

- **Lehrenden**, die in thematischer Verbindung zu Lehrinhalten bereits entwickelte Lehrmaterialien finden können, um sie im Rahmen eigener Lehre einzusetzen,
- **Forschenden**, die durch die Nachnutzung der Materialien neue Methoden und Ressourcen schnell entdecken und sich aneignen können. Durch die Bereitstellung von eigenen Materialien können sie aber auch aktiv z. B. neue Methoden und Ressourcen vorstellen, ihr Wissen teilen, eigene Materialien verbreiten, diese bekannt machen und die Sichtbarkeit eigener Arbeitsergebnisse erhöhen,
- einer **interessierten Öffentlichkeit**, indem der Tutorial Finder nicht nur der wissenschaftlichen Gemeinschaft vorbehalten ist, sondern die Inhalte offen zur Verfügung stellt und einen schnellen Überblick über diese ermöglicht.

Tutorial Finder

Aufbau des Tutorial Finders

Der CLARIAH-DE Tutorial Finder baut auf einer fachlichen Adaption der in DARIAH-DE entwickelten *Generischen Suche*⁹ mit den Komponenten der *Collection Registry*¹⁰ (CR) und des *Data Modelling Environment*¹¹ (DME) auf¹². Die Generische Suche indiziert digitale Inhalte der Sammlungen gemeinsam mit vorhandenen Metainformationen und ermöglicht einen Zugriff auf diese. Dafür muss die Sammlung zunächst in der CR registriert und beschrieben werden. Die Beschreibung in der CR geschieht anhand des *DARIAH Collection Description Data Model* (DCDDM)¹³. In den vordefinierten, aber nur in wenigen Fällen obligatorischen Feldern (siehe Abbildung 2) werden die Sammlung, der Ort der Sammlung, die Lizenz sowie die technische Schnittstelle bzw. die Zugriffsmöglichkeit auf die Inhalte der Sammlung beschrieben.

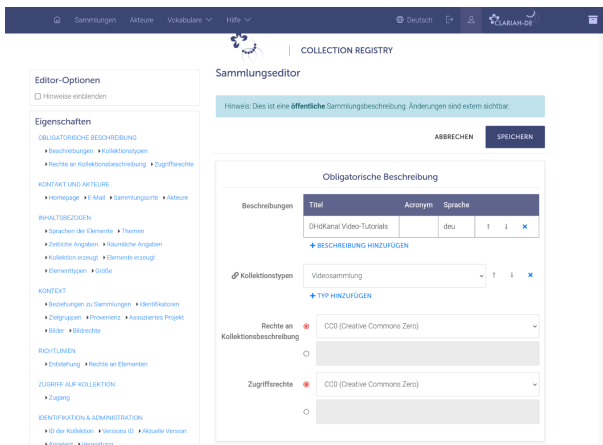


Abb. 2: Sammlungseditor der CR im Tutorial Finder

Nachdem eine Sammlung beschrieben, gespeichert und veröffentlicht wurde, erscheint sie in der CR des Tutorial Finders (siehe Abbildung 3).

Bild	Titel	Kollektionstypen	Zugang	Aktuelle Version
	Teaching and Learning Materials Collection (TeLeMaCo)	Lehrmaterialsammlung	Online file	18. November 2020, 15:54:02
	Digital Research Infrastructure for the Arts and Humanities (DARIAH-DE)	Lehrmaterialsammlung	Online file (2)	18. November 2020, 14:08:01
	DARIAH Campus	Lehrmaterialsammlung	Git Repository	18. November 2020, 14:59:09
	The Programming Historian	Lehrmaterialsammlung	Git Repository	18. November 2020, 15:01:18
	DARIAH Video-Tutorials	Videosammlung	Online file	18. Mai 2021, 12:28:27
	Parhense (Youtube Channel)	Videosammlung	Online file	18. Mai 2021, 13:38:00

Abb. 3: In die CR des Tutorial Finder eingetragene Sammlungen

Anschließend werden mithilfe der DME die Daten der Sammlung unter Beachtung ihres disziplinären Kontexts, ihrer Struktur und Komplexität modelliert, Verknüpfungen erstellt und in Form von Zugriffstellen bereitgestellt. Die Generische Suche kann danach auf Datenmodelle und Mappings der DME zugreifen, Informationen zu den Daten weiterverarbeiten, transformieren und integrieren. Hierdurch entstehen Brücken zwischen heterogenen Sammlungen – unabhängig von dort jeweils angebotenen Datenformaten. Volltexte und Metadaten aus TeLeMaCo können so beispielsweise gemeinsam mit den Transkripten aus dem Digital Medieval Webinar Repository (DMWR)¹⁴ durchsucht und ausgewertet werden.

