



Claudio Atzori, Gina Pavone

Institute of Information Science and Technologies

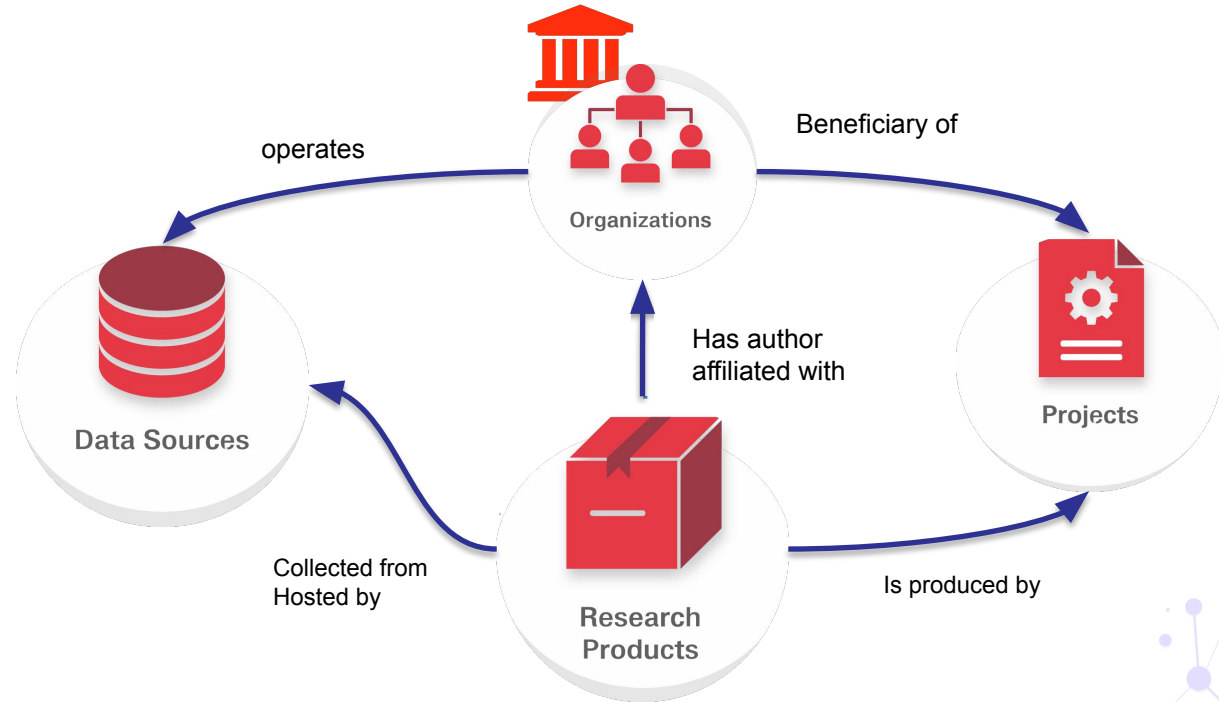
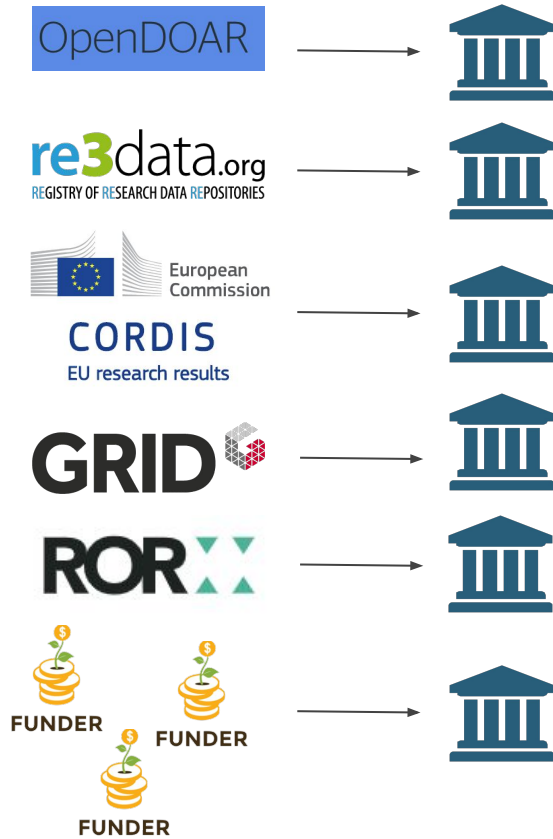
Consiglio Nazionale delle Ricerche - IT

Bridging registries of research organizations

supporting disambiguation and improving the quality of data

OpenAIRE Tech-Clinic 2021-06-21

Organizations in OpenAIRE



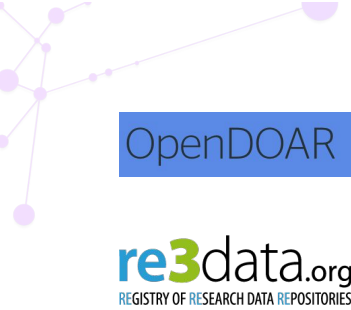
Problem statement

[1] import phase

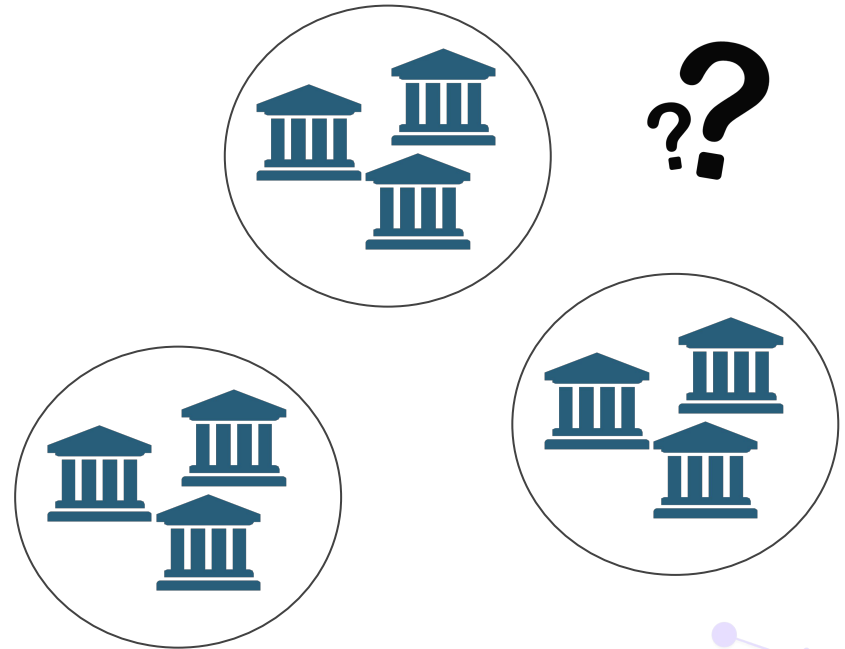
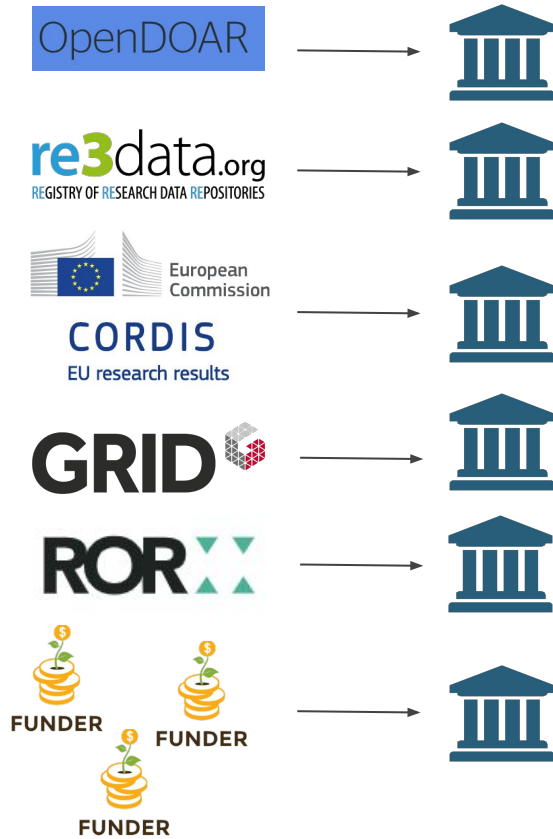
Organizations (mentions)

