



Contribution ID: 17

Type: Poster

DBRepo: A Database Repository to Support Research Activities

Tuesday, 20 September 2022 19:00 (1 hour)

Many institutions start having dedicated data stewards curate data in close collaboration with researchers who collect, compute or distribute the data (e.g. as part of supplementary material to a journal article). Contrary to traditional data dumps, this is a challenge for structured data in databases where data evolves over time as tuples are added in data streams, updated or deleted. Outside of large-scale infrastructures designed to host e.g. climate or genome data, researchers usually have to maintain their own, local database and take care of regular software updates, configurations and feeding data, before being able to do research. Curation activities such as collecting metadata or preservation, if at all, happen only after the project is finished when the database is exported to a file repository turning it into a static dump that cannot be trivially queried anymore.

We present DBRepo, a repository for relational databases in a private cloud setting to support research activities in four dimensions: (1) keep research data in relational databases from the beginning of a project and offer application programming interfaces to access the data; (2) provide separation of concerns that allows experts to handle database management tasks and let researchers focus on conducting their research work; (3) improve FAIRness of data (Findability by collecting ontology-mapped metadata centrally and issuing persistent identifiers to queries; Accessibility by providing HTTP/AMQP/JDBC protocols; Interoperability by mapping to controlled vocabularies; and Reusability by offering metadata and attaching a license to each database); and (4) support reproducibility and persistent identification of arbitrary subsets of data by implementing the RDA WGDC recommendations. DBRepo's source code is available in GitLab (<https://gitlab.phaidra.org/fair-data-austria-db-repository/fda-services>), we also operate a public demo instance (<https://dbrepo.ossdip.at>).

Any relevant links

<https://gitlab.phaidra.org/fair-data-austria-db-repository/fda-services>

<https://dbrepo.ossdip.at/>

<https://dbrepo-docs.ossdip.at/>

Topic

Data Spaces

Primary authors: WEISE, Martin (TU Wien); Mr STAUDINGER, Moritz (TU Wien); Mrs MICHLITS, Cornelia (TU Wien); Mrs GERGELY, Eva (University of Vienna); Mr STYTSENKO, Kirill (University of Vienna); Mr GANGULY, Raman (University of Vienna); Prof. RAUBER, Andreas (TU Wien)

Presenter: WEISE, Martin (TU Wien)

Session Classification: Posters (presenters at poster)

Track Classification: Data Spaces