Einmal kontextuell erschlossen und modelliert können Datenmodelle auf weitere integrative Formate abgebildet werden. Im Fall des Tutorial Finders wurde zunächst ein – um die Fähigkeit Volltext aufzunehmen – erweitertes Schema von DataCite¹⁵ gewählt. Über die grundlegende Suchfunktionalität hinausgehende Darstellungen in Form von Wissensgraphen werden derzeit evaluiert und weitergeführt.

Suchfunktionalität des Tutorial Finders

Nachdem eine Sammlung eingetragen ist und ihre Inhalte modelliert sind, kann sie über die Suchfunktionalität des Tutorial Finders durchsucht werden. Hierfür stehen eine *einfache* und eine *erweiterte Suche* zur Verfügung. Die *einfache Suche* (siehe Abbildung 4) erlaubt eine Volltextsuche sowohl auf der Ebene der Primär- als auch der Metadaten – unabhängig von der Existenz von Mappings zwischen den betrachteten Sammlungsdaten. Sie kann auf sämtlichen verfügbaren Sammlungen mit Lehr- und Schulungsmaterialien oder auf einer benutzerdefinierten Teilmenge hieraus ausgeführt werden.

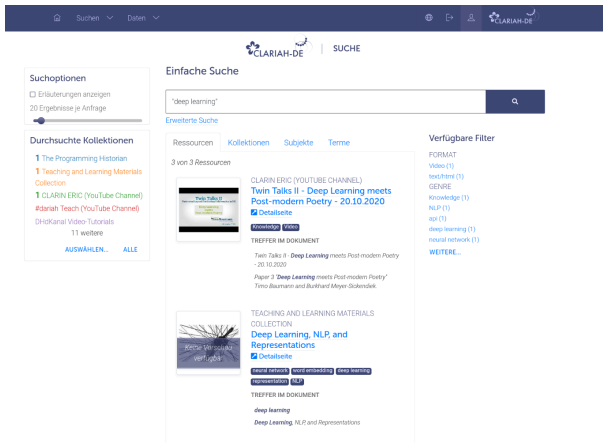


Abb. 4: Suchergebnisse des Tutorial Finders zu „deep learning“ mittels einfacher Suche

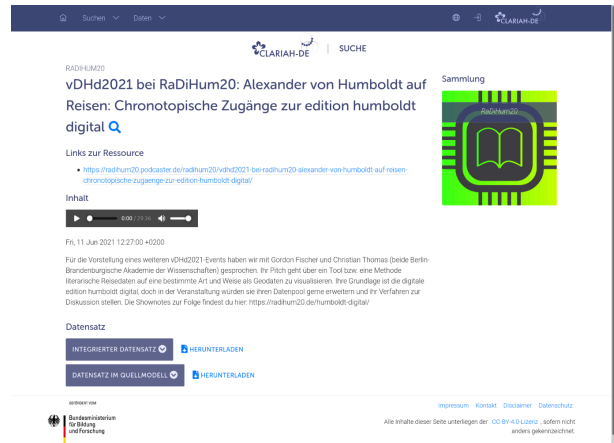


Abb. 6: Detailsicht zu einem Suchergebnis im Tutorial Finder

Die *erweiterte Suche* (siehe Abbildung 5) bietet neben der Volltextsuche eine gezielte Suche, bei der spezifische Kriterien zur Einschränkung der Suchergebnisse angewendet werden können, die auf dem Schema von DataCite basieren. So können Metadatenfelder des Datenmodells beispielsweise herangezogen werden, um Suchergebnisse nach Orten oder Genres zu filtern.

