

MUSE4Anything

Ontologiebasierte Generierung von Werkzeugen zur strukturierten Erfassung von Daten

Bühler, Fabian

fabian.buehler@iaas.uni-stuttgart.de
Institut für Architektur von Anwendungssystemen (IAAS) -
Universität Stuttgart

Barzen, Johanna

johanna.barzen@iaas.uni-stuttgart.de
Institut für Architektur von Anwendungssystemen (IAAS) -
Universität Stuttgart

Leymann, Frank

Leymann@iaas.uni-stuttgart.de
Institut für Architektur von Anwendungssystemen (IAAS) -
Universität Stuttgart

Standl, Bernhard

bernhard.standl@ph-karlsruhe.de
Institut für Informatik und digitale Bildung - Pädagogischen
Hochschule Karlsruhe

Schlomske-Bodenstein, Nadine

nadine.schlomske-bodenstein@ph-karlsruhe.de
Institut für Informatik und digitale Bildung - Pädagogischen
Hochschule Karlsruhe

Einleitung

Zur Unterstützung der Erforschung von Text und Sprache als einer der Kernbereiche der Digital Humanities (Reichert 2017: 13), haben sich bereits verschiedene Methoden und Werkzeuge etabliert, wie EU Infrastrukturprojekte wie CLARIN (Hinrichs/Krauer 2014) verdeutlichen. Um sich darüber hinaus Artefakten wie beispielsweise Filmen, Musik, Gemälden oder Architekturen, aber auch abstrakten Artefakten wie Lehr-Lern-Szenarien, mittels digitaler Ansätze zu nähern, ist oftmals die Überführung dieser Artefakte in maschinenlesbare Repräsentationen nötig, was aufwendige und allzu häufig kostspielige Individuallösungen erfordert. Dieser Problematik wollen wir mit der hier vorgestellten Werkzeugunterstützung “MUSE4Anything” begegnen. MUSE4Anything soll die strukturierte und detaillierte Erfassung von unterschiedlichen Artefakten unterstützen. Hierzu erlaubt MUSE4Anything angepasste Eingabemasken zur Erfassung der relevanten Parameter zu erstellen. Diese werden automatisch aus der von dem Benutzer erstellten Ontologie (Ontologie nach dem Verständnis von Furrer (2014: 308-309) als Werkzeug der Wissensrepräsentation) generiert.

Für die Anforderungsanalyse als Basis von MUSE4Anything haben wir die domänenspezifischen Repositorien aus unseren

langfristig laufenden Projekten MUSE¹, welches vestimentäre Kommunikation im Film untersucht (Barzen et al. 2018a, 2018b) und MUSE4Music², das sich mit symphonischer Musik des 19. Jahrhunderts befasst (Barzen et al. 2017), herangezogen.

Anforderungsanalyse: MUSE und MUSE4Music

MUSE und MUSE4Music zielen auf die Identifikation von Mustern (nach dem Musterbegriff von Alexander et al. (1977)), um zu einem besseren Verständnis des jeweiligen Untersuchungsgegenstandes beizutragen. Dazu müssen in großem Umfang Daten erfasst, analysiert und interpretiert werden. Das Vorgehen und die dazugehörige Werkzeugumgebung unterstützen dabei: (i) die Definition der potenziell relevante Parameter mittels Ontologie, (ii) die Auswahl der Kriterien zur Zusammenstellung des Untersuchungsgegenstandes, (iii) die Erfassung der Daten mittels des Repositoriums (iv) die, auf die Art und Struktur der Daten abgestimmte Analyse der Daten und (v) die Überführung der validierten Ergebnisse in Muster (Barzen et al. 2018b).

Beide Projekte nutzen Mind-Map Programme, um die Ontologie zu verwalten (Schritt (i)) und ein jeweils speziell entwickeltes Datenrepositorium zur Erfassung der Daten (Schritt (iii)). Der in MUSE erfasste umfangreiche Datensatz (Barzen et al. 2021b), wird aktuell mit verschiedenen Ansätzen aus den Bereichen des Machine Learning erschlossen (Barzen 2021a).

Um das hier erarbeitete Vorgehen für weitere Anwendungsfälle zu erschließen, haben wir die vorhandenen Implementierungen systematisch analysiert und Kernanforderungen an ein generisches Werkzeug extrahiert. Diese beinhalten z. B. die Möglichkeit komplex strukturierte und teilweise tief verschachtelte Eigenschaften zu erfassen und die Ontologie direkt im Werkzeug zu erstellen und zu bearbeiten (Bühler 2021: 9–18). Zusätzlich haben wir existierende Lösungen zur Verwaltung von Ontologien, wie Protégé (Musen 2015), OWLGrEd (OWLGrEd 2021) oder VocBench (Stellato et al. 2020), evaluiert (Bühler 2021: 19–24) und uns in Hinblick auf die identifizierten Anforderungen für eine eigene Lösung entschieden.

