

## ***Breaking Down Barriers: The Transition of ODIS from a Relational to a Triple Store Database***

**Bert Aernouts and Joris Colla (KADOC-KU Leuven / ODIS)**

ODIS is a contextual relational database on the history of civil society. It is used by a growing number of research and heritage organizations in Flanders and Brussels and aims to stimulate cross-fertilization between humanities researchers and the custodians of cultural heritage collections. The database was set up in 2000-2003 thanks to a grant of the Research Foundation Flanders (FWO). It is administered by the non-profit association ODIS vzw, in which the universities of Antwerp (UA), Brussels (VUB), Ghent (UGent) and Leuven (KU Leuven) are represented. Its daily management is ensured by KADOC, the Documentation and Research Centre on Religion, Culture and Society of KU Leuven.

ODIS, offering content both in Dutch and English, proves the added value of a joint authority database for the contextual disclosure of analog and digital heritage collections. The members of the ODIS partnership use the system both to provide basic information on their documentary heritage (archives, periodical publications) and to build elaborate data sets on organizations, persons, families, buildings and events related to that heritage, with (deep)links to and from primary catalogs and digital repositories. By means of the database's search functions and tools, scholars can analyze its content. But they can also use ODIS to store, pool, validate, publish and/or analyze their own data sets. By doing so, they preserve them from oblivion and guarantee the reproducibility of their research. At the same time, they enrich the data series developed by heritage organizations.

In 2022, ODIS has embarked on a transformative journey, propelled by a research infrastructure grant of FWO and benefiting from a technical collaboration with LIBIS-KU Leuven. At the core of this 'ODIS renewal project 2022-2027' is the development of a *new triple store database* and of a *new public interface (OPAC)*. This stems from a desire to:

- 1) Develop *new ways to query, analyze and visualize data sets* (such as network visualization and geographical visualization), allowing the growing number of scholarly users of ODIS to respond to the challenges of present-day digital humanities research.
- 2) Develop an *interoperability and discovery layer*, offering more and better connections with other catalogs, research instruments, platforms and Linked Open Data resources. This is the cornerstone of a durable open access policy, based on the FAIR principles. It will also facilitate the semantic enrichment of the ODIS content through interconnections with complementary data sets.

The transition from a relational database to a Virtuoso triple store is a challenging procedure. It involves organizing all fields and objects that form the data model in a Google Sheets spreadsheet, taking into account the multilingualism of ODIS. This sheet is then used to generate SHACL files, which in turn are used to raise the architecture of the triple store database and the JSON API that will allow interoperability between the database, the new frontend and third-party users. The Google Sheets spreadsheet acts as a bridge between the functional and the technical side of the project: it allows people with more functional knowledge of the database to interact with the model and help shaping it in an accessible way. This approach also allows easy adjustments to the model in the first stages of setting up the new database environment. Indeed, the process provides the opportunity to reassess the

current metadata model and update certain fields. The model used in the current ODIS database consists of elements of ISAAR(CPF), ISAD(G) and ISDAIAH. Those standards have been welded together by the International Council on Archives (ICA) into the linked data ontology Records in Context (RiC). In the redesigning of the model, compliance with RiC and with the Open Standards for Linking Organizations (OSLO) of the Flemish authorities is an important aspect.

The shift to a triple store approach allows to break down barriers between ODIS and other data repositories, but also between the separate silos within the database. ODIS consists of 8 modules: organizations, persons, families, buildings, events, publications, archives and repositories. By redesigning the data model, those separate modules are opened up for more (inter)connectivity in the future.

Together with redesigning the database structure, the public interface of ODIS is given a revamp. Apart from updating the search and filter functionalities, we are aiming to facilitate a network visualization of the relations between records. This will open up new ways of presenting the vast amount of validated data in ODIS and will allow both researchers and heritage organizations to approach the data from multifold angles, facilitating new insights. Furthermore, the investment program aims to develop tools for the geographical analysis and visualization of the ODIS data series. External authorities such as OpenStreetMap are used to refine the existing geographical thesaurus.

To date, the data model is finished and the data conversion is in full swing. This is happening module by module because each module poses its own specific set of challenges. Work on the public layer has started with wireframes and mock-ups for the data input module as well as for the search engine and result list view. The new public interface will gradually be released for rigorous testing within the ODIS community. The network visualization and the geographical tools are planned for a later stage since more in-depth research is necessary for their development and implementation.

In our contribution at the DH Benelux Conference, we will briefly introduce ODIS and its renewal project 2022-2027. Our main focus will be the transition from a relational to a triple store database as a way to foster interconnectivity, highlighting the main challenges and pitfalls of this process.

**Bert Aernouts** (KADOC-KU Leuven) is collaborator of the ODIS renewal project 2022-2027 and expert in digital data.

<https://www.kuleuven.be/wieiswie/nl/person/00128460>

<https://orcid.org/0009-0001-2042-6397>

**Joris Colla** (KADOC-KU Leuven) is data manager of ODIS, coordinator of the ODIS renewal project 2022-2027 and heritage expert.

<https://www.kuleuven.be/wieiswie/nl/person/00064754>

<https://orcid.org/0000-0003-3826-9593>