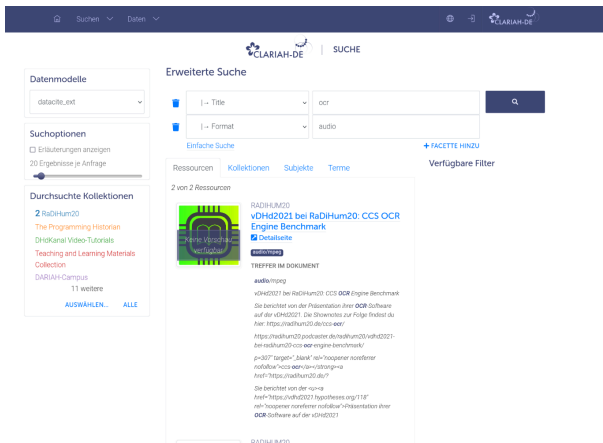


Abb. 5: Gezielte Suche nach Audiodateien zu „ocr“ mittels erweiterter Suche

Wählt man ein Objekt der Ergebnisliste aus, gelangt man zu Informationen über die Zugehörigkeit zu einer Sammlung (siehe Abbildung 6). Insbesondere wird auch die URL präsentiert, unter der das Objekt in seinem ursprünglichen Sammlungskontext zu finden ist. Verfügbare Metainformationen zum Objekt können angesehen und heruntergeladen werden.

Dabei hostet der Tutorial Finder die Lehr- und Schulungsmaterialien nicht selbst, sondern ermöglicht bzw. erleichtert lediglich den Zugang auf diese über Metadaten und Volltexte. Es sind daher die zugrunde liegenden Lizenzen für die Nutzung in den unterschiedlichen Lehr- und Lernkontexten zu beachten.

Angebot und Erweiterung des Tutorial Finders

Aktueller Stand

Der Tutorial Finder ist ein offenes Angebot und befindet sich im stetigen Ausbau. Er wurde zunächst als Dienst für CLARIAH-DE konzipiert, der eine übergreifende Suche nach den Materialien aus den zwei vorangegangenen Projekten – CLARIN-D¹⁶ und DARIAH-DE¹⁷ – wie TeLeMaCo von CLARIN oder DARIAH-Campus ermöglichen sollte. Im Laufe der Entwicklung und Erweiterung wurden in den Tutorial Finder Sammlungen und Materialien aufgenommen, die aus anderen für die Community wichtigen nationalen und internationalen Projekten, Vorhaben und Initiativen, z. B. The Programming Historian¹⁸ oder linguisticsweb.org, stammen. Weiterhin finden sich darin einzeln außerhalb jeglicher Sammlungen entwickelte Tutorials und Anleitungen, die jedoch methodisch aufbereitet sind und eine wertvolle Lehr- und Lernquelle zu aktuellen Themen der (Digitalen) Geisteswissenschaften und benachbarter Disziplinen bieten.

Erweiterung des Angebots

Die Sammlung des Tutorial Finders wird nicht nur im Rahmen von Text+ erweitert. Auch die Community der Digital Humanities kann hierzu einen Beitrag leisten und Vorschläge zur Angebots-erweiterung des Tutorial Finders machen: Interessierte, die eine Sammlung von Lehr- und Schulungsmaterialien oder einzelne Tutorials einem möglichst breiten Nutzerkreis zur Verfügung stellen wollen, können diese über den CLARIAH-DE Helpdesk¹⁹ vorschlagen oder sie als authentifizierte Nutzer*innen über den Sammlungseditor der CR (siehe Abbildung 2) im Tutorial Finder selbst anlegen. Sofern erforderlich, kann über den Helpdesk jederzeit auch Unterstützung angefordert werden. So können die

Nutzer*innen das Spektrum der angebotenen Materialien unkompliziert erweitern und über den Tutorial Finder auffindbar, durchsuchbar und zugänglich machen.

Neben der thematischen Relevanz für den Bereich der Digital Humanities bestehen einige grundlegende Voraussetzungen für die Aufnahme einer Sammlung von Lehr- und Schulungsmaterialien in den Tutorial Finder:

- Die Materialien einer Sammlung müssen unter einem **freien Lizenzierungsmodell** zur Verfügung stehen.
- Die Sammlung muss **öffentlich zugänglich** sein.
- Die Sammlung muss **aktuell gehalten** werden.
- Es muss eine **Ansprechperson benannt** werden.