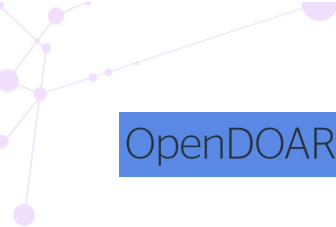
1. Responsible body for the management of a data provider
2. Project participant
3. Extracted from the article full-text as author affiliation

Lack of common identifiers!



Problem statement [2] disambiguation





OpenDOAR



re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

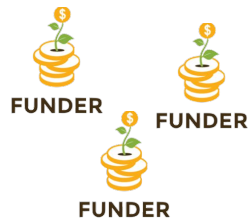


CORDIS
EU research results

GRID



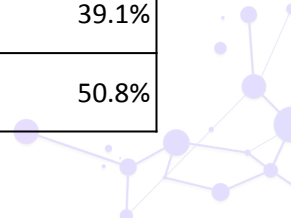
ROR



Problem statement

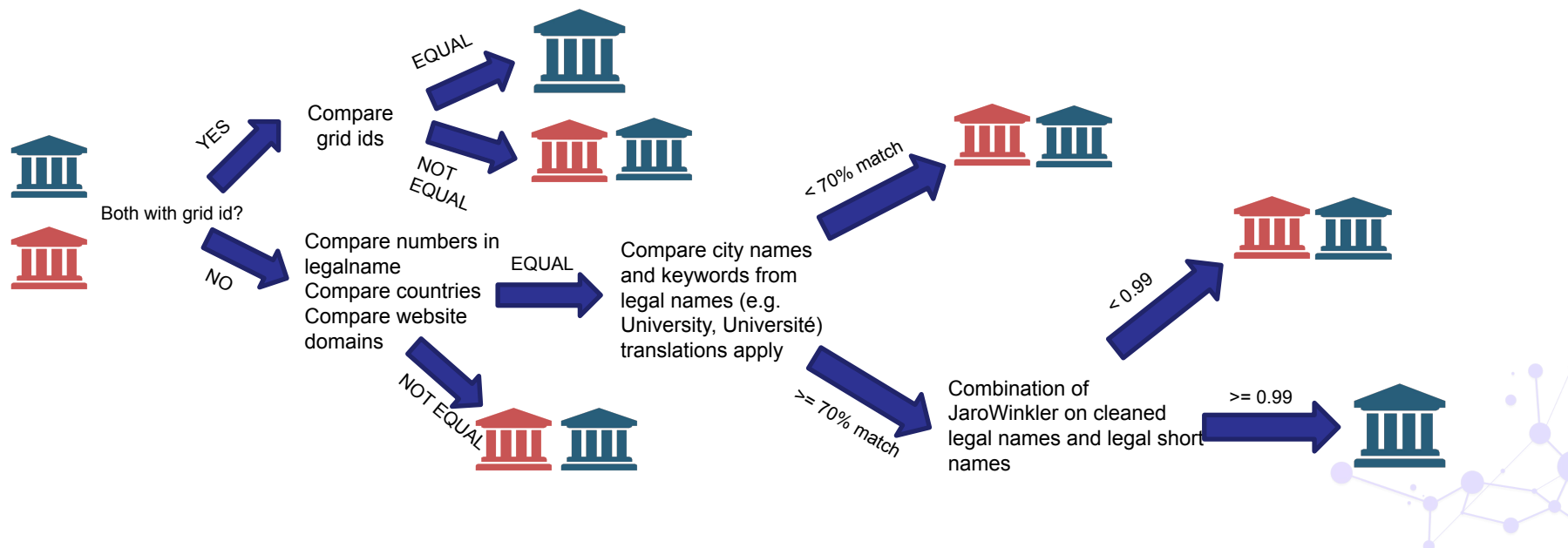
[3] quality / completeness

organization field completeness		
field	missing	% missing
legalname	254	0.1%
legalshortname	156,923	43.2%
alternativenames	319,825	88.0%
country	82,002	22.6%
websiteurl	142,047	39.1%
pid	184,536	50.8%



The duplicate identification algorithm in a nutshell

- Group organisations by legalname and by website URL (4 clustering functions)
- For each cluster we perform pairwise comparisons with early exit strategies.



