

---

# DKCHATBOT

---

Prepared by LERO

Ofentswe Lebogo

DUE DATE: 20 FEBRUARY 2018

“To strengthen the connection between your conscious and subconscious, is to gain access to a map and compass, as you travel through parallel worlds.”

– **Kevin Michel**

Project Name: DKChatBot

Dataset Name: **Comple\_dataset\_2017\_05\_no\_duplicates.csv**

Tools to use

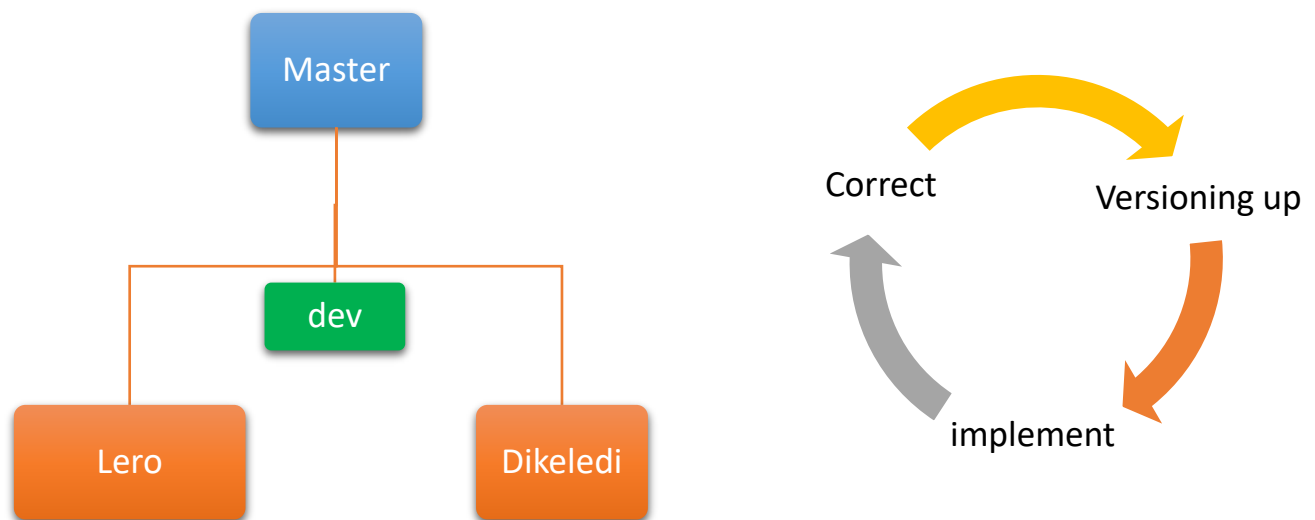
Names	Links
Anaconda 5.0.x (Python 3.6.x)	<a href="https://www.anaconda.com/download/">https://www.anaconda.com/download/</a> (Anaconda includes more than 400 packages that are parallelized so you don't need to download more packages Except the below packages mentioned, you can install them by using pip, to install the packages you can do <b>pip install &lt;package-name below&gt;</b> )
Django 2.0.x	<a href="https://www.djangoproject.com/">https://www.djangoproject.com/</a> (Python framework that allows us to communicate with the bot via the web)
Django Rest Framework 3.7.x	<a href="http://www.django-rest-framework.org/">http://www.django-rest-framework.org/</a> (Python restful framework for developing restful API's)
Keras 2.1.x	<a href="https://keras.io/">https://keras.io/</a> (Deep Learning Package that allows us to build robust models to communicate with the bot)
Tensor flow 1.4.x	<a href="https://www.tensorflow.org/">https://www.tensorflow.org/</a> (Deep Learning Package that allows us to build robust models to communicate with the bot)

## Description of data

The data was mined from this site using BeautifulSoup in python <http://www.perverted-justice.com/?con=full> to develop the Chabot to talk to, **NB: The Data contains sensitive information as it was about the predators grooming young girls, with that in mind the information contains** two fields namely Name and Message/Text, For the purpose of the individual Chabot you will have to organise the text the way you like, **SUGGESTION:** you can organise the text with the keywords me and name or you can request a script that I mined original Dataset with in order to see the flow of the conversation,

## Project Scope

The Goal of DKChatBot is to use the current data we have to have a model that answers questions from the user, with the like of machine learning we are coming far from this kind of this models and today we have capabilities to use technology to implement and generalize question in terms of Natural Language Processing, This System will be developed from scratch using perverted justice data, You are allowed to add any data to the current data we have at hand, This System as part of Learning is requested to be built on Python Language and It will be converted to Android app much simpler with the effort Django rest framework have, The project is intended to take approximately 4-8 Weeks as such you are allowed to chat with me on daily basis when you are Stuck, I would like us to use version control to avoid any technical failures, DKChatBot will have Four Branches, Namely (Lero, Dikeledi, dev and Master) refer to **figure 1**



**Figure 1**

By that we will be correcting each other and pushing to dev and once is permitted the we push it to master, the stages are called implement, correct and versioning up, it will make us have a better and efficient way of working together virtually, with that at the end of the day the models will have to reply to our questions, below is the description on how DKChatBot System will work, refer **Figure 2**, So the Project can be found on GitHub <https://github.com/Chatbort/DKChatBot> the project is empty I will initialize the projects and Let you go through the code while trying to adapt to new languages, I will raise issues on GitHub expecting you to correct the code or improving it, as such I will monitor the Repo on daily basis to keep track of our progress I will suggest you to start exploring Jupiter notebooks in python it should take you two or three days and You will add everything to the repo that is regarding the project. Keep coding

Here are useful links to get started, I recommend reading documentation and use pep8 style thank you.

[http://scikit-learn.org/stable/tutorial/text\\_analytics/working\\_with\\_text\\_data.html](http://scikit-learn.org/stable/tutorial/text_analytics/working_with_text_data.html)

<https://jupyter-notebook-beginner-guide.readthedocs.io/en/latest/>