Lizenzauswahl für Lehr- und Schulungsmaterialien

Die für die Lehr- und Schulungsmaterialien am häufigsten genutzten Standardlizenzen sind die Creative-Commons-(CC)-Lizenzen²⁰. Sie werden in den Geistes-, Kultur- und Sozialwissenschaften und in Bildungseinrichtungen allgemein für die Lizenzierung von Texten, Bildern und Videos oder Multimediawerken genutzt. Zu den Vorteilen der CC-Lizenzen gehört, dass sie im Vergleich zu anderen Open-Content-Lizenzen verhältnismäßig einfach zu verstehen sowie international anerkannt und anwendbar sind. Den Nutzenden ermöglichen sie, schnell zu erkennen, ob und unter welchen Bedingungen sie das Material verwenden können. Den Autor*innen bieten sie eine Auswahl von modular aufgebauten Standard-Lizenzen, mit denen sie ihre Materialien nach eigenen Bedürfnissen lizenzieren können (vgl. Kamocki & Ketzan, 2014 für mehr Details zu CC-Lizenzen). Für umfassende Informationen zu verschiedenen Lizenzthemen und zur Auswahl einer passenden Lizenz steht die *CLARIN ERIC Legal Information Platform*²¹ zur Verfügung.

Zugänglichkeit der Sammlung mit Lehr- und Schulungsmaterialien

Die Infrastruktur des Tutorial Finders ermöglicht eine Abfrage von Angeboten einer Sammlung über eine standardisierte Schnittstelle, z.B. OAI-PMH. Die Bereitstellung einer solchen standardisierten Schnittstelle ist jedoch keine Voraussetzung. Wichtig ist nur, dass die Inhalte der Sammlung online zugreifbar sind und so in den Tutorial Finder eingebunden werden können. So verfügt beispielsweise die TeLeMaCo-Sammlung über keine dedizierte Zugriffsschnittstelle. Die Daten wurden in diesem Fall direkt von der Webseite in der DME mithilfe eines iterativen Web-Crawling erschlossen, modelliert und extrahiert und so in den Tutorial Finder übernommen. Ein anderes Beispiel bildet die Anbindung von Sammlungen, die ihre Materialien in einem öffentlich zugänglichen Git-basierten Repositorium²² zusammenstellen und verwalten, wie *The Programming Historian* oder *DARIAH-Campus*. In diesem Fall kann das Git-Protokoll als Schnittstelle angebunden werden.

Aktualität der Sammlung mit Lehr- und Schulungsmaterialien

Da der Tutorial Finder lediglich den Zugriff auf Sammlungen mit Schulungs- und Lehrmaterialien ermöglicht, die bereits öffentlich zugänglich sind, und sie selbst weder hostet noch pflegt, wird die regelmäßige Überprüfung und Aktualisierung der Samm-

lungsinhalte von den jeweiligen Autor*innen bzw. den dafür zuständigen Personen vollzogen.

Klärung der zuständigen Ansprechpersonen

Bei der Beschreibung einer Sammlung im Sammlungseditor der CR (siehe Abbildung 2) werden die Autor*innen bzw. die dafür zuständigen Ansprechpersonen ggf. auch mit der Angabe der Affiliation genannt. Diese Angaben sind notwendig, um z. B. Fragen zu Hintergründen der Sammlung, Zugriffsmöglichkeiten oder allgemeinen und technischen Voraussetzungen für die Aufnahmen der Sammlung in den Tutorial Finder zu klären.

Zusammenfassung und Ausblick

In diesem Beitrag wurde der Tutorial Finder vorgestellt, der als Angebot des CLARIAH-DE Projekts zum zentralen Durchsuchen von vielfältigen Lehr- und Schulungsmaterialien zu digitalen Methoden, Forschungsdaten, Ressourcen und Diensten im Bereich der Digital Humanities entwickelt wurde. Er ersetzt nicht die bereits vorhandenen Sammlungen von Lehr- und Schulungsmaterialien, da der Tutorial Finder die Materialien nicht selbst hostet, sondern die Suche in verteilten Inhalten ermöglicht, die Sichtbarkeit von Materialien erhöht und ihre Nachnutzbarkeit steigert. Der Tutorial Finder ist für verschiedene Nutzergruppen – Forschende, Lehrende und Studierende aus dem Bereich der Digital Humanities – relevant und kann unter ihrer aktiven Teilnahme erweitert werden. Der Dienst wird derzeit für die weitere inhaltliche und technische Entwicklung in die neu geschaffenen Strukturen der Text+²³ Forschungsdateninfrastruktur überführt und weiterentwickelt – insbesondere mit Blick auf die organisatorische und fachliche Einbindung in die NFDI Konsortien von Text+ und NFDI4Culture²⁴