Drawbacks



OpenDOAR



re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES



European
Commission



CORDIS
EU research results

GRID



ROR



FUNDER
FUNDER
FUNDER



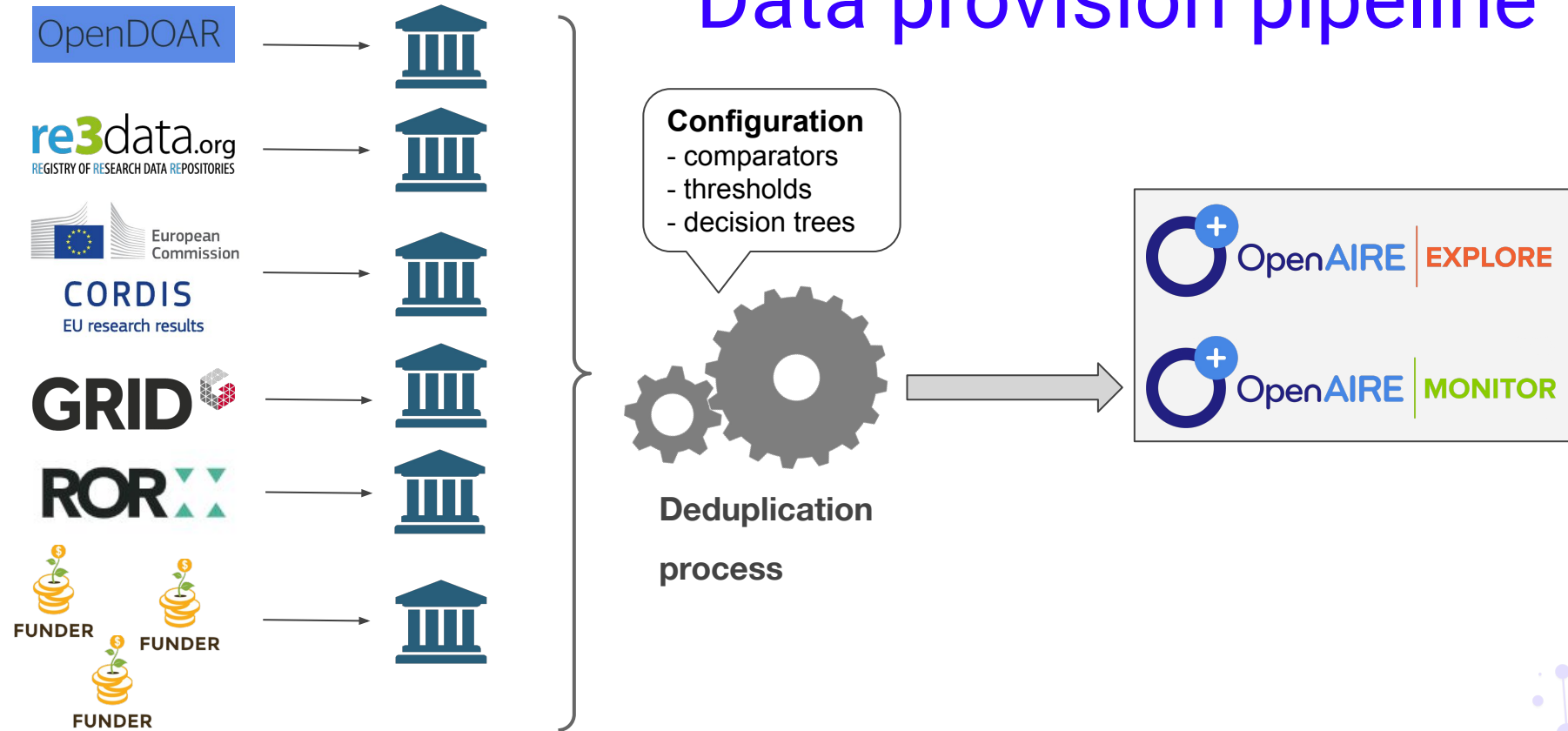
False positives

- Inaccurate statistics (e.g. for institutional and country monitors)
- Inaccurate search results (e.g. finding only the US Concordia University and not the Canadian one)

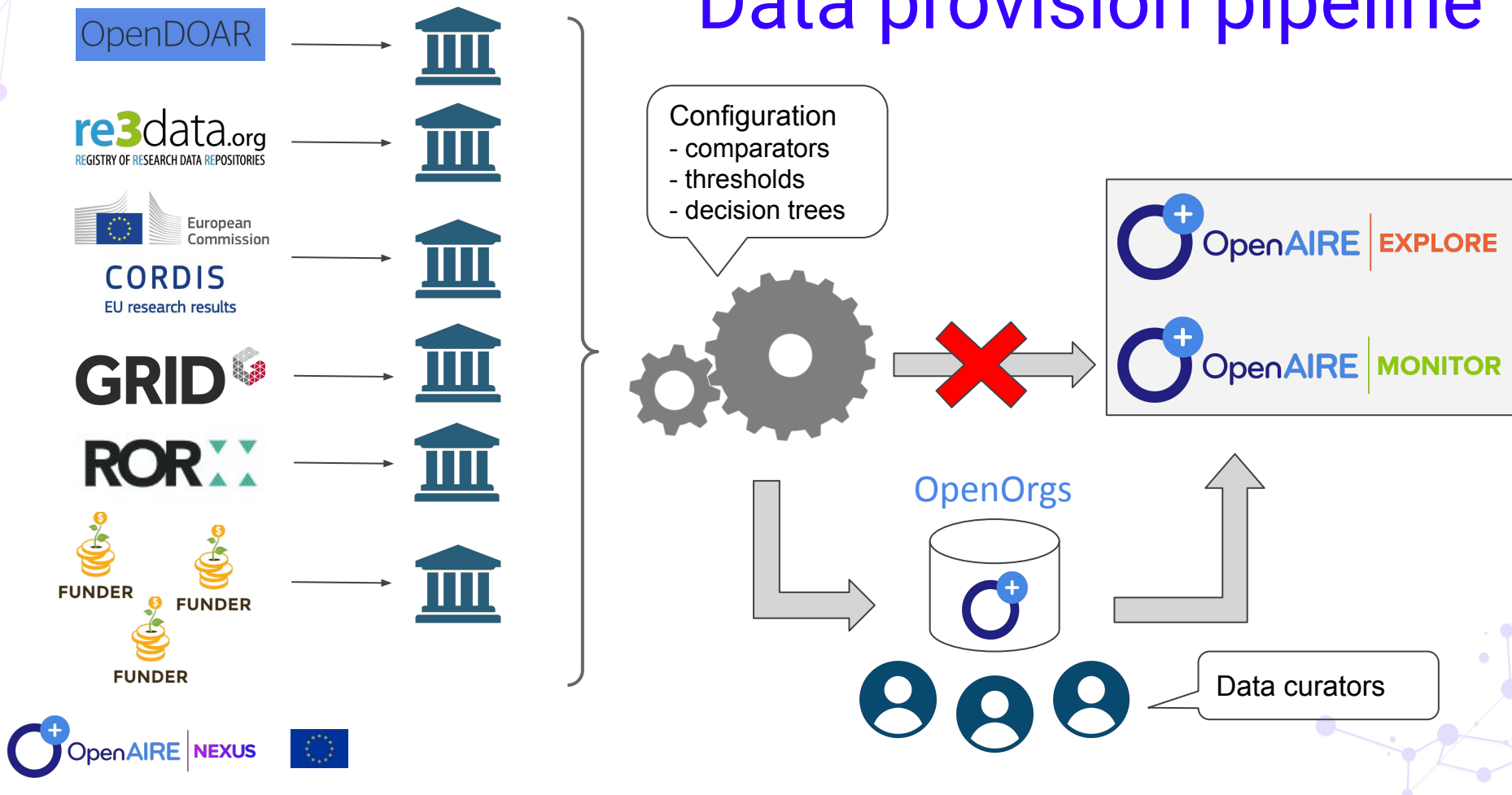
False negatives

- One instance is linked to the papers, one to the project... but it is the same organization!

Data provision pipeline



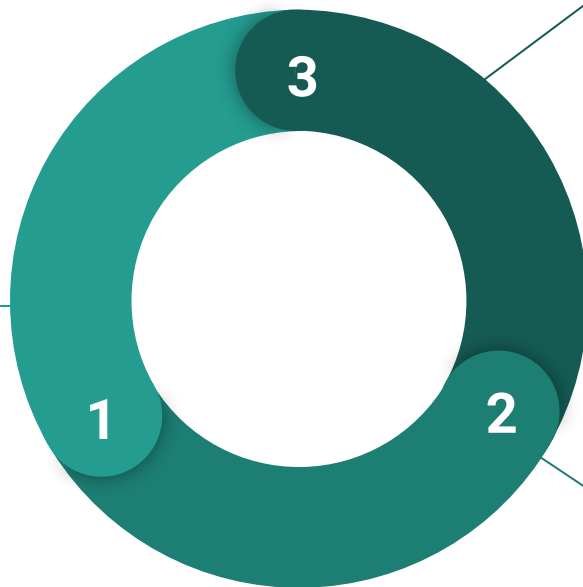
Data provision pipeline



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. This will happen when a new source of organization objects is registered in OpenAIRE (typically a new funder providing descriptions about projects and the relative beneficiaries)



Metadata curation

The quality of the data is a key aspect enabling the realization of data-centric services. 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.

