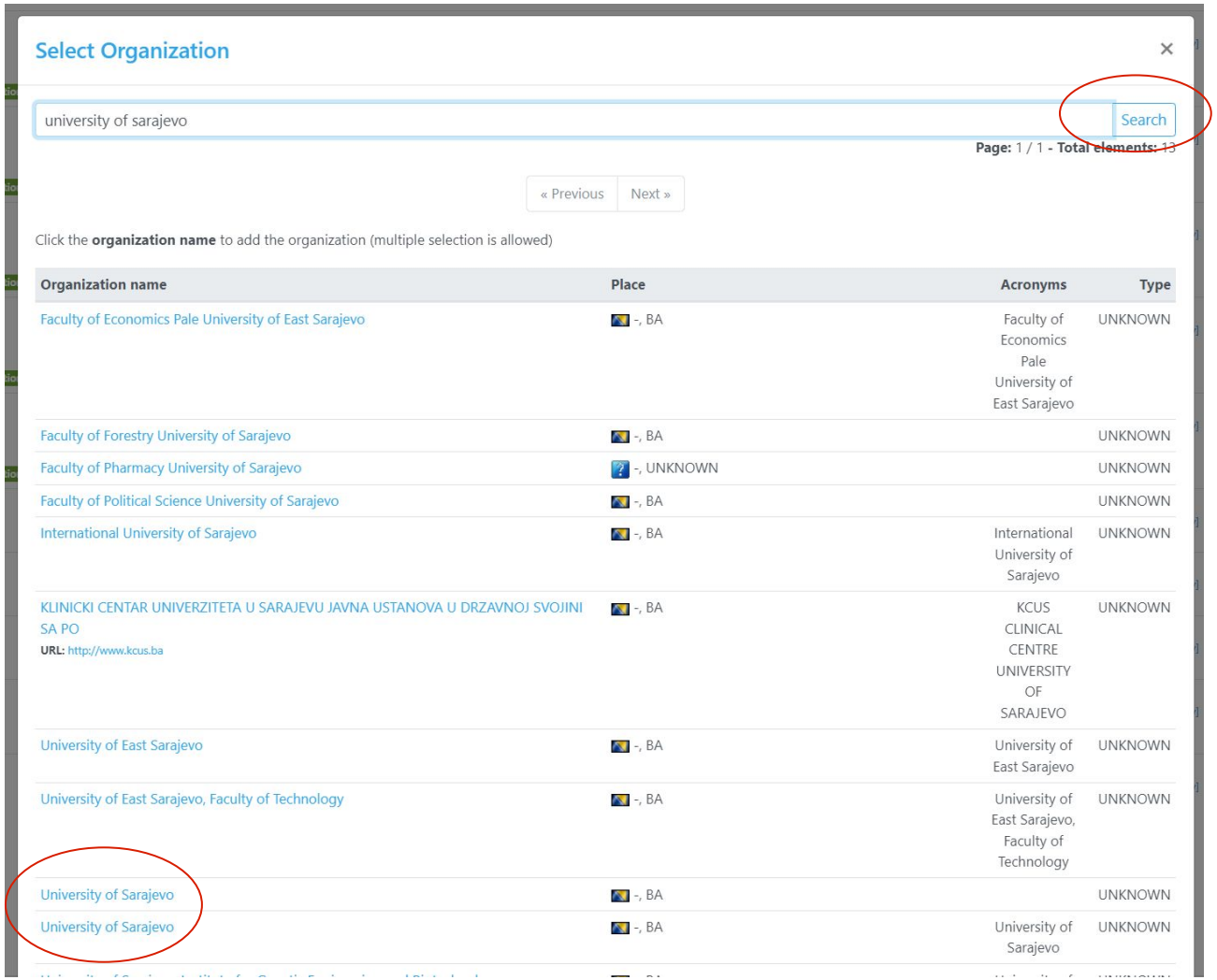
Name	University of Zagreb, Faculty of Agriculture
Type	Education legal body legal person non profit higher education
Place	Zagreb, HR
Acronyms	
Also known as	Sveučilište u Zagrebu, Agronomski fakultet
Urls	http://www.agr.unizg.hr/
Other identifiers	999534561 **(PIC)**

Duplicates

Related organization	Acronym	Country	Source
SVEUCILISTE U ZAGREBU - GEODETSKI FAKULTET **PID (PIC):** 973026207 **URL:** http://www.geof.hr legal body legal person non profit research organization higher education	UNIVERSITY OF ZAGREB-FACULTY OF GEODESY	HR	**Original Id:** corda_____::973026207 **OA Graph Node ID:** corda_____::84dde38fa21b40fad483aa5267d51239 [try] **Provenance:** CORDA - COmmon Research DAta Warehouse **Added by:** Martina.Zugaj@irb.hr

