

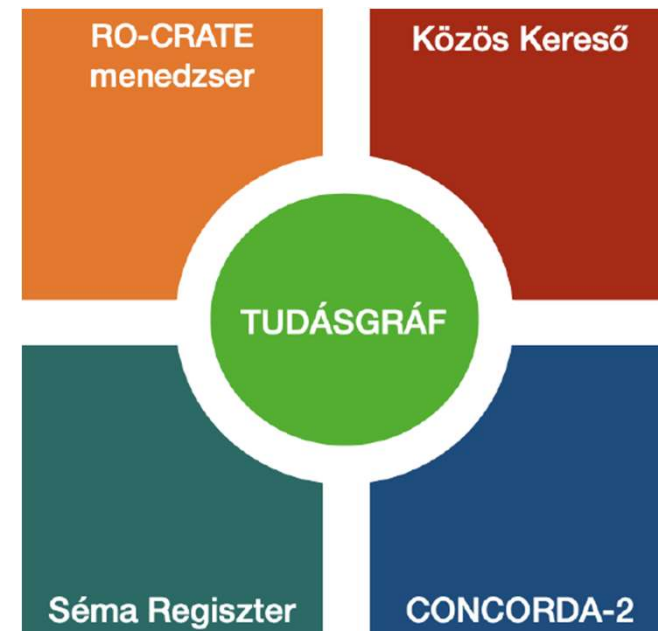


# Multidiszciplináris metaadatséma kezelés integrálása az ARP platformba

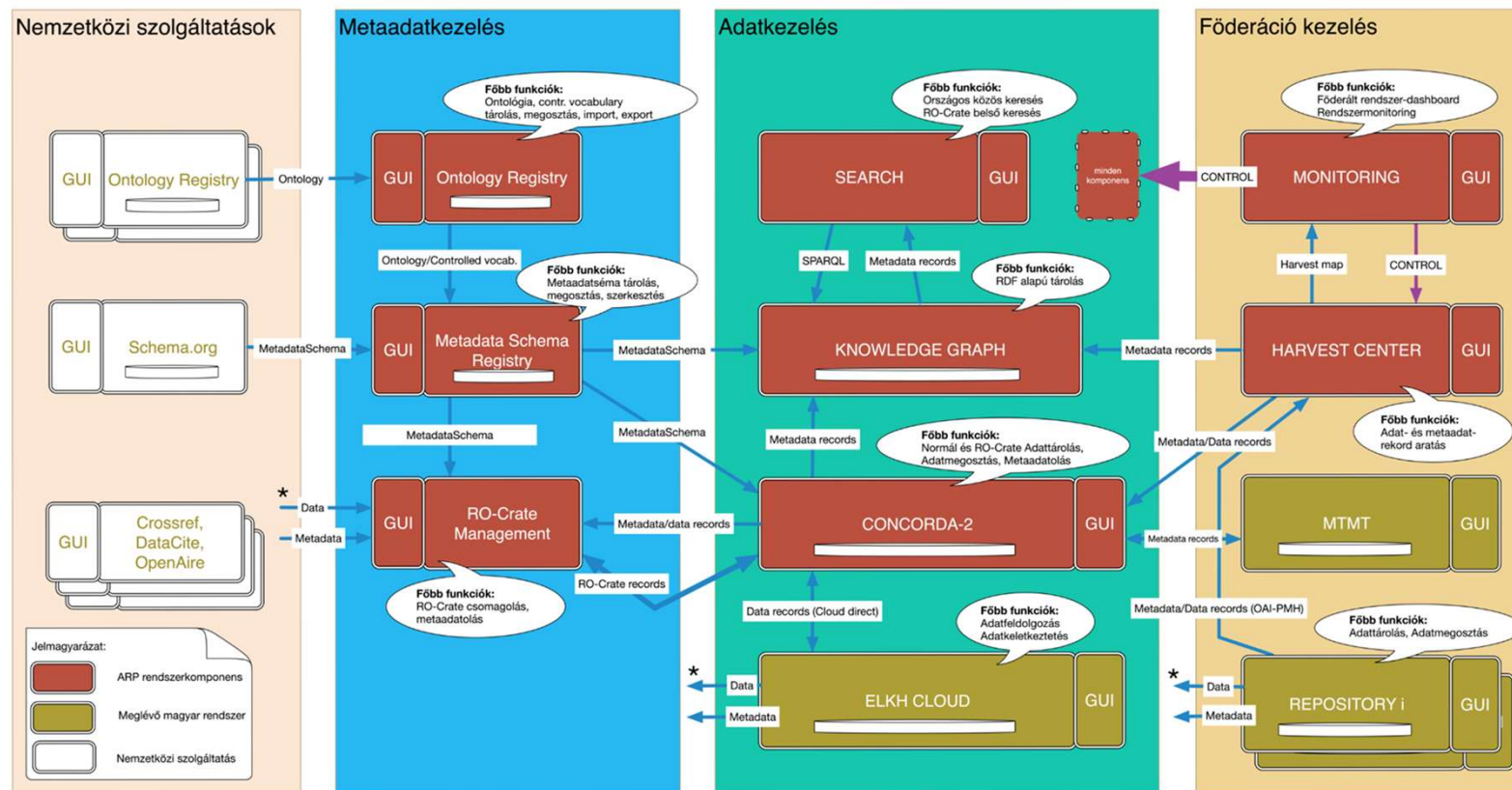
Pataki Balázs  
SZTAKI – DSD  
[pataki@sztaki.hu](mailto:pataki@sztaki.hu)

# ELKH Adat Repozitórium Projekt

- Cél
  - o A teljes ELKH kutatóhálózat számára folyamatos és hosszú távú kutatási repozitóriumi infrastruktúra-szolgáltatás megvalósítása
- Résztvevők
  - o Számítástechnikai és Automatizálási Kutatóintézet (SZTAKI)
  - o Társadalomtudományi Kutatóközpont (TK)
  - o Wigner Fizikai Kutatóközpont (Wigner FK)
- Időtartam
  - o 2021.11.01. - 2023.12.31.