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

User roles



Simple User

Scope limited to a set of countries

Metadata editing & enrichment

Creation of Org → To be approved by national admin/super admin

Approval of suggestions from the duplicate identification algorithm



National admin

Is a user (limited to a set of countries)

Approval of pending orgs → generates approved Orgs

Grant access to users for the Nation

Resolution of conflicts



Super admin

Is a National Admin for all countries

Grant access to National Admins

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 (suggested) org: is an organization to be approved by the national admin; the approval produces an OpenOrgs ID.

- Suggested by the dedup algorithm
- Suggested by the user

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...).

- Suggested by the dedup algorithm
- Created by the user

Hidden org: is an organization that should not be shown in the OpenAIRE public portals

Conflict: when two or more “approved” orgs are:

- Identified as duplicates by de-duplication
- Created by users approving already existing organizations
- They can be resolved through dedicated user actions that set as hidden the conflicting orgs

Enrolling data curators

1. Create an OpenAIRE account @ <https://services.openaire.eu/uoa-user-management/register.jsp>
2. Sign in @ <https://beta.orgs.openaire.eu>
3. You may choose different countries that you want to curate
4. Get in contact with admins: **openorgs-admin@openaire.eu**

[Sign In](#) [Info](#)

Access Guide

To use this service you have to perform the following steps:

1. Register on the OpenAIRE portal.
2. Login using the OpenAIRE credentials
3. Compile the Authorization Request Form
4. An administrator will authorize you as soon as possible

beta.orgs.openaire.eu/authorizationRequest

OpenOrgs Database

Authorization request

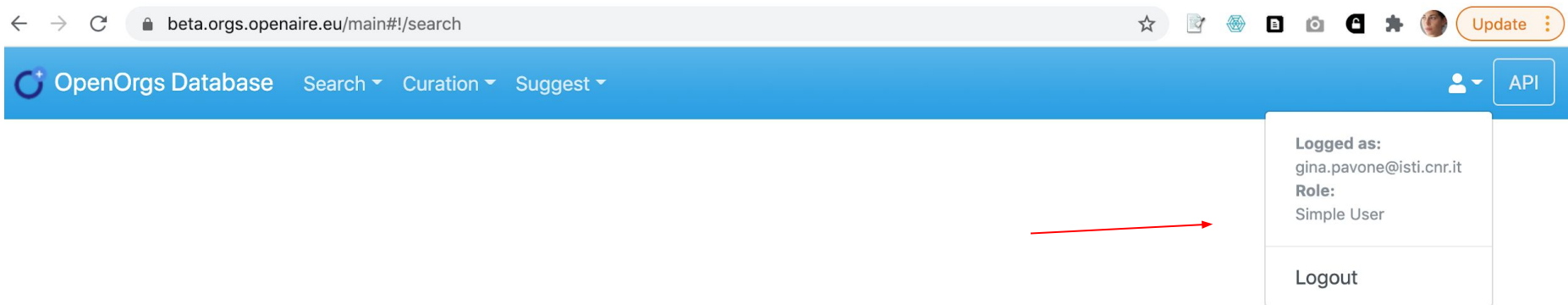
Hello 'gina.pavone@isti.cnr.it', you don't have a role yet
To apply as data curator compile the form below, an administrator will authorize you as soon as possible.

Select your countries

☐ Afghanistan	☐ Åland Islands	☐ Albania
☐ Algeria	☐ American Samoa	☐ Andorra
☐ Angola	☐ Anguilla	☐ Antarctica
☐ Antigua and Barbuda	☐ Argentina	☐ Armenia
☐ Aruba	☐ Australia	☐ Austria
☐ Azerbaijan	☐ Bahamas	☐ Bahrain
☐ Bangladesh	☐ Barbados	☐ Belarus
☐ Belgium	☐ Belize	☐ Benin
☐ Bermuda	☐ Bhutan	☐ Bolivia
☐ Bonaire, Sint Eustatius and Saba	☐ Bosnia and Herzegovina	☐ Botswana
☐ Bouvet Island	☐ Brazil	☐ British Indian Ocean Territory
☐ Brazil	☐ Bulgaria	☐ Bulgaria

Simple user operations

<https://beta.orgs.openaire.eu>



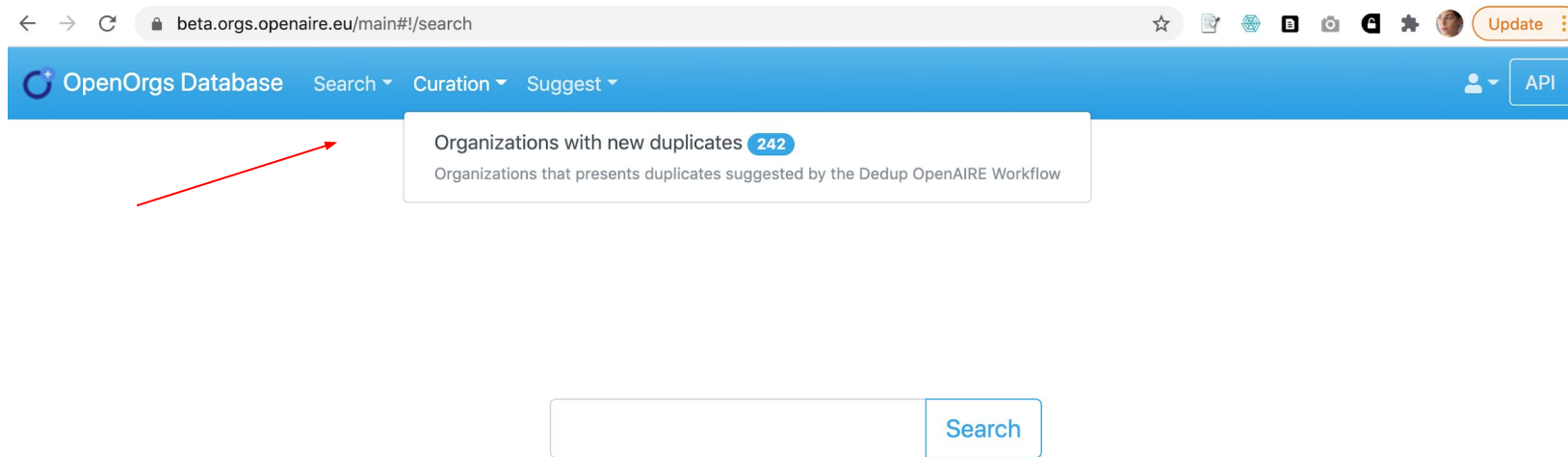
The screenshot shows the OpenOrgs Database website interface. The browser address bar displays `beta.orgs.openaire.eu/main#!/search`. The website header is blue and contains the OpenOrgs Database logo, navigation links for Search, Curation, and Suggest, and a user profile dropdown menu. The user profile dropdown menu is open, showing the user's logged-in status as `gina.pavone@isti.cnr.it` with the role of `Simple User`. A red arrow points to the user profile dropdown menu.

OpenOrgs Database Search Curation Suggest

Logged as:
gina.pavone@isti.cnr.it
Role:
Simple User

Logout

Simple user task: resolution of duplicates



Menu: Curation → Organizations with new duplicates.

An automated deduplication process has spotted similarities between two or more organizations and it suggests to confirm they are the same organization

Curation → organization with new duplicates

← → ↻ beta.orgs.openaire.eu/main#!/duplicates/IT ☆        [Update](#) ⋮

 OpenOrgs Database Search ▾ Curation ▾ Suggest ▾  [API](#)

Duplicates

Current country: IT ▾

Filter...

