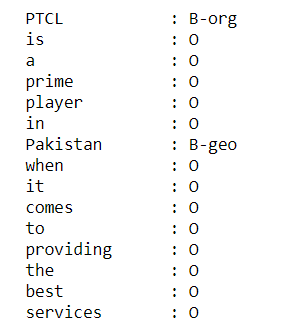
**Named Entity Recognition (NER)**

Named Entity recognition is useful in solving a lot of research problems in NLP such as QA and information extraction. The task consists of identifying and extracting named entities in the text to some defined categories such as PERSON, ORGANIZATION, LOCATION, etc. For this purpose, various methods can be adopted and one of such is a deep learning approach based on a Bi-directional LSTM-CRF which we have used in this notebook.

**Dataset**

The dataset used is the GMB(Groningen Meaning Bank) corpus mixed with some of our own custom annotated data of PTCL. In this dataset, a BIO annotation scheme is used with main entities tagged with a "B" (Beginning) prefix, multi-words tagged with "B" and "I" (Inside) prefixes while all other words are tagged with an "O" (Other). e.g the sentence "PTCL is a prime player in Pakistan when it comes to providing the best services" is tagged as :



**Predicting Entities**

For predicting entities, the data first gets preprocessed into a suitable format. After that, a Bi\_LSTM\_CRF model is built and trained on specific parameters. Finally, the model is evaluated for accuracy and then tested on some sample sentences

**Resources and Acknowledgments**

[**https://github.com/floydhub/named-entity-recognition-template**](https://github.com/floydhub/named-entity-recognition-template)

[**https://www.kaggle.com/abhinavwalia95/entity-annotated-corpus?select=ner\_dataset.csv**](https://www.kaggle.com/abhinavwalia95/entity-annotated-corpus?select=ner_dataset.csv)