

SHORT TERM SCIENTIFIC MISSION (STSM) – SCIENTIFIC REPORT

Action number: IS1312

STSM title: Web Portal Development

STSM start and end date: 15/01/2018 to 25/02/2018

COST Applicant: Ahmet Üstün, Middle East Technical University

Grantee name: Prof. Dr. Bonnie Webber, University of Edinburgh

PURPOSE OF THE STSM

The main objective behind the “Web Portal Development” STSMs is to prepare the Textlink Web Portal which will ultimately provide a various search, visualization and dissemination capabilities for researchers. The current STSM is the last phase and specifically aimed to finalize the existing portal development. Therefore, improvement on user interface, development regarding to statistics of annotated corpora which are uploaded to the portal, extending search options according to specific information such as keyword and enable to download and save the search results are main topic for this STSM.

In addition to development regarding to search utilities and user experience, another aim of this STMS is to make Textlink Web Portal publicly available so that it can be used by researchers.

DESCRIPTION OF WORK CARRIED OUT DURING THE STSMS

First improvement in the search menu is to enable search on multiple file at the same time. In order to use this feature, user should upload annotation/raw file pairs to the portal. A tabbed menu is designed to display each file on another tab so that user can analyse each file separately at the same time.

Second improvement in the search menu is adding the keyword search utility. Two text area are put on the search menu so that user can search specific words in argument 1 and/or argument 2 accordingly. In order to search words in arguments, each word is indexed with respect to argument.

Thirdly, in order to download and save search results on users own PC, a download option is added to search menu. Users can choose a filename and download its result from portal to save with this feature. Downloaded results is in pipe-delimited structure so that users can use this results in PDTB annotator for the further annotations.

As for the statistics, two statistics menu have been developed --- one for individual files and other for the total statistics. Google JavaScript graph library is used to display statistics in pie chart format. By using first menu (File Statistics), users can compute discourse type distribution over senses or sense of discourse connective over discourse type for each file which is displayed separately on tabbed menu. In second menu (Total Statistics), users can compute same statistics regarding to all file information.

For the better user experience, a home page and CSS template is designed. All menus and buttons shares this CSS template.

Lastly, in order to access Textlink Web Portal publicly, a domain name is used which is automatically generated by AWS servers. Configurations required for JQuery calls and Django backend have been prepared and deployed on AWS server that contains portal source code and executables. Therefore, researchers can access Textlink Web Portal on **http://ec2-18-219-79-53.us-east-2.compute.amazonaws.com:8000/home**

This automatically generated domain name will be mapped a generic web site address which is arranged by TextLink management committee.

CONCLUSION

In summary, during the STSM, the existing web-portal is extending regarding to more search facilities such as multiple file option and keyword search. Then download feature is implemented to download and save search results in pipe-delimited structure. Two statistics menu is designed and implemented for separate file statistics and total statistics accordingly. Lastly, Textlink Web Portal made open to public access and usage.