Organization	Place	# pending duplicates
Agenzia Regionale per la Protezione Ambientale 	 Turin, IT	1
Alfasigma (Italy) 	 Bologna, IT	1
Ambiente Italia (Italy) 	 Milan, IT	2
AM Testing (Italy) 	 Pisa, IT	2
Archivio di Stato di Venezia 	 Venice, IT	1
Associazione della Croce Rossa Italiana 	 Rome, IT	1
Associazione Italiana per la Ricerca Industriale 	 Rome, IT	2
Azienda Mobilità e Trasporti 	 Genoa, IT	1
Azienda Ospedaliero-Universitaria Careggi 	 Florence, IT	1
Azienda Trasporti Milanesi (Italy) 	 Milan, IT	1
Azienda Unità Sanitaria Locale 11 di Empoli 	 Empoli, IT	1
Base Protection (Italy) 	 Barletta, IT	1

Confirming identity between orgs

Polytechnic University of Milan



Registered organization

Name	<u>Polytechnic University of Milan</u>
Type	Education
Place	<u>Milan, IT</u>
Acronyms	
Also known as	Politecnico di Milano Polytechnikum Mailand Polytechnic University of Milan École polytechnique de milan
Urls	<u>http://www.polimi.it/en/</u>

Duplicates

Related organization	Acronym	Country	Source	
Politecnico di MILANO <u>URL:</u>		 IT	Original Id: miur_____: :45646ba06dcfbb446dee13fdf2dfc 053 Provenance: Ministero dell'Istruzione dell'Università e della Ricerca	☒ ☒ ☒
Politecnico di Milano <u>URL:</u> <u>http://www.polimi.it/</u>		 IT	Original Id: opendoar_____: :Politecnico_di_Milano_IT Provenance: OpenDOAR	☒ ☒ ☒
Polytechnic University of Milan <u>URL:</u>		 IT	Original Id: ukri_____: :84803376-D657-41A3- A5DB-E1341282462E Provenance: UK Research and Innovation	☒ ☒ ☒

Confirming identity between organizations

← → ↺

beta.orgs.openaire.eu/main#!/duplicates/IT

☆ 📄 🌐 📁 📷 📺 🛠️ 📄 🧑

Update ⋮

Registered organization

Name	Polytechnic University of Milan
Type	Education
Place	Milan, IT
Acronyms	
Also known as	Politecnico di Milano Polytechnikum Mailand Polytechnic University of Milan École polytechnique de milan
Urls	http://www.polimi.it/en/

Duplicates

Related organization	Acronym	Country	Source	
<u>Politecnico di MILANO</u> URL:		IT	Original Id: miur_____: 45646ba06dcfbb446dee13fdf2dfc053 Provenance: Ministero dell'Istruzione dell'Università e della Ricerca	☐ ☐ ☒
Politecnico di Milano URL: http://www.polimi.it/		IT	Original Id: opendoar_____: Politecnico_di_Milano_IT Provenance: OpenDOAR	☐ ☐ ☒
Polytechnic University of Milan URL:		IT	Original Id: ukri_____: 84803376-D657-41A3-A50B-E1341282462E Provenance: UK Research and Innovation	☐ ☐ ☒

Close Save changes

Deduplication errors

Agenzia Regionale per la Protezione Ambientale



Registered organization

Name	Agenzia Regionale per la Protezione Ambientale
Type	Government
Place	Turin, IT
Acronyms	ARPA
Also known as	Agenzia Regionale per la Protezione Ambientale Regional Environmental Protection Agency
Urls	https://www.arpa.piemonte.gov.it/english-version

Duplicates

Related organization	Acronym	Country	Source
AGENZIA REGIONALE PROTEZIONE CIVILE- <u>EMILIA ROMAGNA</u> URL:		 IT	Original Id: corda_____: 998349803 Provenance: CORDA - Common Research Data Warehouse



Close

Save changes

Summary of situation per organization

#Dups = orgs already marked as different, new duplicate suggestions and orgs already marked as the same

Organization name	Place	Acronyms	# Dups diff/sugg/sim	Type
Alcatel-Lucent (Italy)	 Vimercate, IT		- / - / -	Company
Alenia Aermacchi	 Varese, IT		- / - / 4	Company
Alessandro Manzoni Hospital	 Lecco, IT		- / - / 1	Healthcare
Alfasigma (Italy)	 Bologna, IT		- / - / 2	Company
Alitalia (Italy)	 Rome, IT		- / - / 1	Company
ALK-Abelló (Italy)	 Lainate, IT	ALK	- / - / 1	Company
Alleanza Contro il Cancro	 Rome, IT	ACC	- / - / 3	Nonprofit
Allegria	 Milan, IT		- / - / 1	Nonprofit
Alma Mater Studiorum, Università di Bologna	 Bologna, IT	UNIBO	2 / 2 / 10	Education
AlmavivA (Italy)	 Palermo, IT		- / - / 1	Company

There may be uncertainty

Duplicates				
Related organization	Acronym	Country	Source	
THE UNIVERSITY OF <u>THE WEST INDIES</u> URL:	UWI	 JM	Original Id: corda_____: :999860287 Provenance: OpenOrgs Database	☐ ☐ ☒
THE UNIVERSITY OF THE WEST INDIES <u>U WI*</u> URL: http://www.sta.uwi.edu	UWI	 JM	Original Id: corda_____: :999871248 Provenance: CORDA - COmmon Research DAta Warehouse	☐ ☐ ☒
THE UNIVERSITY OF THE WEST INDIES <u>U WI*</u> URL: http://www.sta.uwi.edu	UWI	 JM	Original Id: corda_h2020: :999871248 Provenance: CORDA - COmmon Research DAta Warehouse - Horizon 2020	☐ ☐ ☒
The University of the West Indies URL:	The University of the West Indies	 JM	Original Id: doajarticles::The_University_of_the_West_Indies Provenance: DOAJ-Articles	☐ ☐ ☒
University of the West Indies URL:		 JM	Original Id: nsf_____: :University_of_the_West_Indies Provenance: NSF - National Science Foundation	☐ ☐ ☒
University of the West Indies at Mona URL: http://www.mona.uwi.edu/		 JM	Original Id: opendoar_____: :University_of_the_West_Indies_at_Mona_JM Provenance: OpenDOAR	☐ ☐ ☒
University of the West Indies URL:		 JM	Original Id: ukri_____: :FD944103-8D9E-4A58-B8BC-79B246196C29 Provenance: UK Research and Innovation	☐ ☐ ☒

Creation of a new org

- Simple user → suggests a new org
- National admin → creates

← → ↻ beta.orgs.openaire.eu/main#!/new 🔍 ☆ 📄 🌐 📁 📷 📄 🧩 👤 ⋮

Metadata

Official name and type

name	type
organization name...	type...

Geographical location

city	country	lat	lng
	country...	0	0

Other names and identifiers

Acronyms

new acronym...

Aliases

name	language
new alias...	language...

Identifiers

id	type
new id...	type...

Uris

http://...

Relations

Related organization	Type
No relations	
add a related organization...	rel type...