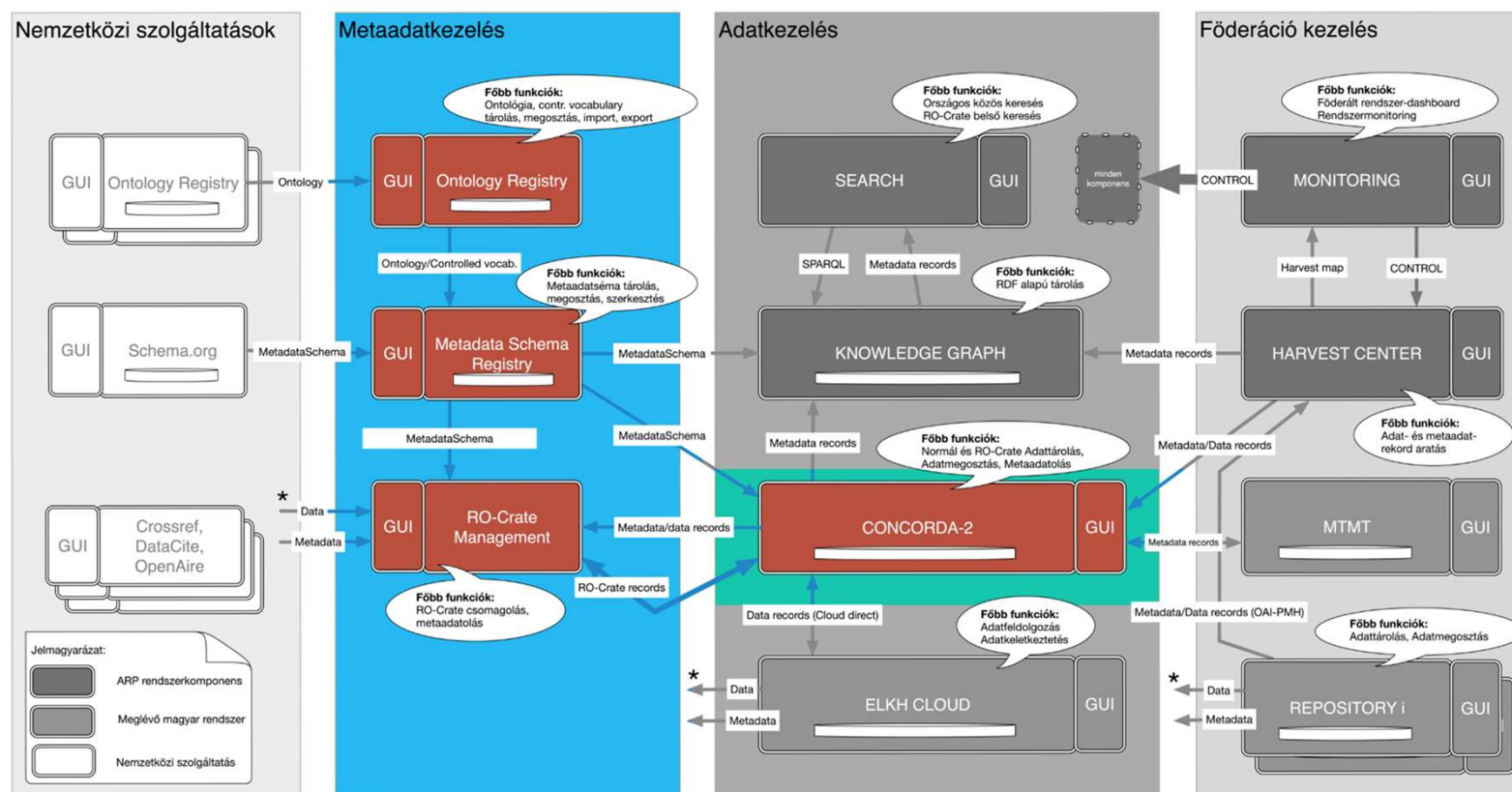


# Architektúra



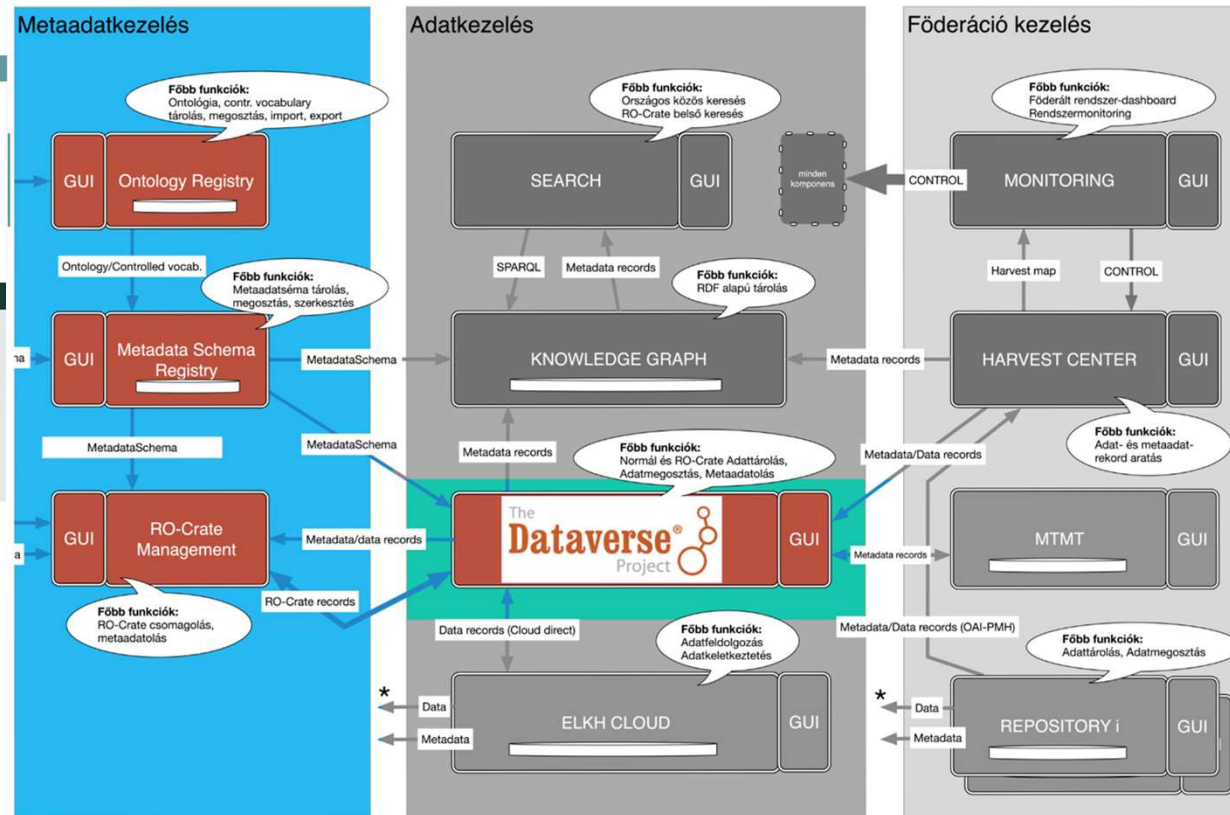
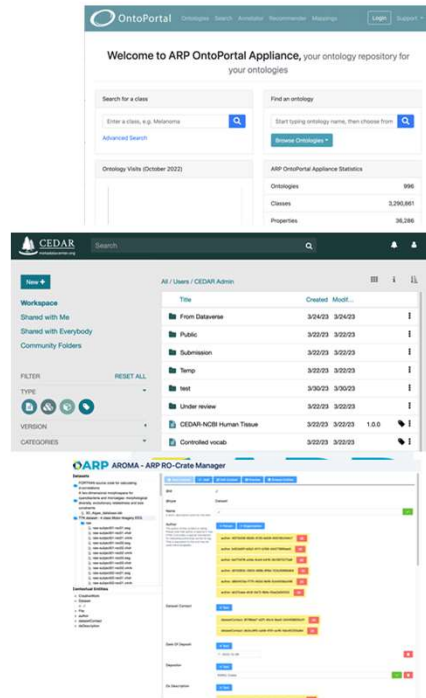
Copyright © 2022 SZTAKI DSD