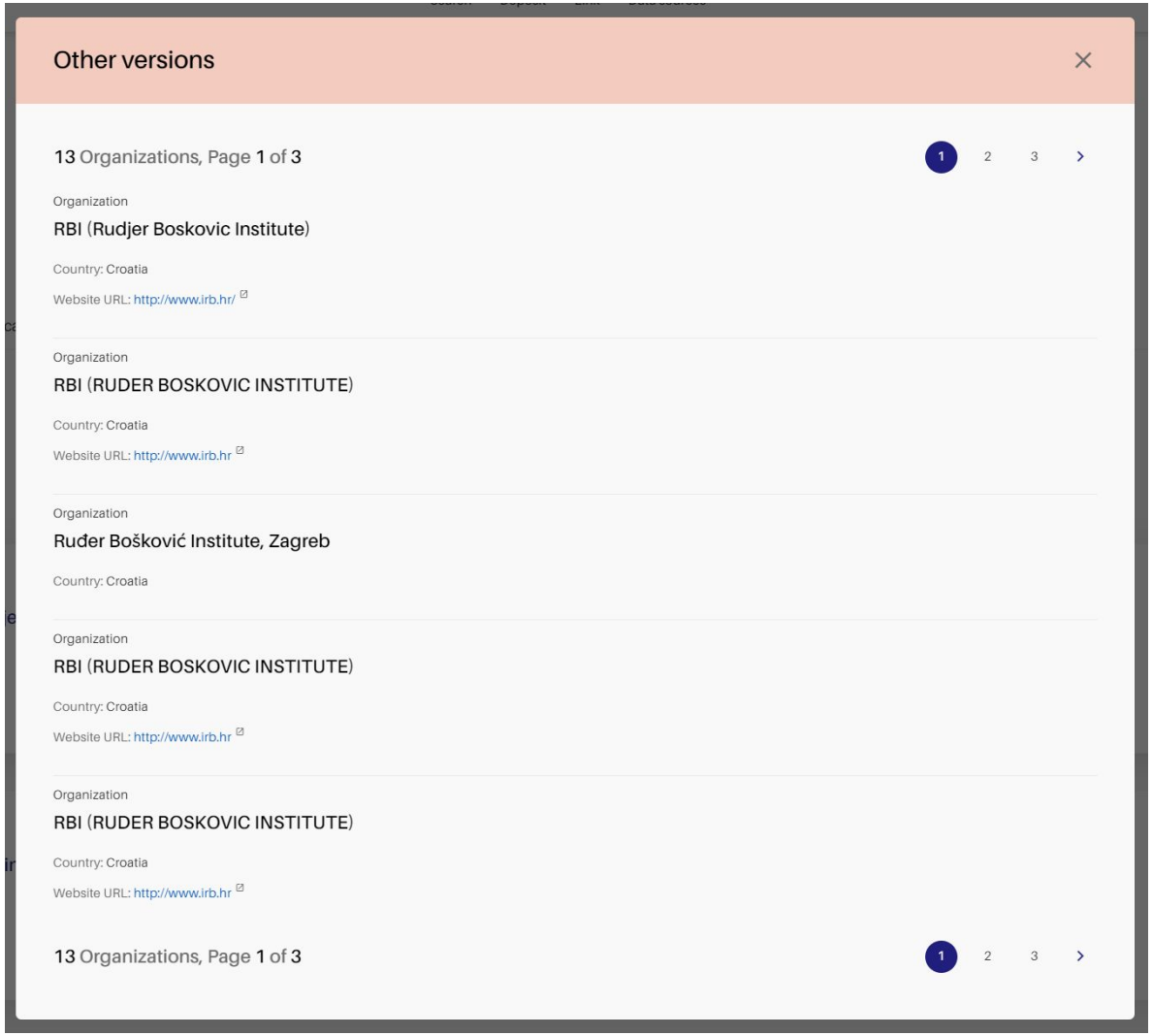
Manual entry of duplicates

Single org page → Duplicates tab → Add button (bottom of the page) → type the search



Duplicates on Openaire

Explore



Resolution of conflicts

A conflict is created when the same organization is approved more than once, thus giving to the same org different OpenOrgs IDs.

- The deduplication **algorithm suggests potential conflicts**, which require confirmation from the curator.
- To add a conflict, curators can select the main organization and add the conflicting org through the **conflicts tab** on the org page.
- Resolving potential conflicts hides the conflicting orgs (**'hidden'**) and creates a **new OpenOrg ID**.

Conflicts

Current country: PT

Filter...