Suggest a new Organization

Metadata information

Official name and type:

- Name (mandatory)
- Type (mandatory)

(description of types: <https://www.grid.ac/pages/policies>)

Geographical location:

- City (not mandatory, recommended)
- Country (mandatory)
- Lat and long (not mandatory)

Other names and identifiers

Acronyms (not mandatory)

Aliases:

- Alternative names and language of the alternative name (not mandatory)

Identifiers:

Id and type (automated compilation)

Urls (not mandatory)

Relations:

- Related organizations and relation type:
(<https://www.grid.ac/pages/policies>)

Ca Foscari University of Venice

ID: openorgs____:0000097956

Created at July 16, 2020 11:48:19 by import:grid.ac

Modified at July 16, 2020 11:48:19 by import:grid.ac

Metadata Management

Duplicates

Conflicts

Official name and type

name Ca Foscari University of Venice type Education

Geographical location

city Venice country Italy lat 45.4345 lng 12.3265

Other names and identifiers

Acronyms

new acronym...

Aliases

name	language	
Université ca' foscari de venise	fr	
Università Ca' Foscari Venezia	it	
Università Ca' Foscari	UNKNOWN	
Ca Foscari University of Venice	en	
Univerza v Benetkah	sl	
Universität Venedig	de	
new alias...	language...	

Identifiers

Id	type	
0000 0004 1763 0578	ISNI	
https://ror.org/04yzxz566	ROR	
463087	OrgRef	
grid.7240.1	GRID	

Metadata curation

- **It is not necessary to complete all metadata information**

You can choose to complete as much of the metadata as possible, enter only what you feel is necessary, or simply approve organizations (without completing the metadata).

- It is up to you how much information to add depending on the individual case, or depending on the rules established at a national level.
- We can decide together which information should be managed at national level and which should have international consistency.
- We recommend to curate as best as possible the following fields:
 - **Name of the org:** UTF-8 encoded and Latin letters
 - **Other names:** if possible, add the name in English and the acronym
 - **Country** - when working with orgs that do not have an assigned country



National admin operations

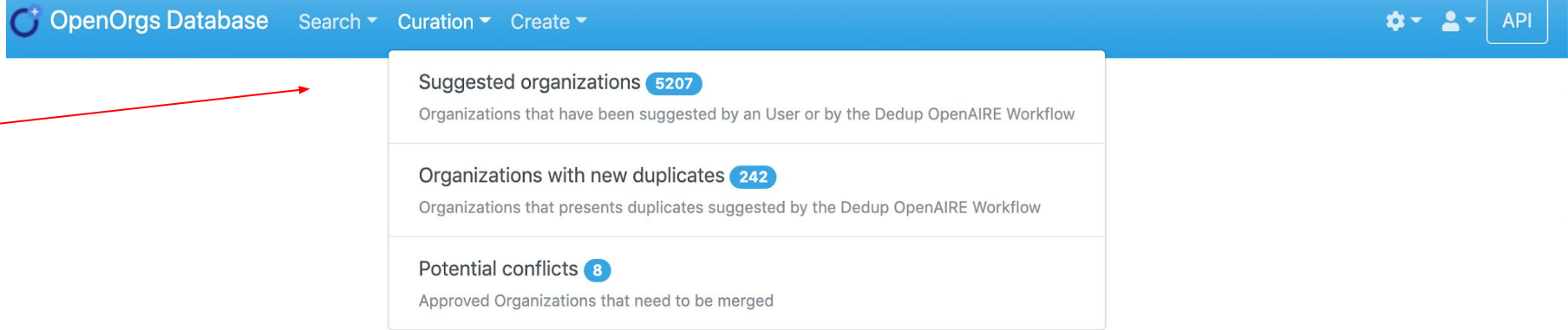
<https://beta.orgs.openaire.eu/>

The screenshot shows the OpenOrgs Database interface. The top navigation bar is blue and contains the following elements from left to right: the OpenOrgs Database logo, a search dropdown menu, a 'Curation' dropdown menu, a 'Create' dropdown menu, a settings gear icon, a user profile icon, and an 'API' button. A red arrow points from the user profile icon to a dropdown menu that displays the following information:

- Logged as:** noad-it@openaire.eu
- Role:** National Admin
- Logout** (button)

Below the navigation bar, there is a large white search area with a text input field and a blue 'Search' button.

National admin curation menu



Menu: Curation → Suggested organizations:

Organizations to be confirmed by a curator. They will be persisted with a stable identifier (OpenOrgs ID) and it will be possible to enrich/curate their metadata fields.

Single organization page

Centro de Estudios en Biodiversidad

ID: pending_org_:f8a225eaf885b93498ea21b27057b3f3

Created at May 7, 2021 08:07:19 by dedupWf

Modified at May 7, 2021 08:07:19 by dedupWf

Metadata Management

Duplicates new

Note

History ↻

This organization is not yet subsumed by an OpenOrg. You can resolve this anomaly by:

- creating a new OpenOrg ID for this organization, or
- adding the organization as a duplicate of an existing OpenOrg

Official name and type

name Centro de Estudios en Biodiversidad

type UNKNOWN

EC flags

Curation → Suggested organizations → single org page

Suggested organizations:

- Metadata Management: to manage descriptive metadata information.
- Duplicates: where possible duplicates of the same organisation are suggested.
- Notes: something to remember
- History: activity log (action and who did it)



National admin → Editing of metadata

University of Rome I (La Sapienza)

ID: ukri_____:473F7B5C-70FE-4079-A7A4-9B33054D3495

Created at March 25, 2021 09:29:53 by dedupWf

Modified at March 25, 2021 09:29:53 by dedupWf

Metadata Management

Duplicates

This organization is not yet subsumed by an OpenOrg. You can resolve this anomaly by:

- creating a new OpenOrg ID for this organization, or
- adding the organization as a duplicate of an existing OpenOrg

Official name and type

name University of Rome I (La Sapienza)

type UNKNOWN

Geographical location

city

country Italy

lat 0.0

lng 0.0

Editing of metadata

← → ↻ beta.orgs.openaire.eu/main#!/edit/0/corda__h2020::938221540 ☆ 📄 🌐 📄 📄 📄 📄 📄

OpenOrgs Database Search ▾ Curation ▾ Create ▾ ⚙️ 👤 API

ABACO SPA

ID: corda__h2020::938221540
Created at March 25, 2021 09:29:53 by dedupWf
Modified at March 25, 2021 09:29:53 by dedupWf

Metadata Management Duplicates

This organization is not yet subsumed by an OpenOrg. You can resolve this anomaly by:

- creating a new OpenOrg ID for this organization, or
- adding the organization as a duplicate of an existing OpenOrg

Official name and type

name ABACO SPA type ✓ UNKNOWN EC flags

Archive
Company
Education
Facility
Government
Healthcare
Nonprofit
Other

Organization Type

Flags from the EC can help in determining the type of the org

beta.orgs.openaire.eu/main#/edit/0/corda_____::999718570

AB.ACUS SRL

ID: corda_____::999718570
Created at March 25, 2021 09:29:53 by dedupWf
Modified at March 25, 2021 09:29:53 by dedupWf

Metadata Management Duplicates

This organization is not yet subsumed by an OpenOrg. You can resolve this anomaly by:

- creating a new OpenOrg ID for this organization, or
- adding the organization as a duplicate of an existing OpenOrg

Official name and type

name AB.ACUS SRL type UNKNOWN EC flags

Geographical location

city country Italy lat 0.0 lng 0.0

Other names and identifiers

Acronyms

ABACUS

new acronym...