# Metaadatkezelés moduljai



Copyright © 2022 SZTAKI DSD

# Szoftver komponensek



Copyright © 2022 SZTAKI DSD

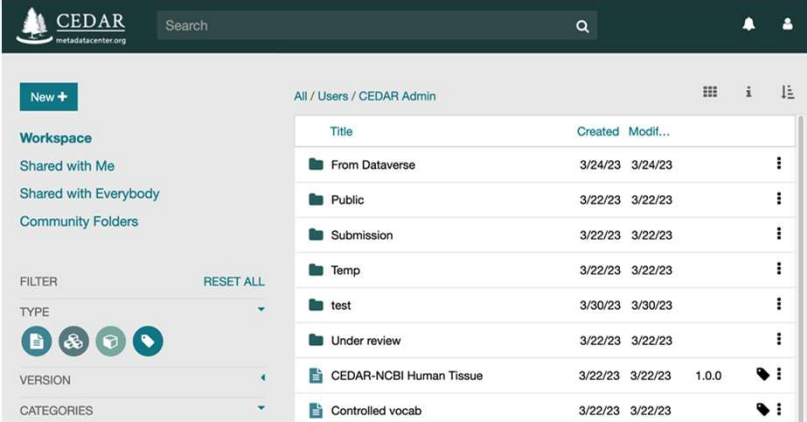
# Dataverse

- Adattár
- Metaadattár
- Kibővített séma kezelés a CEDAR alapokon
- RO-Crate export-import



# CEDAR

- Séma szerkesztő és sématár
- Sémák mezőkből és elemekből (blokkokból) építhetők
- A séma mezői szemantikusan annotálhatók ontológiák alapján



The screenshot shows the CEDAR metadata center interface. The top navigation bar includes the CEDAR logo, a search bar, and user profile icons. The main content area is titled "All / Users / CEDAR Admin" and displays a table of metadata items. The table has columns for "Title", "Created", and "Modif...". The items listed are:

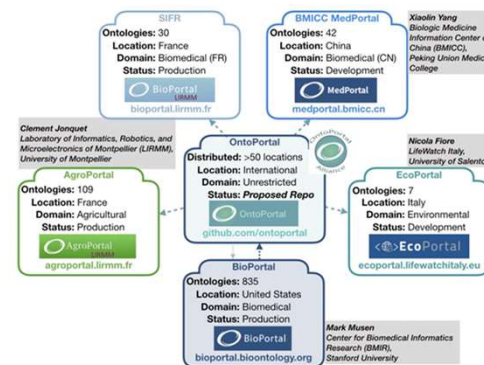
Title	Created	Modif...
From Dataverse	3/24/23	3/24/23
Public	3/22/23	3/22/23
Submission	3/22/23	3/22/23
Temp	3/22/23	3/22/23
test	3/30/23	3/30/23
Under review	3/22/23	3/22/23
CEDAR-NCBI Human Tissue	3/22/23	3/22/23 1.0.0
Controlled vocab	3/22/23	3/22/23



# OntoPortal

- Ontológia tár
- Ontológiák keresése, böngészése
- Témakör specifikus portálok a világban: BioPortal, MedPortal, AgroPortal, EcoPortal, stb.

ARP OntoPortal Appliance Statistics	
Ontologies	996
Classes	3,290,861
Properties	36,286



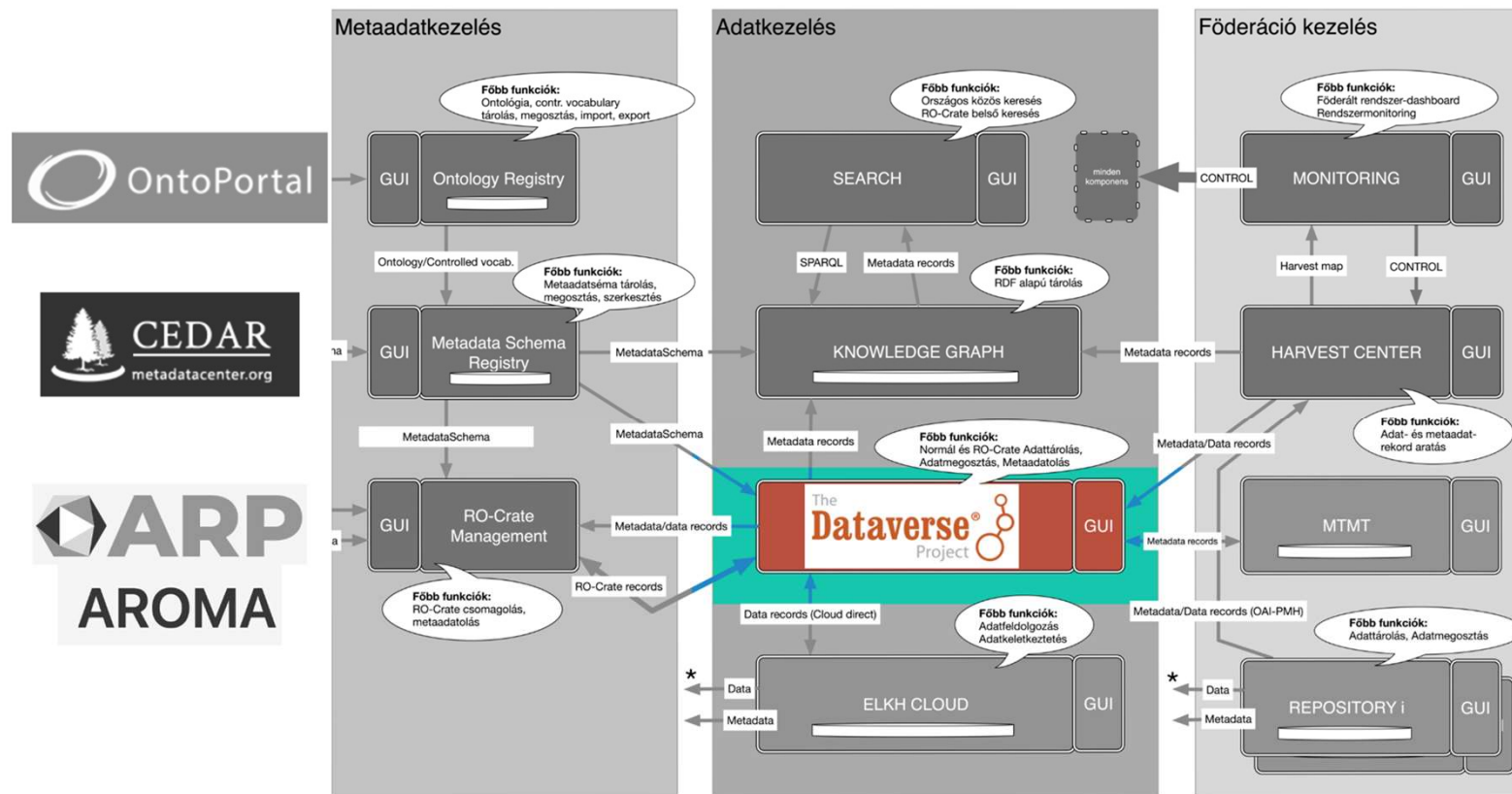


# AROMA - ARP RO-Crate Manager

- A Describo RO-Crate kezelő komponensen alapul
- RO-Crate szabványnak megfelelő metaadatolás támogatása.
- Profile (séma) alapú működés, ami a CEDAR sémákon alapul

