Detailseite

DataPLANT Research Data Management and Publication Platform (ARChub) (ARChub)

Die DataPLANT Research Data Management and Publication Platform (ARChub) ist eine auf der Cloud- und Speicherinfrastrukturen aufbauender Forschungsdatenmanagement-Dienst, der als universelles ScienceGateway die Zusammenarbeit von Einzel über Labor bis hin zu interdisziplinären Forschungsgruppen auf Basis des Annotated Research Contexts unterstützt. Es kombiniert die Open Source Frameworks GitLab und InvenioRDM in einer für das FDM angepassten Form mit verschiedenen Workflow-Systemen (der Bio-Informatik), wie beispielsweise Galaxy. Durch die eingebaute GitVersionierung, IssueTracking, CI/CD und Workflow-Integration ermöglicht der ARChub lückenloses ProvenienzTracking, verlässliches Teilen von Forschungsdaten und ein simples ProjektManagement, das den gesamten DatenLebenszyklus (Erzeugung, Annotation, Versionierung, Veröffentlichung) abdeckt. Die Anbindung an darunterliegende stabile Speichersysteme liefert eine skalierbaren, gesicherte Grundlage für große Datensätze, während die integrierten REST und WebHooks eine nahtlose Einbindung in bestehende ForschungsWorkflows und automatisierte DatenPipelines erlauben. InvenioRDM ist eine weitverbreitete Datenpublikationsplattform, die am CERN entwickelt wird und als Grundlage für die mit persistenten Identifiern (DOI) veröffentlichten Datensätze dient. Die cloud-basierte Plattform ist offen angelegt, so dass neben Galaxy weitere Workflow-Systeme integriert werden können.

Adresse: Hermann-Herder-Str. 10
79104 Freiburg
Baden-Württemberg
Deutschland
Zur Webseite

Träger

Universität Freiburg
Friedrichstraße 39
79098 Freiburg
Baden-Württemberg
Deutschland
https://uni-freiburg.de/

Wissenschaftsgebiet

Hauptgebiete:
- Biologie
- Agrar-, Forstwissenschaften, Gartenbau und Tiermedizin

Nebengebiete:
- Medizin
- Geowissenschaften (einschl. Geographie)

Kategorie

Forschungsdaten-Repositorien

Wissenschaftliche Dienstleistungen

1. Bereitstellung einer stabilen und langfristigen Speicherplattform für (größe) Datenmengen im Terabyte++ Bereich 2. Betrieb einer Datenmanagement- und Datenpublikationsplattform mit Zugang über (Community) AAI 3. Vergabe, Registrierung und Auflösung langfristiger persistenter Identifier (DOI) (für ARC RO-Crates) 4. Bereitstellung von APIs für komplexe Suchen, AI-Anbindung, Datenföderation 5. Anbindung an Workflow-Systeme wie Galaxy, Unterstützung bei Anbindung an weitere Infrastrukturen 6. Ausbildung und Training: Einführung in die Nutzung der Plattform, Trainings zum Forschungsdatenmanagement auf Basis des Annotated Research Contexts, Beratung zum Datenmanagement, Datenveröffentlichung und -nachnutzung, Beratung zur Datenaufbereitung, -paketierung, Datensparsamkeit, Lizenzen 7. Supportangebote: * Erstellung einer projektspezifischen Arbeitsumgebung und langfristige Ausgestaltung * Einbindung von Labor- und arbeitsgruppenspezifischen Workflows * Einbindung neuer Analysetools (in Galaxy oder weiteren Workflow-Systemen) * Unterstützung beim Qualitätsmanagement der Daten und deren Veröffentlichung * Nutzung und Weiterentwicklung des ARC-Schemas für eigene Anwendung

Wissenschaftliche Geräte

- bwSFS-Speichersystem: Netzwerk-Shares und Storage-Grid (S3)
- de.NBI-Infrastruktur (Server für Galaxy-Workflows)
- bwCloud (Virtuelle Maschinen für (Micro-)Services)
- Virtualisierungs-Cluster
- HPC System

Schlagworte

- Datenmanagementplattform
- Datenpublikationssystem
- Galaxy Workflows
- Annotated Research Context
- ARC RO-crate
- Bioinformatische Tools

Netzwerke

NFDI-Konsortium DataPLANT
https://nfdi4plants.org/

de.NBI - Deutsches Netzwerk Bioinformatik
https://www.denbi.de/

Galaxy Europe
https://usegalaxy.eu/

Nutzer/Jahr

Interne Nutzer: ~ 30
Externe Nutzer gesamt: ~60
Externe Nutzer in Deutschland: 50+
Externe Nutzer im europ. Ausland: ~5
Externe Nutzer außerhalb Europas:

© DFG