

Texte für CLARIN-D zur Verfügung stellen

10/13/2024 08:26:17

FAQ Article Print

Category:	FAQ Clarin::Dienste	Votes:	1
State:	public (all)	Result:	75.00 %
Language:	de	Last update:	11:27:27 - 07/25/2016 (Europe/Berlin)

Keywords

Service

Symptom (public)

Ich verfüge über digitalisierte Texte, die ich CLARIN-D gerne zur Verfügung stellen würde. Wie muss ich vorgehen?

Problem (public)

Solution (public)

Hier gibt es verschiedene Möglichkeiten, von denen im Folgenden zwei genannt werden können.

Die Universität Leipzig entwickelt unter dem Namen „TEI-Integrator“ ein Programm, welches aus einem xml-Ausgangstext (halb)automatisch Metadaten und relevante Textdaten extrahiert und diese sowohl in das Virtual Language Observatory (VLO) von CLARIN-D überführt als auch als XML/TEI abspeichert.

Das CLARIN-D Zentrum der BBAW bietet über das dort angesiedelte Deutsche Textarchiv ([1]DTA) das Modul DTA-Erweiterungen ([2]DTAE) an. Mit dem DTAE-Modul können historische Texte, die in anderen Projektkontexten digitalisiert wurden, dem DTA-Kernkorpus angelagert werden, um so die Textbasis für ein Referenzkorpus des historischen Neuhochdeutschen (Zeitraum ~1650–1900) zu schaffen. Mit DTAE verbunden ist ein Workflow, um die zur Verfügung gestellten Texte so aufzubereiten, dass sie DTA- und somit CLARIN-kompatibel sind. Die Texte werden dann zusammen mit allen anderen DTA-Korpusdaten in CLARIN verfügbar gemacht. Falls Sie in diesem Rahmen Ihre Texte verfügbar machen wollen, wenden Sie sich bitte an das DTA-Team (BBAW, Berlin, [3]dta@bbaw.de).

[1] <http://www.deutschestextarchiv.de/>

[2] <http://www.deutschestextarchiv.de/dtae>

[3] <mailto:dta@bbaw.de>