In [1]:
```
#@title Licensed under the Apache License, Version 2.0 (the "Licens
# you may not use this file except in compliance with the License.
# You may obtain a copy of the License at
#
# https://www.apache.org/licenses/LICENSE-2.0
#
# Unless required by applicable law or agreed to in writing, softwa
# distributed under the License is distributed on an "AS IS" BASIS,
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or i
# See the License for the specific language governing permissions a
# limitations under the License.
```

CO Open in Colab

(https://colab.research.google.com/github/lmoroney/dlaicourse/blob/master/TensorF
%20NLP/Course%203%20-%20Week%201%20-%20Lesson%201.ipynb)

Copyright 2019 The TensorFlow Authors.

In [ ]:
```
#@title Licensed under the Apache License, Version 2.0 (the "Licens
# you may not use this file except in compliance with the License.
# You may obtain a copy of the License at
#
# https://www.apache.org/licenses/LICENSE-2.0
#
# Unless required by applicable law or agreed to in writing, softwa
# distributed under the License is distributed on an "AS IS" BASIS,
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or i
# See the License for the specific language governing permissions a
# limitations under the License.
```

In [1]:
```
from tensorflow.keras.preprocessing.text import Tokenizer

sentences = [
    'i love my dog',
    'I, love my cat',
    'You love my dog!'
]

#number of words to tokenize, since sentence's are short doesn't ma
tokenizer = Tokenizer(num_words = 100)
#this tokenizes the words into encoding values
tokenizer.fit_on_texts(sentences)
#creates the key value dictionary of the word and it's token
word_index = tokenizer.word_index
print(word_index)
```

```
{'love': 1, 'my': 2, 'i': 3, 'dog': 4, 'cat': 5, 'you': 6}
```