MUSE4Anything

Auf Basis der identifizierten Anforderungen haben wir MUSE4Anything entwickelt. MUSE4Anything (Bühler 2021) ist ein generisches Datenrepositorium und unterstützt die Schritte (i) und (iii) aus dem vorhin beschriebenen MUSE-Vorgehen (Abbildung 1). Die benutzerdefinierte Ontologie zur Definition der Domäne, sowie die ggf. enthaltenen Taxonomien, können in dem Werkzeug erstellt (Schritt (i)) und jederzeit bearbeitet werden (Abbildung 2). Aus der Ontologie werden automatisch korrespondierende Eingabemasken für die Datenerfassung (Schritt (iii)) und Detailansichten der Objekteigenschaften generiert (Abbildung 3).

Alle Funktionen sind über eine Web-Oberfläche und mit der HTTP-API nutzbar. Dabei wurde im Besonderen für die Datenerfassung auf die Benutzbarkeit des Werkzeugs ohne umfangreiche Vorkenntnisse geachtet. Um die Nachvollziehbarkeit von Änderungen zu gewährleisten, sind alle Einträge, inklusive der Ontologie, versioniert. Die Funktionen des Repositoriums wurden anhand eines Anwendungsfalls basierend auf einem Ausschnitt der MUSE4Music Ontologie (Eusterbrock et al. 2017) evaluiert. Wei-

tere Details finden sich in (Bühler 2021: 25 ff., 41 ff.). Der Prototyp ist auf GitHub unter einer Open-Source-Lizenz frei verfügbar.³

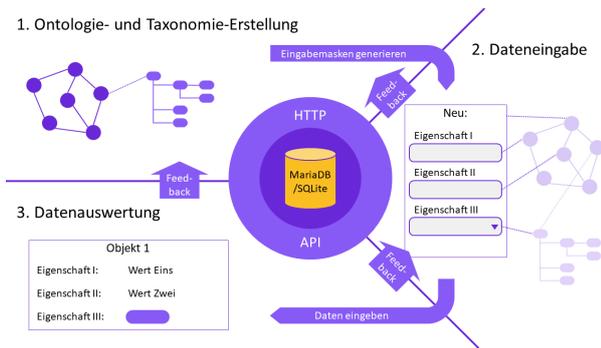


Abb. 1: Übersicht der von MUSE4Anything unterstützten Arbeitsschritte.

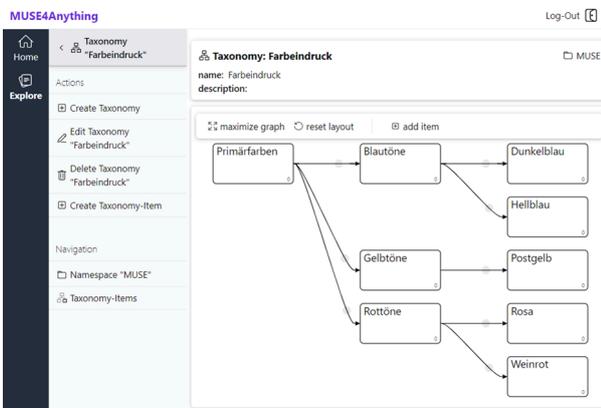


Abb. 2: Screenshot des Taxonomie-Editors mit einer stark vereinfachten Taxonomie aus der MUSE Ontologie.

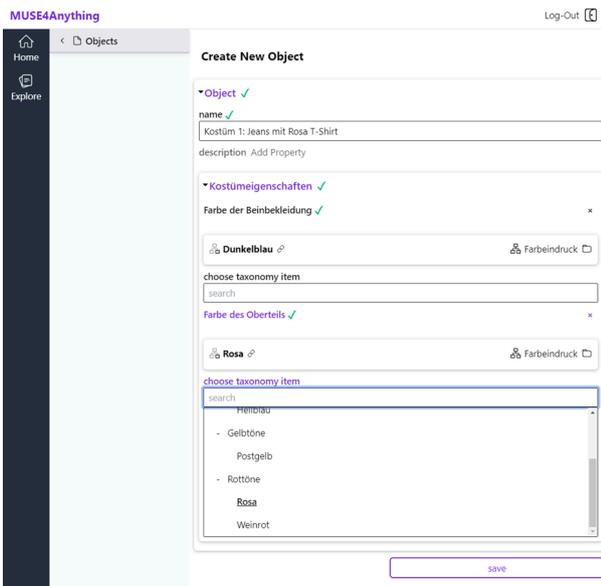


Abb. 3: Screenshot einer generierten Eingabemaske.

Ausblick

MUSE4Anything befindet sich noch teils in der Entwicklung. Weitere Funktionen, darunter eine grafische Visualisierung der Ontologie, ein einfacherer Editor für die Ontologie und eine minimale Benutzerverwaltung, sind bereits in Arbeit. In Zukunft ist auch eine Anbindung an die Auswertungswerkzeuge, die für MUSE entwickelt wurden (Barzen 2021a: 35), vorgesehen.