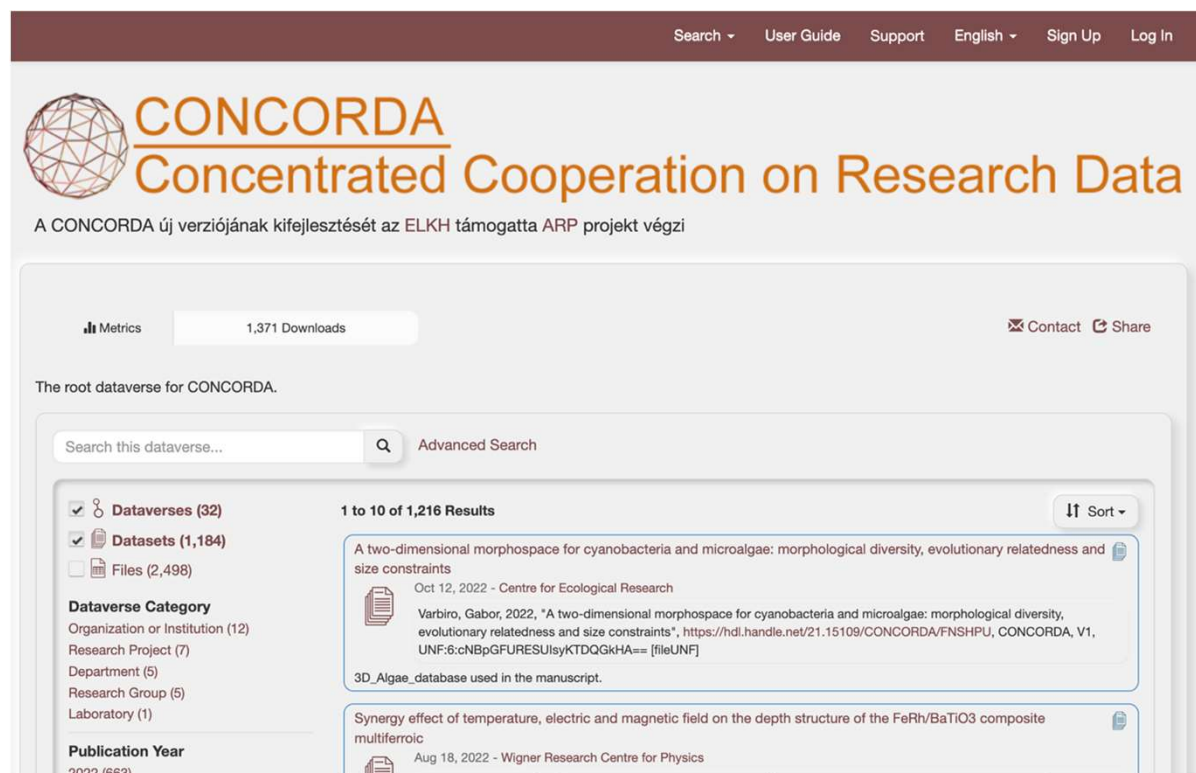
The screenshot displays the ARP AROMA - ARP RO-Crate Manager interface. The left sidebar shows a tree view of datasets, including 'FORTRAN source code for calculating d-correlations', 'A two-dimensional morphospace for cyanobacteria and microalgae: morphological diversity, evolutionary relatedness and size constraints', '3D\_Algae\_database.tab', and 'TK dataset - 4 class Motor-Imagery EEG'. The main area shows the 'Dataset' form with fields for '@id', '@type', 'Name', 'Author', 'Dataset Contact', 'Date Of Deposit', 'Depositor', and 'Ds Description'. The 'Author' field is populated with a list of authors, each with a red 'ORCID' icon. The 'Dataset Contact' field is populated with a list of contacts, each with a red 'ORCID' icon. The 'Date Of Deposit' field is set to '2022-12-08'. The 'Depositor' field is set to 'KIB601, Csaba'. The 'Ds Description' field is empty.

# Use case 1: Adatfeltöltés a Dataverse-be



Copyright © 2022 SZTAKI DSD

# Use case 1: Adatfeltöltés a Dataverse-be



The screenshot displays the CONCORDA Dataverse website. At the top, there is a navigation bar with links for Search, User Guide, Support, English, Sign Up, and Log In. The main header features the CONCORDA logo and the text "Concentrated Cooperation on Research Data". Below this, a message states: "A CONCORDA új verziójának kifejlesztését az ELKH támogatta ARP projekt végzi".

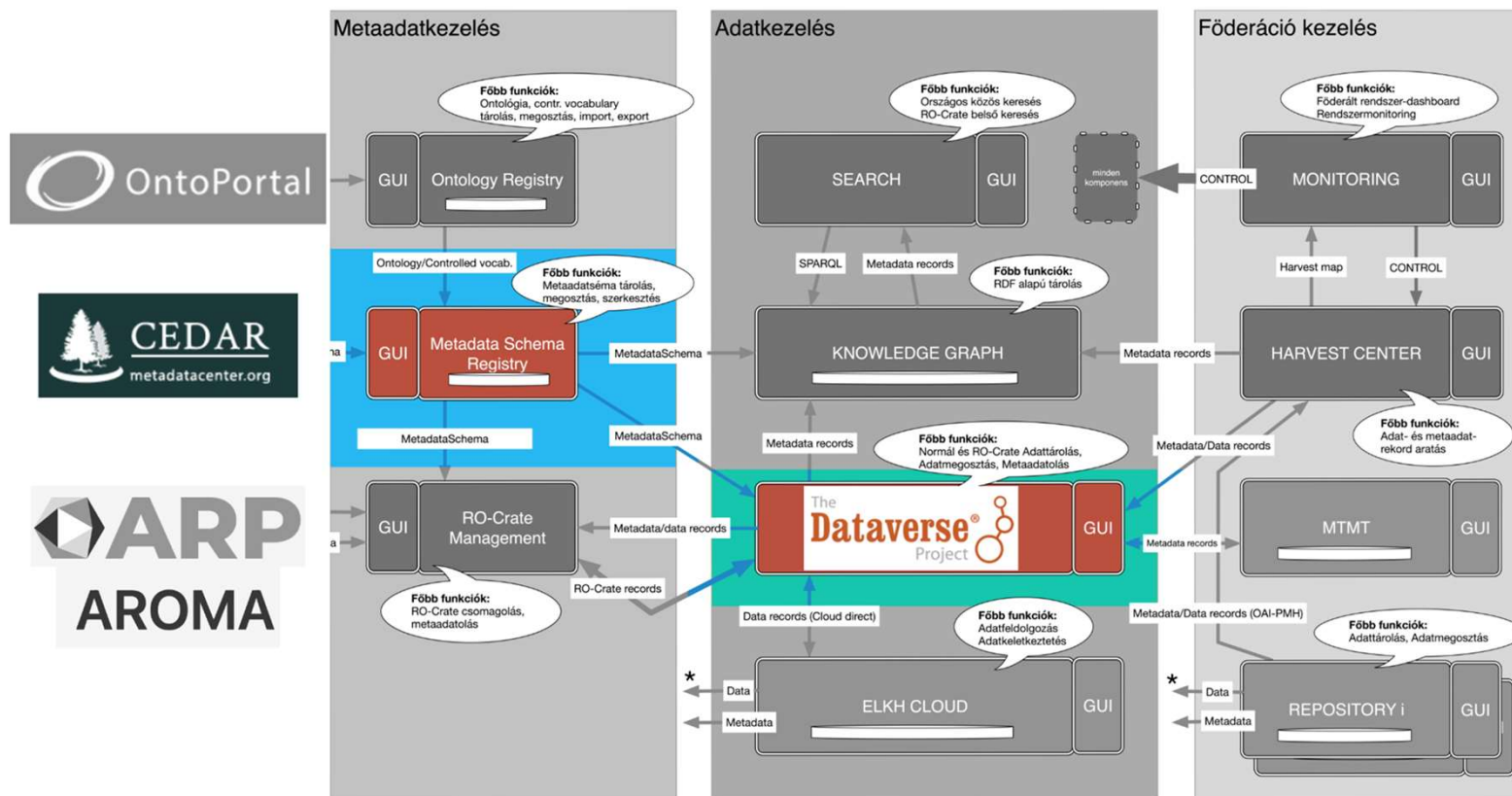
The interface shows a search bar with the text "Search this dataverse..." and a search icon. To the right of the search bar, there are links for "Advanced Search", "Contact", and "Share". Below the search bar, there is a section for "Metrics" showing "1,371 Downloads".