Fußnoten

1. <https://telemaco.clarin-d.uni-saarland.de/hub/> [letzter Zugriff auf alle Verweise in diesem Dokument am 15. Juli 2021]
2. <https://campus.dariah.eu/>
3. <http://linguisticsweb.org/doku.php?id=start>
4. <https://de.dariah.eu/tei-tutorial>
5. <https://dariah-de.github.io/DARIAH-DKPro-Wrapper/tutorial.html>
6. <https://github.com/DARIAH-DE/TopicsExplorer>
7. CLARIAH-DE war ein vom Bundesministerium für Bildung und Forschung (BMBF) gefördertes, zweijähriges (2019–2021) Verbundprojekt, in dessen Rahmen die Ressourcen der Forschungsinfrastrukturen CLARIN-D und DARIAH-DE zusammengeführt wurden.
8. <https://teaching.clariah.de/search/>
9. <https://dfa.de.dariah.eu/doc/search/>
10. <https://dfa.de.dariah.eu/doc/colreg/>
11. <https://dfa.de.dariah.eu/doc/dme/>
12. Für mehr Informationen zu den Komponenten der GS siehe auch Gradl (2019) sowie Henrich und Gradl (2021).
13. <https://github.com/DARIAH-DE/DCDDM>
14. <https://zenodo.org/communities/dmwr/>
15. <https://schema.datacite.org>
16. CLARIN-D ist der deutsche Beitrag zur *Common Language Resources and Technology Infrastructure* (CLARIN), die als Eu-

European Research Infrastructure Consortium (ERIC) ein Netzwerk aus 20 Ländern neben weiteren Partnern bildet, URL: <https://www.clarin-d.net/de>

17. DARIAH-DE ist der deutsche Partner der *Digital Research Infrastructure for the Arts and Humanities* (DARIAH-EU), die mit 19 Ländern und weiteren Partnern ein ERIC auf europäischer Ebene ist, URL: <https://de.dariah.eu>

18. <https://programminghistorian.org/>

19. <https://www.clariah.de/support/>; die Überführung in Text+Strukturen ist in Arbeit

20. <https://creativecommons.org>

21. <https://www.clarin.eu/content/legal-information-platform>

22. Für die ausführliche Anleitung zur Nachnutzung Git-basierter Sammlungen im Rahmen der Infrastrukturdienste von CLARIAH-DE siehe Gradl & Jegan (2021).

23. <https://www.text-plus.org/>

24. <https://nfdi4culture.de/>

Bibliographie

Annisius, Marie / Bock, Sina / Gradl, Tobias / Schopf, Juliane / Stegmeier, Jörn / Werthmann, Antonina (2021): „CLARIAH-DE Work Package 3: Skills Training and Promotion of Junior Researchers“ in: *CLARIAH-DE 2. Vollversammlung / 2nd General Assembly*. Göttingen: Zenodo. DOI: 10.5281/zenodo.4578414

Eckart, Thomas / Werthmann, Antonina / Buddenbohm, Stefan / Sambale, Heidemarie / Helfer, Felix / Jegan, Robin (2021): „CLARIAH-DE Work Package 4 Technical Integration and Coordination of Technical Developments“ in: *CLARIAH-DE 2. Vollversammlung / 2nd General Assembly*. Göttingen: Zenodo. DOI: 10.5281/zenodo.4572610

Gradl, Tobias (2019): „Dokumentation der Datenföderationsarchitektur“ in: *DARIAH-DE Working Papers Nr. 39*. Göttingen: DARIAH-DE, URN: urn:nbn:de:gbv:7-dariah-2019-11-7

Gradl, Tobias / Robin Jegan (2021): „Nachnutzung Git-basierter Sammlungen im Rahmen der Infrastrukturdienste von CLARIAH-DE“ in: *DARIAH-DE Working Papers Nr. 42*. Göttingen: DARIAH-DE. URN: urn:nbn:de:gbv:7-dariah-2021-2-5

Henrich, Andreas / Tobias Gradl (2021): „Integration von Forschungsdaten. Wie können Forschungsinfrastrukturen helfen?“ in: *Innovation in der Bauwirtschaft Innovation in the Building Industry*. Berlin, Boston: De Gruyter, 749-762. DOI: 10.1515/9783110538915-039

Kamocki, Paweł / Erik Ketzan (2014): „Creative commons and language resources: general issues and what's new in CC 4.0“. https://ids-pub.bsz-bw.de/frontdoor/deliver/index/docId/3224/file/Kamocki_Ketzan_Creative_Commons_and_Language_Resources_2014.pdf