Zusätzlich planen wir für die Evaluierung von MUSE4Anything die Anwendung in einer völlig anderen Domäne: An der Pädagogischen Hochschule Karlsruhe werden im Rahmen des von der "Qualitätsoffensive Lehrerbildung" geförderten Hochschulentwicklungsprojektes "Nachhaltige Integration von fachdidaktischen digitalen Lehr-Lern-Konzepten" (InDiKo) aus unterschiedlichen Fächern innovative digitale Lehr-Lern-Szenarien identifiziert. Um daraus fächerübergreifende didaktische Muster zu ermitteln (Standl/Schlomske-Bodenstein 2021) und für eine Wiederverwendung in der Lehre zu beschreiben, wird als Grundlage zunächst eine interdisziplinäre Taxonomie in einem Konsensvalidierungsprozess entwickelt und in MUSE4Anything abgebildet werden.

Fußnoten

1. www.iaas.uni-stuttgart.de/forschung/projekte/muse/
2. www.iaas.uni-stuttgart.de/forschung/projekte/muse4music/
3. [www.github.com/Muster-Suchen-und-Erkennen/muse-for-anything](https://github.com/Muster-Suchen-und-Erkennen/muse-for-anything)

Bibliographie

Alexander, Christopher / Ishikawa, Sara / Silverstein, Murray / Jacobson, Max / Fiksdahl-King, Igrid / Angel, Shlomo (1977): „A Pattern Language: Towns, Buildings, Constructions.“ New York: Oxford University Press.

Barzen, Johanna (2018a): „Wenn Kostüme sprechen – Musterforschung in den Digital Humanities am Beispiel vestimentärer Kommunikation im Film.“ Dissertation, Universität zu Köln, <https://kups.ub.uni-koeln.de/9134/> [letzter Zugriff 6. Juli 2021].

Barzen, Johanna (2021a): „From Digital Humanities to Quantum Humanities: Potentials and Applications“ in: Miranda / E. R. (ed.): *An Introduction to Core Concepts, Theory and Applications*, Cham: Springer Nature. (Zur Veröffentlichung angenommen), Preprint <https://arxiv.org/abs/2103.11825> [letzter Zugriff 6. Juli 2021].

Barzen, Johanna / Breitenbücher, Uwe / Eusterbrock, Linus / Falkenthal, Michael / Hentschel, Frank / Leymann, Frank (2017): „The vision for MUSE4Music. Applying the MUSE method in musicology“ in: Bernhard Mitschang (eds.): *Computer Science - Research and Development. Advancements of Service Computing: Proceedings of SummerSoC 2016*, 32/3-4, 323-328.

Barzen, Johanna / Falkenthal, Michael / Leymann, Frank (2018b): „Wenn Kostüme sprechen könnten: MUSE - Ein musterbasierter Ansatz an die vestimentäre Kommunikation im Film“, in: Bockwinkel P et al (ed.) *Digital Humanities. Perspektiven der Praxis*, Frank & Timme, 223-241.

Barzen, Johanna / Bühler, Fabian / Leymann, Frank (2021b): „MUSE Datenset“. DaRUS, V1, 10.18419/darus-1805.

Bühler, Fabian (2021): „MUSE4Anything“ Masterarbeit, Universität Stuttgart 10.18419/opus-11410.

Eusterbrock, Linus / Barzen, Johanna / Hentschel, Frank (2017): „Eine Ontologie symphonischer Musik des 19. Jahrhunderts“. Technischer Bericht Nr. 2017/02, Universität Stuttgart.

Furrer, Frank J. (2014): „Eine kurze Geschichte der Ontologie. Von der Philosophie zur modernen Informatik.“ In: *Informatik Spektrum*. Organ der Gesellschaft für Informatik e.V. und mit ihr assoziierter Organisationen 37 (2014), H. 4, 308-317.

Hinrichs, Erhard / Krauwer, Steven (2014): „The CLARIN Research Infrastructure: Resources and Tools for E-Humanities Scholars.“ In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*, Mai 2014, 1525-31, <http://dspace.library.uu.nl/handle/1874/307981> [letzter Zugriff 6. Juli 2021].

Musen, M.A. (2015): „The Protégé project: A look back and a look forward.“ In: *AI Matters*. Association of Computing Machinery Specific Interest Group in Artificial Intelligence, 1(4), June 2015. 10.1145/2757001.2757003.

OWLGrEd (2021): <http://owlgred.lumii.lv> [letzter Zugriff 6. Juli 2021]

Reichert, Ramón (2017): „Digital Humanities als Wissenschaft“ in: Jannidis, Fotis / Kohle, Hubertus / Rehbein, Malte (eds.): *Digital Humanities: Eine Einführung*. Stuttgart: J. B. Metzler, 13-34.

Standl, Bernhard / Schlomske-Bodenstein, Nadine (2021): „A Pattern Mining Method for Teaching Practices.“ In: *Future Internet* 13(5). 10.3390/fi13050106.

Stellato, Armando / Fiorelli, Manuel / Turbati, Andrea / Lorenzetti, Tiziano / Gemert, Willem / Dechandon, Denis / Laaboudi-Spoiden, Christine / Gerencsér, Anikó / Waniart, Anne / Costetchi, Eugeniu / Keizer, Johannes (2020): „VocBench 3: A collaborative Semantic Web editor for ontologies, thesauri and lexicons“, in: *Semantic Web* 11(5): 855-881, 10.3233/SW-200370.