**Descriptions of the files in this public repository**

**Journal Name**: Journal of Computer Languages

**Paper Title**: Ask or Tell: An Empirical Study on Modeling Challenges From LabVIEW Community

**Authors**: Xin Zhao (Corresponding Author), Gurshan Rai, Saheed Popoola

**Contact Information**: xzhao1@seattleu.edu

There are two folders in this repository.

Folder One: Analysis Results.

This folder stores all the results we analyzed in our paper. To be more specific, there are three Excel files inside the folder.

1. DataCollection.xlsx. This is the file that contains the raw data we collected from the LabVIEW official discussion forum. Reference: Section 4.2.2.
2. 500PostsManualAnalysis.xlsx. This is the file we used to manually define the concept/topic for LabVIEW posts. Reference: Section 4.2.4.
3. FinalRefinedTopicModelingResults.xlsx. This is the file that contains all the posts that belong to one of the eight categories we discussed in the paper. Reference: Section 4.2.5.

Folder Two: Script.

This folder contains the LDA modeling code. To be more specific, there are two Python notebook files in this folder.

1. Language.ipynb. This is the code we used to remove the posts that are not written in English during the process of post-cleaning. Reference: Section 4.2.3.
2. TopicModeling.ipynb. This is the code we used to conduct topic modeling by using the LDA approach. Reference: Section 4.2.4.

Usage: Both these scripts are executed by Python notebook – Jupyter Notebook. The running requires the installation of a Jupyter Notebook. After installation, users can import these files to replicate the modeling process by using the LDA approach.