

European Summer School for Digital Humanities

Text Mining with Canonical Text Services

Practical Tasks 1

Managing CTS Instances

First Steps

The CTS Admin Tool is available as a deployable .warfile. It can be installed and deinstalled without any impact on the published CTS instances. If it breaks for any reason, you can simply delete and redeploy the .war file. After deploying the .war file, the password and user name is both “cts”. The CTS template may have to be updated after deployment.

Access the pre installed CTS Admin Tool using the user name “cts” and the password “cts”.

Change the password to something else. Since it is a training workshop, it is advised to choose a simple password. There is no password recovery process. If the password is forgotten, the tool has to be re installed.

Deploy a fresh CTS instance with the name “myfirstCTS”.

Rename it to a name of your choosing.

Deploy a new Canonical Text Service

Deploy a fresh CTS instance with a name of your choosing

Familiarize yourself with the parameters in the tab *Data Import*.

Configure the import process according to given information and publish the pre configured data set.

Create a new CTS instance and load the data from the TED Talk Transcripts (CTS Namespace urn:cts:ted) in a language of your choosing. Limit the document count to 10 documents.

Change the configuration in such a way that one of the other data sets is added to the TED CTS instance. Make sure that the previously imported data is not lost.

Ctrace

Find the Ctracer in the Browser Tab and open it in Fullscreen mode.

Use the top hierarchy to select a document of your previously installed CTS instance.

Use the structure view on the left to navigate through a document.

Create a valid text passage request to the CTS server.

Namespace Resolver

Find out, what the CTS Namespace Resolver is and why it is necessary. You can find information at <http://cts.informatik.uni-leipzig.de/>.

Request a list of supported namespaces from the Namespace Resolver.

Find information about the CTS instance *berlinletters*.

Final Task

Open our local office document or create a new one and add a valid CTS GetPassage request to your previously published data set as a hyperlink reference. If you want you can exchange the documents with other participants to let them request data from your CTS instance.