

## **Deliverable T.4: Final report**

### **As software provider you shared your resource(s) as (choose):**

Docker image: <https://hub.docker.com/r/andrelamurias/bolstm/>

### **Briefly, explain/justify your choice:**

We shared our component as a docker image because it makes it easier to pack and distribute software written in any language, without having to be restricted to any framework. It can be executed both on a personal computer or on a cloud server. Furthermore, the code is openly available too, and it can be integrated with other tools using the command line, without being restricted to any framework.

### **List the highlights and key success factors of the project:**

The project adapted LSTMs, a popular deep learning technique, to biomedical ontologies to extract relations from biomedical text. The incorporation of semantic information based on biomedical ontologies improves the quality of the relations extracted when compared to using only lexical features.

### **List and identify anything you found useful:**

We found that the infrastructure provided by the OMTD platform was useful in providing a repository of text mining tools, and the Google groups and GitHub pages were useful to solve technical issues. Furthermore, the fact that we could share our component as a docker container was also important, so that we did not have to restrict our component to a specific framework.

### **List and describe any unexpected events**

We encountered some problems with the time schedule since we originally developed the project for a duration of 6 months, to test more parameters and case studies. However, we were able to complete the project for one case study. The guidelines to produce each deliverable were not fully detailed

### **List any issue you had regarding (interoperability) requirements related to**

A web interface was provided to generate the required metadata for our component, however, some of the fields were ambiguous and no examples were given. It was not clear how our component would be used in the context of an application, i.e. how it should process the input and output documents. It was also difficult to know when to make the component public since it was no longer possible to edit the metadata after that.

### **The amount of work to adapt and integrate your tools into OMTD was**

The amount of work necessary to integrate our platform was as expected for a project of this dimension and with the objective of integrating various text mining components.

### **Do you think the effort done is justified? (yes, no, to what extend)**

In text mining there are often issues with dependencies and compatibility between components, so the effort required to make a tool more compatible with other tools is perfectly justifiable.

**List and summarize any lessons learned from this project.**

The main lesson from this project was the importance of domain-specific ontologies and deep learning models for classification of biomedical relations. Considering that biomedical ontologies are openly available and regularly updated as the knowledge on the domain progresses, they should be considered important information sources for relation extraction. Furthermore, it was simple to adapt the code and release it as a container, while this effort had the benefit of making our project more useful for other researchers.

**Regarding the use case you presented, list the aspects that OMTD can (or could) benefit the use case.**

The use case we presented consisted of using the ChEBI ontology to extract drug-drug interactions from biomedical text. Since our component requires the input text to be pre-annotated with chemical entities mapped to ChEBI, OMTD will benefit by connecting our component to a component that pre-annotates with that information. Furthermore, the results of the tool could be used by other components for other tasks.

**List what you would have done differently and/or your recommendations for improvement:**

With more planning and time, we would have implemented more file formats and use cases with our component, for example, protein-protein and phenotype-gene interactions.