The main content area displays search results for "1 to 10 of 1,216 Results". On the left side, there is a sidebar with filters for "Dataverses (32)", "Datasets (1,184)", and "Files (2,498)". Under "Dataverse Category", there are options for "Organization or Institution (12)", "Research Project (7)", "Department (5)", "Research Group (5)", and "Laboratory (1)". Under "Publication Year", there is an option for "2022 (663)".

The search results list two items:

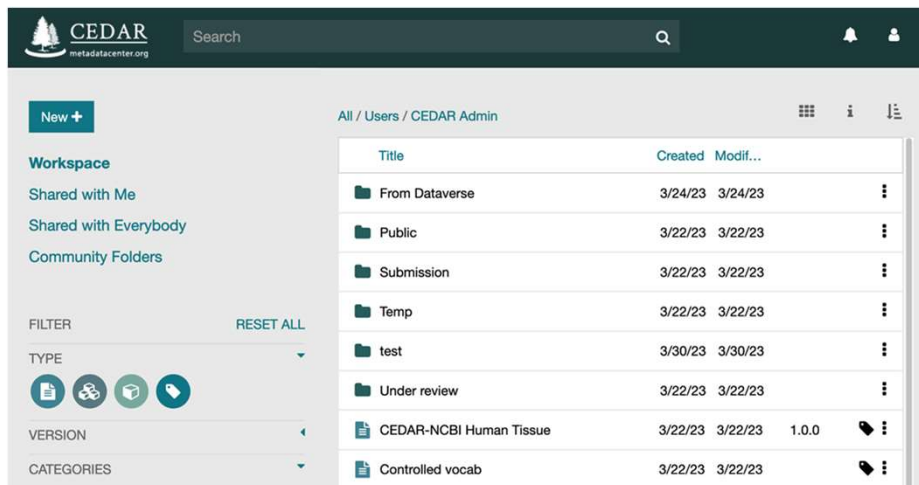
- A two-dimensional morphospace for cyanobacteria and microalgae: morphological diversity, evolutionary relatedness and size constraints**  
Oct 12, 2022 - Centre for Ecological Research  
Varbiro, Gabor, 2022, "A two-dimensional morphospace for cyanobacteria and microalgae: morphological diversity, evolutionary relatedness and size constraints", <https://hdl.handle.net/21.15109/CONCORDA/FNSHPU>, CONCORDA, V1, UNF:6:cNBpGFURESUlsyKTDQGkHA== [fileUNF]  
3D\_Algae\_database used in the manuscript.
- Synergy effect of temperature, electric and magnetic field on the depth structure of the FeRh/BaTiO3 composite multiferroic**  
Aug 18, 2022 - Wigner Research Centre for Physics

# Use case 2: Adatfeltöltés a Dataverse-be saját sémával



Copyright © 2022 SZTAKI DSD

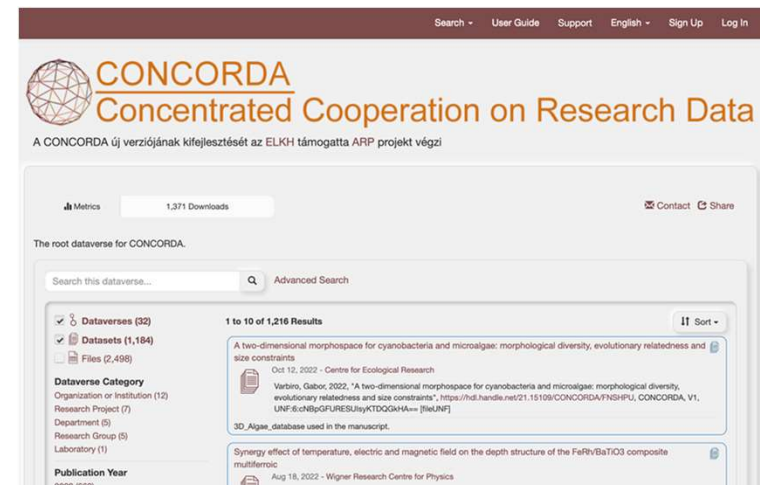
# Use case 2: Adatfeltöltés a Dataverse-be saját sémával



The screenshot shows the CEDAR Dataverse interface. The top navigation bar includes the CEDAR logo, a search bar, and user icons. The main content area displays a list of datasets under the path 'All / Users / CEDAR Admin'. The list has columns for Title, Created, and Modified. The datasets listed are:

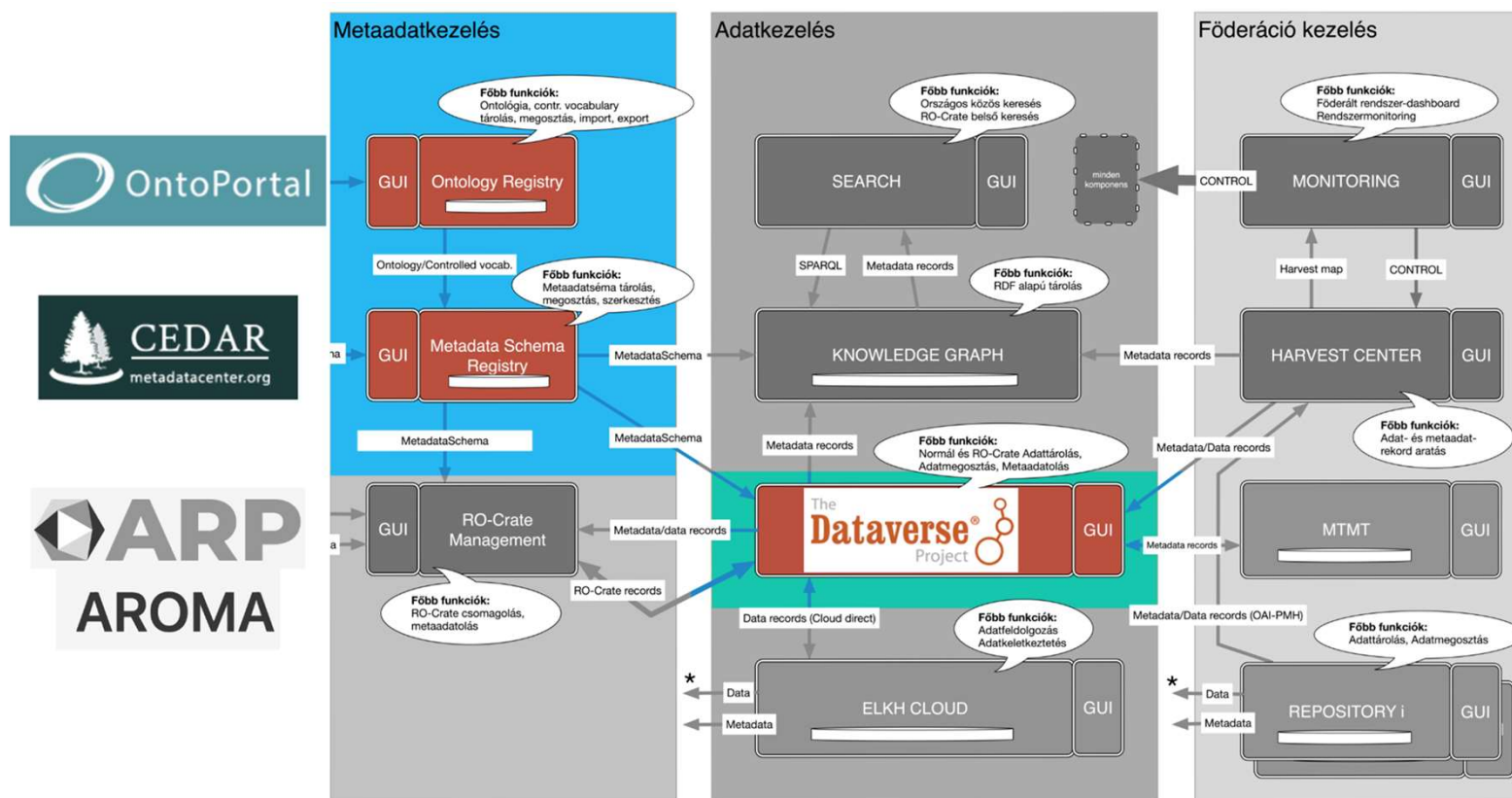
Title	Created	Modif...
From Dataverse	3/24/23	3/24/23
Public	3/22/23	3/22/23
Submission	3/22/23	3/22/23
Temp	3/22/23	3/22/23
test	3/30/23	3/30/23
Under review	3/22/23	3/22/23
CEDAR-NCBI Human Tissue	3/22/23	3/22/23 1.0.0
Controlled vocab	3/22/23	3/22/23

Séma



The screenshot shows the CONCORDA Dataverse interface. The top navigation bar includes the CONCORDA logo, a search bar, and user icons. The main content area displays search results for '1 to 10 of 1,216 Results'. The results are filtered by 'Dataverses (32)', 'Datasets (1,184)', and 'Files (2,498)'. The first result is a dataset titled 'A two-dimensional morphospace for cyanobacteria and microalgae: morphological diversity, evolutionary relatedness and size constraints' by Verber, Gábor, 2022. The second result is a dataset titled 'Synergy effect of temperature, electric and magnetic field on the depth structure of the FeRivBaTK03 composite multiferric' by Wigner Research Centre for Physics, 2022.

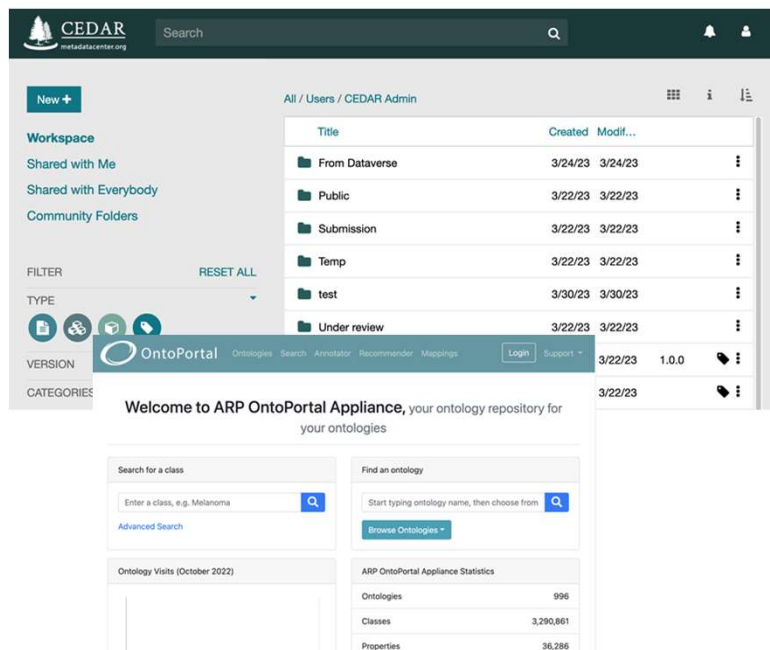
# Use case 3: Adatfeltöltés a Dataverse-be saját sémával és ontológiával



Copyright © 2022 SZTAKI DSD



# Use case 3: Adatfeltöltés a Dataverse-be saját sémával és ontológiával



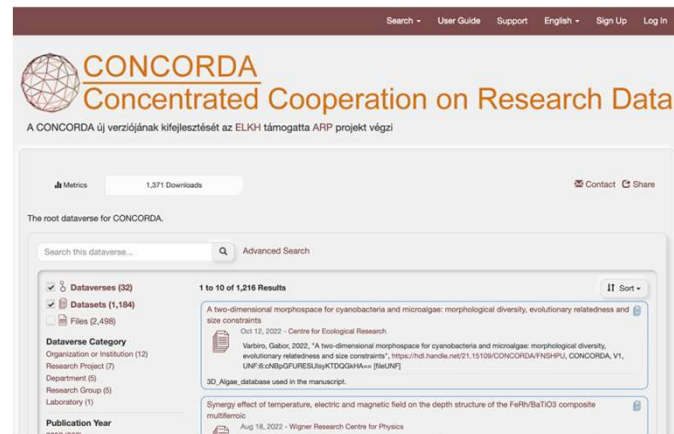
The screenshot shows the CEDAR metadata center interface. At the top, there is a search bar and user information. Below, a sidebar on the left contains navigation options like 'Workspace', 'Shared with Me', and 'Community Folders'. The main area displays a list of folders under 'All / Users / CEDAR Admin'. The folders are:

Title	Created	Modif...
From Dataverse	3/24/23	3/24/23
Public	3/22/23	3/22/23
Submission	3/22/23	3/22/23
Temp	3/22/23	3/22/23
test	3/30/23	3/30/23
Under review	3/22/23	3/22/23

Below the folder list, there is a 'Welcome to ARP OntoPortal Appliance' section with search and browse options. At the bottom right, there are statistics for 'ARP OntoPortal Appliance Statistics':

Category	Count
Ontologies	996
Classes	3,290,861
Properties	36,286

Séma

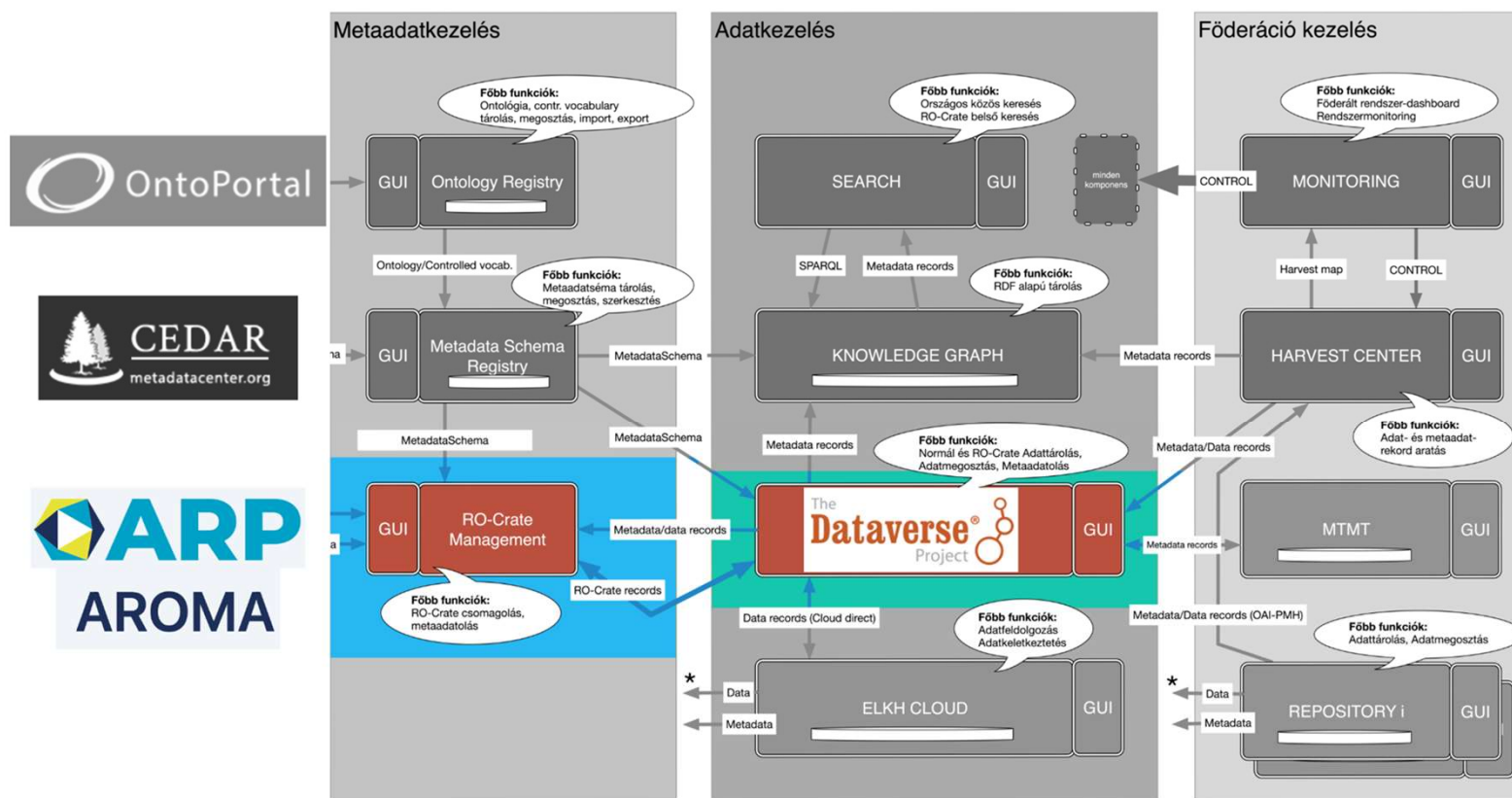


The screenshot shows the CONCORDA interface. At the top, there is a search bar and user information. Below, the header reads 'CONCORDA Concentrated Cooperation on Research Data'. The main area displays search results for a dataset. The results include:

- Dataverse Category:** Organisation or institution (12), Department (5), Research Project (7), Research Group (5), Laboratory (1)
- Publication Year:** 2022 (663)
- 1 to 10 of 1,216 Results:** A two-dimensional morphospace for cyanobacteria and microalgae: morphological diversity, evolutionary relatedness and size constraints. Oct 12, 2022 - Centre for Ecological Research. Várkonyi, Gábor, 2022. "A two-dimensional morphospace for cyanobacteria and microalgae: morphological diversity, evolutionary relatedness and size constraints". [https://hdl.handle.net/21.15108/CONCORDA-P1681PU\\_CONCORDA\\_V1\\_UNF:6:0H5GJFURESJuyK7DQGH4= \[BlUjNF\]](https://hdl.handle.net/21.15108/CONCORDA-P1681PU_CONCORDA_V1_UNF:6:0H5GJFURESJuyK7DQGH4= [BlUjNF]) 3D\_Algaes\_database used in the manuscript.