Group 1			
#1	Escola Superior de Saúde da Cruz Vermelha Portuguesa	Lisbon, PT	Education
#2	Escola Superior de Saúde Norte da Cruz Vermelha Portuguesa	Oliveira de Azeméis, PT	Education

add | resolve manually | merge all | all different

Group 2			
#1	Hospital Santa Cruz	Carnaxide, PT	Healthcare
#2	Hospital de Egas Moniz	Lisbon, PT	Healthcare
#3	Centro Hospitalar de Lisboa Ocidental	Lisbon, PT	Healthcare

add | resolve manually | merge all | all different

Group 3			
#1	Iscte - Instituto Universitário de Lisboa	Lisbon, PT	Education
#2	ISCTE - Instituto Universitário de Lisboa	-, PT	Education

add | resolve manually | merge all | all different

Resolve conflicts

Organizations

openorgs___::0000043562

Name	Iscte - Instituto Universitário de Lisboa
Type	Education
Place	Lisbon, PT
Acronyms	Iscte
Also known as	Instituto Universitário de Lisboa Lisbon University Institute Higher Institute of Business and Labour Sciences Instituto Superior de Ciências do Trabalho e da Empresa
Urls	https://www.iscte-iul.pt/
Other identifiers	0000 0001 2220 8863 **(ISNI)** grid.45349.3f **(GRID)** https://ror.org/014837179 **(ROR)** 21625121 **(OrgRef)** Q2886844 **(Wikidata)**

select >>

openorgs___::0000098645

Name	ISCTE - Instituto Universitário de Lisboa
Type	Education
Place	-, PT
Acronyms	ISCTE
Also known as	
Urls	http://www.iscte-iul.pt
Other identifiers	PT0014 **(OrgReg)**

select >>

New group

Main organization

Name	Iscte - Instituto Universitário de Lisboa
Type	Education
Place	Lisbon, PT
Acronyms	Iscte
Also known as	Instituto Universitário de Lisboa Lisbon University Institute Higher Institute of Business and Labour Sciences Instituto Superior de Ciências do Trabalho e da Empresa
Urls	https://www.iscte-iul.pt/
Other identifiers	0000 0001 2220 8863 **(ISNI)** grid.45349.3f **(GRID)** https://ror.org/014837179 **(ROR)** 21625121 **(OrgRef)** Q2886844 **(Wikidata)**

Secondary organization 1

Name	ISCTE - Instituto Universitário de Lisboa
Type	Education
Place	-, PT
Acronyms	ISCTE
Also known as	
Urls	http://www.iscte-iul.pt
Other identifiers	PT0014 **(OrgReg)**

reset

all different | merge

Recap: Main concepts

- **Approved org**: is an organization that was confirmed by a curator. It is persisted with a stable identifier (OpenOrgs ID) and its metadata fields can be enriched/curated;
- **Pending or suggested org**: is an organization to be approved by the national admin; the approval produces an OpenOrgs ID.
- **Duplicate**: possible duplication of the same organization, to be resolved by the user. It is created between an approved org and a raw org (an external org with no OpenOrg ID, ie coming from Corda, Re3data, ROR...).
- Duplicates can be suggested by the dedup algorithm or added by the user inside an org page.
- **Raw org**: the organization with the original registry id. It will never change its ID (those changing ID with the approval are the pending or suggested orgs), it identifies the org as it comes from the original source; it can be added as duplicate.
- **Hidden org**: is an organization not shown in the OpenAIRE public portals. It is the result of a resolution of a conflict. A conflict can be created for example when the same organization is approved by mistake more than once inside OpenOrgs (thus giving different stable OpenOrgs IDs to the same organization). If the curator resolves this conflict by merging two organizations with different OpenOrgs IDs, a new org is created and the previous ones are set as hidden.

Q&A
time!

Hands-on session

1. Choose an **organisation to curate** (e.g. your organisation) from the **approved list**.

(Search → Browse by country → #approved)

   a. **Check and edit** (eventually enrich) the metadata.
   b. Set **parent/child** relations (if any)
   c. Go to the duplicate section and **approve/reject duplicates**.
   d. Manually **add duplicates**.
   e. Are there any **conflicts**? If yes, resolve them.

2. Go to the **duplicates list** *(Curation → Organizations with new duplicates)* and resolve at least a few duplicated organisations.

3. Choose an **organisation** from the suggested list **to approve**. *(Curation → Suggested organisations)*
   a. **Check that it hasn't already been approved**:
      i. If yes, add it as a duplicate *(Search → Browse by country → #approved → single org page → duplicates → add)*
      ii. If no, check, edit and approve.

Check out the new website!

openorgs.openaire.eu

Thank you!

get in touch:

ikoncic@irb.hr
mzugaj@irb.hr
bmacan@irb.hr
openorgs-admin@openaire.eu