



**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

### Thema / Zielsetzung

Das Community Cluster ist eine Gruppe von Expert\*innen, die sich mit der Erzeugung, Erhebung und Beschreibung von Daten aus technischer Sicht befassen, um diese für den Knowledge Graph aufbereiten zu können. Die Gruppe arbeitet an der Optimierung von Datenerhebungsprozessen, der Standardisierung von Datenformaten sowie deren Dokumentation. Darüber hinaus werden die Möglichkeiten bzw. die Notwendigkeit einer automatischen Datenvalidierung beleuchtet. Die Vor- und Nachteile verschiedener Datenformate im Hinblick auf Kriterien wie Datenintegrität, Langzeitarchivierung, Zugänglichkeit und Interoperabilität werden ebenfalls untersucht. Durch die Mitarbeit von Expert\*innen mit komplementären Fähigkeiten kann die Gruppe ein breites Spektrum an Wissen abdecken und tiefgreifende Lösungen für komplexe Herausforderungen entwickeln.

Die enge Zusammenarbeit mit der TWG "Development of a common N4O Objects Ontology (N4O OO) and a Minimal Metadata-Set (N4O MMDS)" wird durch die gemeinsame Beteiligung von Chairs und Mitgliedern in beiden Gruppen sichergestellt.

### Abgrenzung

Das Community Cluster beschäftigt sich primär mit Fragen der Generierung, Sammlung und Beschreibung von Daten aus technischer Sicht, um diese auch für einen Wissensgraphen automatisiert aufzubereiten zu können. Die Entwicklung von Ontologien, Thesauri oder Normdaten gehört nicht zu den Aufgaben dieser Arbeitsgruppe. Ebenso bleibt eine semantische



**NFDI4Objects**

Research Data Infrastructure  
for the Material Remains of  
Human History

**Community Cluster**

Repräsentation von Wissen unberücksichtigt. Somit steht eine Sicherstellung der Kohärenz zwischen technischer und semantischer Ebene nicht im Mittelpunkt des Aufgabenspektrums. Software- oder Sensorlösungen zur digitalen Datenerhebung, -verarbeitung und -verwaltung werden nicht analysiert, insbesondere nicht hinsichtlich ihrer Eignung und möglicher Anpassungserfordernisse an die spezifischen Bedürfnisse dieses Konsortiums. Darüber hinaus werden weder Support noch Schulungen zur Nutzung von Softwarelösungen angeboten.

### **Angestrebte Arbeitsergebnisse**

Die Gruppe fokussiert sich auf die Optimierung der Datenerhebungsprozesse im Hinblick auf Effizienz, Genauigkeit und Datenqualität. Daraus ergeben sich fünf Hauptziele des Community Clusters:

1. Bewertung der Vor- und Nachteile von Formaten im Hinblick auf Kriterien wie Datenintegrität, Langzeitarchivierung, Zugänglichkeit und Interoperabilität.
2. Kontaktaufnahme und Zusammenarbeit mit nationalen und internationalen Initiativen zur Standardisierung von Datenformaten im Kulturerbebereich.
3. Entwicklung von Best Practices für die Datenkonversion und -migration.
4. Verfassen von White Papern, die die Ergebnisse der Analyse und Bewertung von Datenformaten zusammenfassen.
5. Erstellung von Blue Papern, die detaillierte Richtlinien und Standards für die Datenformatierung und -speicherung im Kulturerbebereich enthalten.