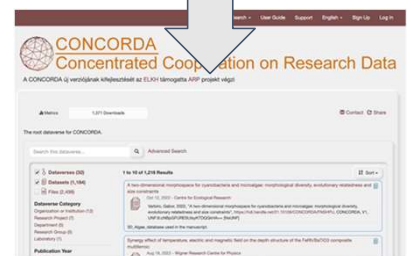
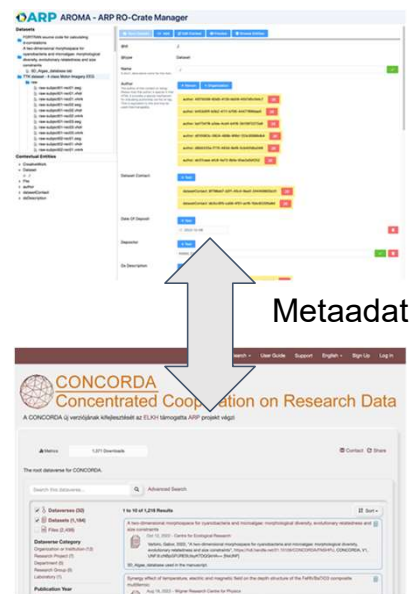
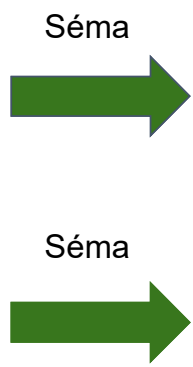
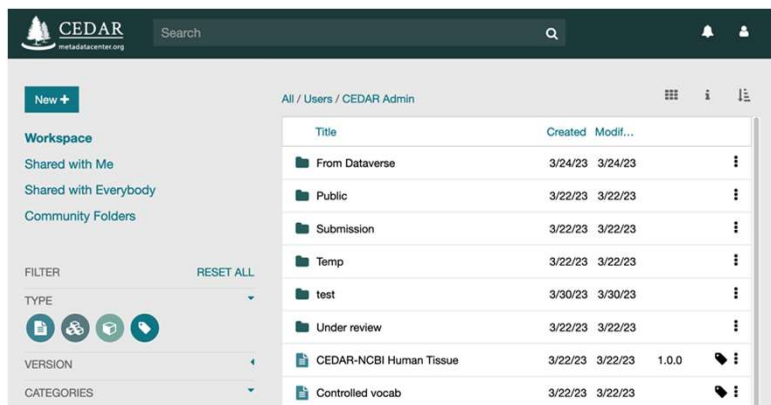


# Use case 4: Bővebb metaadatolás RO-Crate-tel

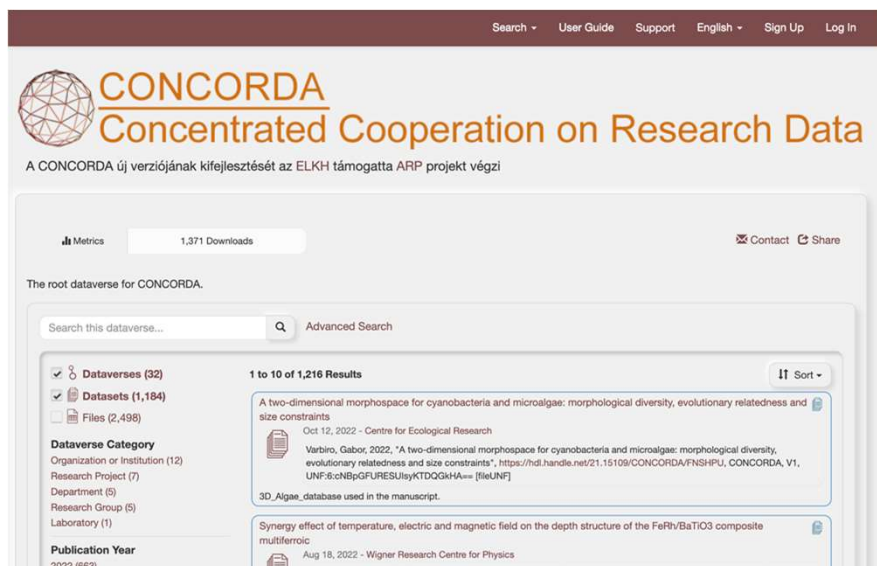


Copyright © 2022 SZTAKI DSD

# Use case 4: Bővebb metaadatolás RO-Crate-tel



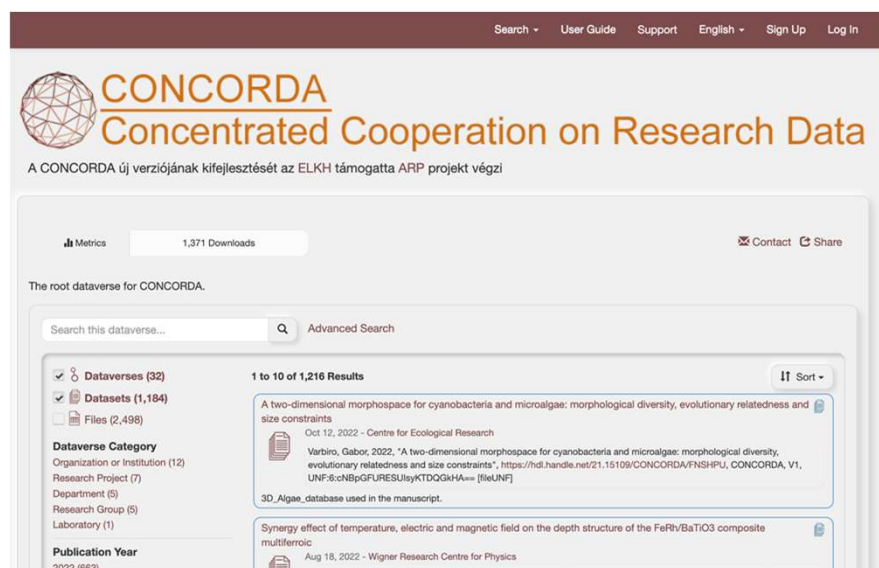
# Use case 5: RO-Crate letöltése Dataverse-ből



RO-Crate csomag



# Use case 6: RO-Crate csomag importja Dataverse-be



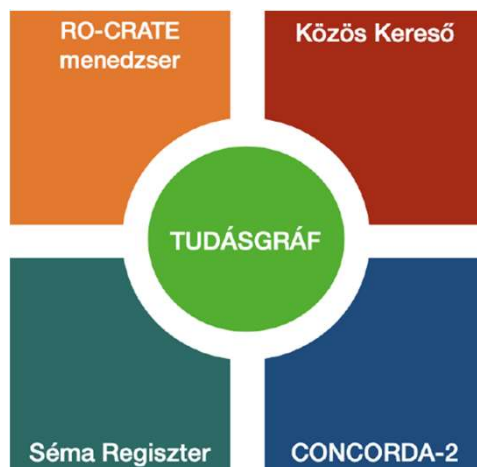
The screenshot shows the CONCORDA Dataverse website. The header includes the CONCORDA logo and the text "Concentrated Cooperation on Research Data". Below the header, there is a search bar and a navigation menu. The main content area displays search results for datasets, with a sidebar on the left for filtering by category and publication year. The results list includes a dataset titled "A two-dimensional morphospace for cyanobacteria and microalgae: morphological diversity, evolutionary relatedness and size constraints" by Varbito, Gabor, 2022.

RO-Crate csomag



# Összefoglalás

- **Központosított metaadatséma** tár, egyszerű vizuális séma szerkesztővel, multidiszciplináris felhasználásra
- **Szemantikusan annotált sémák** támogatása, ami a tudásgráfban való tárolást és keresést segítik
- **RO-Crate alapon bővebb metaadat-kezelés** fájlok esetén, melyek a tudástárban kereshetővé válnak
- **RO-Crate formátum export-import** támogatása a nyílt adatmozgatás támogatására



SZTAKI - <https://www.sztaki.hu/>

ELKH ARP honlap - <https://science-research-data.hu/>

ELKH ARP repozitórium - <https://science-data.hu/>

**Pataki Balázs**  
**SZTAKI – DSD**  
**pataki@sztaki.hu**

