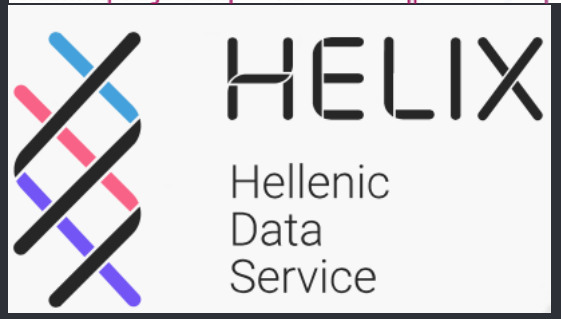
Hellenic Data Service (HELIX) and HARDMIN research data repository for the Academic institutions in Greece ORCID Integration

Michail Alexakis
Athena Research Center
Research Fellow



HELIX - Hellenic Data Service

- Scientific e-Infrastructure for data-intensive research
  - www.hellenicdataservice.gr
- Supports the full lifecycle of scientific data management, processing, sharing, and reuse
- Inherently scalable, cloud-based
- Nation-wide, horizontal, cross-domain
- Multiple roles
  - Open Access/Science
  - FAIR Data
  - Public Data
  - Industrial Data Platform

The 3 pillars of HELIX

- HELIX Pubs
  - Nation-wide, cross-domain discovery of publications
  - Adapt and localize OA OpenAIRE CRIS services
- HELIX Data
  - Data catalogue and repository for FAIR scientific and industrial data
  - Discover, collect, evaluate, download, and use
- HELIX Labs
  - Generic-purpose and domain-specific services and APIs for data analysis, processing, and experimentation

HEAL-Link HARDMIN

- Research Data infrastructure interconnected with HELIX and European Open Science infrastructures
- Collect and curate research data from Academic Institutions in Greece
- Complies with the European principles on research data (FAIR data)
  - Findable
  - Accessible
  - Interoperable
  - Reusable
- Roadmap
  - Currently in development and testing with select HEAL-Link members
  - First public beta available in November 2019
  - Production version expected in February 2020

HARDMIN Features

- Data catalogue and repository for scientific data
  - Repurposes and extends the HELIX Data Catalogue software (economies of scale)
  - Data assets harvested by the HELIX Data Catalogue and integrated to the HELIX service ecosystem (interoperability-by-design)
- Available directly and as-a-service to HEAL-Link's members, i.e., Greek Academic Institutions
  - Harvests data from institutional data catalogues/repositories of HEAL-Link members
  - White-labelled on-demand virtual data catalogues and repositories for HEAL-Link members; data assets automatically harvested to HEAL-Link
- Nation-wide provision of DOIs via DataCite
- ORCID integration
- Deployed at the IT Center of Aristotle University of Thessaloniki

ORCID integration

- Goal
  - Enable HEAL-Link members to be authenticated and authorized using their **Institutional** or **ORCID** iDs interchangeably
  - Built-in support for additional IDPs (e.g., github, google)
- **DELOS** Federation (aai.grnet.gr)
  - National AAI Federation of the academic, research and educational community
- **Keycloak** (keycloak.org)
  - Identity and access management software
  - Deployed to manage AA requests/flows across multiple applications
- **CKAN** (ckan.org)
  - Data catalogue/repository software
  - HARDMIN is based on CKAN

Integration Flow

- Initial user sign-up and authentication via DELOS (institutional accounts)
  - User added to an organization by the organization's admin
  - Non-automated flow in alignment with publishing practices
- User links her institutional ID with ORCID iD
  - The user can optionally link her ORCID iD via the My Account settings page of the catalogue
  - Catalogue invokes the ORCID member API
  - OAuth-based authentication via ORCID; Catalogue authorized by user
  - Catalogue stores and links ORCID iD
  - Catalogue updates user record with her profile in the Catalogue (optional)
- Sign-in
  - User can provide institutional or ORCID iD

Roadmap & Issues

- Testing and PoC’s complete
- The only technical issue has been Keycloak’s handling of JWT tokens
  - <Expired timestamp>; modeled as integer
  - Expected to be addressed in a new version soon. If not:
- Contingency
  - Extend CKAN functionality for multiple IDPs
    - Keycloak + ORCID
  - WSO2 deployment

Thank you

https://hardmin.heal-link.gr