Aliases

European Commission Flags

legal-body	legal person	non-profit
research-organization	higher-education	
international-organization	eur-interests	
international-organization	enterprise	
sme validated	multi-code	

Organization Type: not alway an easy task

OpenOrgs Database Search Curation Create API

Cardiac and Vascular Surgery Fondazione Cardiologico Monzino IRCCS University of Milan

ID: snsf_____:Cardiac_and_Vascular_Surgery_Fondazione_Cardiologico_Monzino_IRCCS
Created at March 25, 2021 09:29:53 by dedupWf
Modified at March 25, 2021 09:29:53 by dedupWf

Metadata Management Duplicates

This organization is not yet subsumed by an OpenOrg. You can resolve this anomaly by:

- creating a new OpenOrg ID for this organization, or
- adding the organization as a duplicate of an existing OpenOrg

Official name and type

name	type
Cardiac and Vascular Surgery Fondazione Cardiologico Monzino IRCCS University of Milan	UNKNOWN

EC flags

European Commission Flags

- legal-body legal-person non-profit
- research-organization higher-education
- international-organization Eur-interests
- international-organization enterprise
- sme-validated nuts-code

National admin → approval of orgs

Curation menu → suggested Orgs → Pending Orgs

OpenOrgs Database Search Curation Create

ID:
Created at by
Modified at by

Metadata Management Duplicates

This organization is not yet subsumed by an OpenOrg. You can resolve this anomaly by:

- creating a new OpenOrg ID for this organization, or
- adding the organization as a duplicate of an existing OpenOrg

Official name and type

name Libera Università 'Vita Salute S.Raffaele' MILANO

Geographical location

city country Italy lat 0.0 lng 0.0

Other names and identifiers

Acronyms

new acronym...

Aliases

name language

new alias... language...

Identifiers

id type

new id... type...

Urls

http://...

Relations

Related organization Type

No relations

add a related organization... rel type...

Approve as new Organization

Archive
Company
Education
Facility
Government
Healthcare
Nonprofit
Other
✓ UNKNOWN

A use case from the OpenAIRE EXPLORE portal



Organization
UNIPD
University of Padua
[View all 12 versions](#)

Web page: <http://www.unipd.it/en/home-page>

Country: Italy

FUNDING / PROJECTS	CONTENT PROVIDERS	PUBLICATIONS	OTHER RESEARCH
9	3	91,200	329

Other versions of

Organization
UNIPD
University of Padua

12 ORGANIZATIONS, PAGE 1 OF 3

1 2 3 >

Organization
UNIPD (Universitat de Pàdua)
Country: Italy
Website URL: <http://www.unipd.it/en/home-page>

Organization
UNIPD (University of Padua)
Country: Italy
Website URL: <http://www.unipd.it/en/home-page>

Organization
UNIPD (University of Padua)
Country: Italy
Website URL: <http://www.unipd.it/en/home-page>

Organization
University of Padua
Country: Italy
Website URL: <https://www.unipd.it/>

Organization
University of Padua
Country: Italy

12 ORGANIZATIONS, PAGE 1 OF 3

1 2 3 >

Simple search by name, acronym or part of the name

University of Padua



Registered organization

Duplicates

Related organization	Acronym	Country	Source	
University of Padova URL:		 IT	Original Id: chistera____: :bf3f5e92d101afe06f9f2ae1fdc6f0d4 Provenance: CHIST-ERA	☐ ☐ ☒
Università degli Studi di PADOVA URL:		 IT	Original Id: miur_____: :bd694496f764d566b055c76a4040bf32 Provenance: Ministero dell'Istruzione dell'Università e della Ricerca	☐ ☐ ☒
University of Padua URL: https://www.unipd.it/		 IT	Original Id: opendoar____: :University_of_Padua_IT Provenance: OpenDOAR	☐ ☐ ☒
University of Padova URL: http://www.unipd.it		 IT	Original Id: re3data____: :a5e8940188729f7f0fbe0d1ef21bc41e Provenance: Registry of Research Data Repository	☐ ☐ ☒
Universita degli Studi di Padova URL: http://www.unipd.it/		 IT	Original Id: re3data____: :f90f258b83fe628a8651e0797c8fa5f5 Provenance: Registry of Research Data Repository	☐ ☐ ☒

Curation → suggested orgs

Beware not to create conflicts while approving new orgs

← → ↺ beta.orgs.openaire.eu/main#/pendings/IT ☆ 📄 🌐 📄 📷 📺 ⚙️ 👤 API

OpenOrgs Database Search ▾ Curation ▾ Create ▾ ⚙️ 👤 API

Pending Organizations

Current country: IT ▾

Filter...

Organization name	Place	Acronyms	Type
8 B SRL	🇮🇹 -, IT		UNKNOWN
A2A CALORE & SERVIZI SRL	🇮🇹 -, IT	ACS	UNKNOWN
A2A RETI ELETTRICHE SPA	🇮🇹 -, IT		UNKNOWN
A2A SPA	🇮🇹 -, IT	A2A	UNKNOWN
A.ABETE SRL	🇮🇹 -, IT		UNKNOWN
A & A F. LLI PARODI SPA	🇮🇹 -, IT	A & A	UNKNOWN
AARHUS GEOFISICA SRL	🇮🇹 -, IT		UNKNOWN
AAT-ADVANCED ANALYTICAL TECHNOLOGIES	🇮🇹 -, IT	AAT	UNKNOWN
ABACO SPA	🇮🇹 -, IT		UNKNOWN
AB.ACUS SRL	🇮🇹 -, IT	ABACUS	UNKNOWN
AB.ACUS SRL	🇮🇹 -, IT	ABACUS	UNKNOWN
AB ANALITICA SRL	🇮🇹 -, IT	AB Analitica	UNKNOWN
ABB SPA	🇮🇹 -, IT	ABB	UNKNOWN
Abdus Salam ICTP	🇮🇹 -, IT		UNKNOWN
A. BENEVENUTA & C.S.P.A.	🇮🇹 -, IT		UNKNOWN
ABICH SRL	🇮🇹 -, IT		UNKNOWN

(All users) Curation → suggested orgs → select the approved org (here AB.ACUS SRL) → duplicates tab → approve the suggested duplicate (if right) and save → **link the additional duplicate by clicking on add**

OpenOrgs Database

Search

Curation

Create

beta.orgs.openaire.eu/main#/edit/2/openorgs____:0000098407

beta.orgs.openaire.eu says
Events updated !!!
OK

AB.ACUS SRL

ID: openorgs____:0000098407
Created at April 7, 2021 15:56:56 by noad-it@openaire.eu
Modified at April 7, 2021 15:57:34 by noad-it@openaire.eu

Metadata Management

Duplicates

Conflicts

Current organization

Name

AB.ACUS SRL

Type

Company

legal person

enterprise

same validated

Place

-, IT

Acronyms

ABACUS

Also known as

Urls

http://www.ab-acus.com

Other identifiers

Duplicates

Related organization

AB.ACUS SRL
URL: http://www.ab-acus.com
legal person enterprise same validated

Acronym

ABACUS

Country

IT

Source

Original Id: corda____:999718570
Added by: noad-it@openaire.eu

add

save

Conflicts

OpenOrgs Database Search Curation Create

Conflicts

Current country: IT

Filter...

