NFDI4Objects

Research Data Infrastructure
for the Material Remains of
Human History

Community Cluster

Data Caption and Creation

Chairs: GERALD STEILEN (VERBUNDZENTRALE DES GBV); ANJA GERBER (KLASSIK STIFTUNG WEIMAR)

SC-Beschluss: 24.05.2024

Beteiligte Organe: TA1, TA2, TA4, TA5, TA6

Externe Anbindung: NFDI4Culture, NFDI4Memory, Text+, BASE4NFDI, Uni Göttingen, Berlin-Brandenburgische Akademie der Wissenschaften

Subject and Objective

The Community Cluster is a group of experts who deal with the generation, collection and description of data from a technical perspective in order to prepare it for the Knowledge Graph. The group works on the optimisation of data collection processes, the standardisation of data formats and their documentation. In addition, the possibilities and necessity of automatic data validation will be analysed. The advantages and disadvantages of different data formats with regard to criteria such as data integrity, long-term archiving, accessibility and interoperability are also being analysed. The collaboration of experts with complementary skills enables the group to cover a broad spectrum of knowledge and develop in-depth solutions to complex challenges.

The close collaboration with the TWG 'Development of a common N4O Objects Ontology (N4O OO) and a Minimal Metadata-Set (N4O MMDS)' is ensured by the joint participation of chairs and members in both groups.

Out of Scope

The Community Cluster is primarily concerned with issues relating to the generation, collection and description of data from a technical perspective in order to be able to process this data automatically for a knowledge graph. The development of ontologies, thesauri or standardised data is not one of the tasks of this working group. The semantic representation of knowledge is also not taken into account.

_____

Ensuring coherence between the technical and semantic levels is therefore not at the centre of the range of tasks. Software or sensor solutions for digital data collection, processing and management are not analysed, particularly with regard to their suitability and possible adaptation requirements to the specific needs of this consortium. Furthermore, neither support nor training on the use of software solutions is offered.

Intended Work Results

The group focuses on optimising data collection processes in terms of efficiency, accuracy and data quality. This results in five main objectives of the Community Cluster:

1. evaluating the advantages and disadvantages of formats with regard to criteria such as data integrity, long-term archiving, accessibility and interoperability.

2. to establish contacts and collaborate with national and international initiatives for the standardisation of data formats in the cultural heritage sector

3. developing best practices for data conversion and migration..

4. writing white papers summarising the results of the analysis and evaluation of data formats.

5. producing blue papers summarising detailed guidelines and standards for data formatting and storage in the cultural heritage sector.