Group 1

#1 European University Institute

#2 University Institute of European Studies

add resolve manually merge all all different

Suggested organizations 5204
Organizations that have been suggested by an User or by the Dedup OpenAIRE Workflow

Organizations with new duplicates 242
Organizations that presents duplicates suggested by the Dedup OpenAIRE Workflow

Potential conflicts 8
Approved Organizations that need to be merged

Florence, IT Education

Turin, IT Education

National admin: Menu → Curation → Potential conflicts

When two or more “approved” orgs are:

- Identified as duplicates by de-duplication algorithm
- Created by users approving already existing organizations

They can be resolved through dedicated actions that set as hidden the conflicting orgs

Conflict resolution: all different

beta.orgs.openaire.eu/main#!/conflicts/IT

OpenOrgs Database Search Curation Create API

Conflicts

Current country: IT

Filter...

Group 1			
#1	European University Institute	Florence, IT	Education
#2	University Institute of European Studies	Turin, IT	Education
add resolve manually merge all all different			

Group 2			
#1	T4tech (Italy)	Genoa, IT	Company
#2	SofTech (Italy)	Milan, IT	Company
add resolve manually merge all all different			

Group 3			
#1	Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise G. Caporale	Teramo, IT	Facility
#2	Istituto Zooprofilattico Sperimentale dell'Umbria e delle Marche	Perugia, IT	Facility
#3	Istituto Zooprofilattico Sperimentale della Sardegna	Sassari, IT	Facility
add resolve manually merge all all different			

Conflict resolution: merge all

Group 5			
#1	Carlo Forlanini Hospital	 Rome, IT	Healthcare
#2	Azienda Ospedaliera San Camillo Forlanini	 Rome, IT	Healthcare
add resolve manually merge all all different			



Conflict resolution: resolve manually

Group 5			
#1	Accademia Europea di Bolzano	 Bolzano, IT	Government
#2	ACCADEMIA EUROPEA DI BOLZANO	 -, IT	Education
add resolve manually merge all all different			



beta.orgs.openaire.eu/main#!/conflicts/IT

Resolve conflicts

Organizations

openorgs____:0000010635

Name	Accademia Europea di Bolzano
Type	Government
Place	Bolzano, IT
Acronyms	EURAC
Also known as	Europäische Akademie Bozen European Academy of Bozen Accademia Europea de Bulsan Accademia Europea di Bolzano
Urls	http://www.eurac.edu/en/Pages/default.aspx
Other identifiers	0000 0001 1089 6435 (ISNI) Q1245390 (Wikidata) grid.418908.c (GRID) https://ror.org/01xt1w755 (ROR)

select >>

New group

select main organization

openorgs____:0000098362

Name	ACCADEMIA EUROPEA DI BOLZANO
Type	Education
Place	~, IT
Acronyms	
Also known as	Europäische Akademie Bozen
Urls	http://www.eurac.edu
Other identifiers	

select >>

Group 3

#1 Istituto Sperimentale per la Zootecnia

#2 Istituto Sperimentale Agronomico

add resolve manually merge all

Group 4

#1 Istituto Ortopedico Galeazzi

#2 Istituto Ortopedico Gaetano Pini

#3 Istituto Ortopedico Rizzoli

add resolve manually merge all

Group 5

#1 Accademia Europea di Bolzano

#2 ACCADEMIA EUROPEA DI BOLZANO

add resolve manually merge all

Group 6

#1 Azienda Ospedaliero Universitaria Pisana

#2 Azienda Ospedaliero-Universitaria Careggi

#3 Azienda Ospedaliera Universitaria Senese

#4 Meyer Children's Hospital

#5 Azienda Ospedaliero Universitario Mater Dei

add resolve manually merge all

Group 7

#1 Carlo Forlanini Hospital

#2 Azienda Ospedaliera San Camillo-Forlanini

add resolve manually merge all

Pisa, IT Company

Rome, IT Facility

Bari, IT Facility

Milan, IT Healthcare

Milan, IT Healthcare

Bologna, IT Healthcare

Bolzano, IT Government

~, IT Education

Pisa, IT Healthcare

Florence, IT Healthcare

Siena, IT Healthcare

Florence, IT Healthcare

Catanzaro, IT Healthcare

Rome, IT Healthcare

Rome, IT Healthcare

#2 Synthema (Italy)

[add](#) [resolve manually](#) [merge all](#) [all d](#)

Group 3

#1 Istituto Sperimentale per la Zootecnia

#2 Istituto Sperimentale Agronomico

[add](#) [resolve manually](#) [merge all](#) [all d](#)

Group 4

#1 Istituto Ortopedico Galeazzi

#2 Istituto Ortopedico Gaetano Pini

#3 Istituto Ortopedico Rizzoli

[add](#) [resolve manually](#) [merge all](#) [all d](#)

Group 5

#1 Accademia Europea di Bolzano

#2 ACCADEMIA EUROPEA DI BOLZANO

[add](#) [resolve manually](#) [merge all](#) [all d](#)

Group 6

#1 Azienda Ospedaliero Universitaria Pisana

#2 Azienda Ospedaliero-Universitaria Careggi

#3 Azienda Ospedaliera Universitaria Senese

#4 Meyer Children's Hospital

#5 Azienda Ospedaliero Universitario Mater D

[add](#) [resolve manually](#) [merge all](#) [all d](#)

Group 7

#1 Carlo Forlanini Hospital

#2 Azienda Ospedaliera San Camillo-Forlanini

[add](#) [resolve manually](#) [merge all](#) [all d](#)

Resolve conflicts

Organizations

no organization(s)

New group

Main organization

Name Accademia Europea di Bolzano

Type Government

Place Bolzano, IT

Acronyms EURAC

Also known as Europäische Akademie Bozen
European Academy of Bozen
Accademia Europea de Bulsan
Accademia Europea di Bolzano

Urls <http://www.eurac.edu/en/Pages/default.aspx>

Other identifiers 0000 0001 1089 6435 (ISNI)
Q1245390 (Wikidata)
grid.418908.c (GRID)
<https://ror.org/01xt1w755> (ROR)

Secondary organization 1

Name ACCADEMIA EUROPEA DI BOLZANO

Type Education

Place -, IT

Acronyms

Also known as Europäische Akademie Bozen

Urls <http://www.eurac.edu>

Other identifiers

[reset](#)

[all different](#) [merge](#)

[Close](#)



Open questions

- Do we need to standardize metadata curation across all users or just at a national level?
- Do we need to curate only research institutions or also companies?
- Orgs main name should be in English or in national language?
- Do we want to adopt national standards for organization names? (uppercase/lowercase letters, e.g. upper cases for each words that is not a preposition as in English; use of prepositions, etc.)
- Should we look up official names of organizations or do we stick to a standard we establish for OpenOrgs?





Next steps

- Continue the disambiguation task (Curation menu → Organizations with new duplicates)
- Approve new orgs (Curation menu → Suggested orgs)
- Resolve conflicts (Curation menu → Potential conflicts)
- Finalise testing on beta
- Have your say on open questions (in the feedback doc)
- Give us feedback:
 - https://docs.google.com/document/d/19_kvibAkQpyAOw8LphnapzblwrV90LwkshsP_HSOW_uQ/edit?usp=sharing





Support

openorgs@openaire.eu

- Intended for registered users, internal communications

openorgs-admin@openaire.eu

- Intended to handle support requests, e.g. user registrations, assignment of countries. Open to all.
- 

Thanks!

claudio.atzori@isti.cnr.it

gina.pavone@isti.cnr.it

Support: openorgs@openaire.eu and openorgs-admin@openaire.eu