**WPBD LAB05 ZADANIA – snakebite oraz map reduce w pythonie**

Przed realizacją zadań uruchom Hadoop w trybie pseudo rozproszonym.

1. Ściągnij bibliotekę snakebite poleceniem:

pip3 install snakebite-py3

1. Korzystając z biblioteki stwórz następujące skrypty w Pythonie:
   1. mkdir.py – który korzystając z biblitoeki snakebite utworzy folder detective\_stories oraz parent directory o nazwie books

Obraz zawierający tekst

Opis wygenerowany automatycznie

* 1. mkdir2.py – który stworzy katalog w HDFS books\_read



* 1. bezpośrednio z hadoopa wgraj plik opowieść.txt do /books/detective\_stories

Obraz zawierający tekst

Opis wygenerowany automatycznie

Obraz zawierający tekst

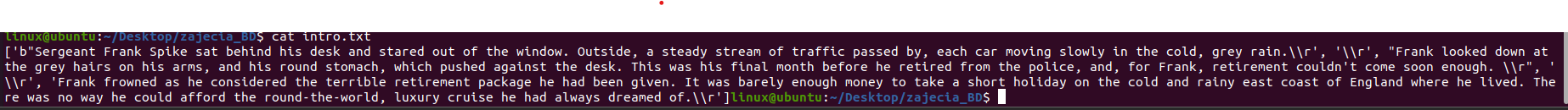
Opis wygenerowany automatycznie

* 1. ls.py – który zwróci listę katalogów i plików HDFS

Obraz zawierający tekst, zewnętrzne

Opis wygenerowany automatycznie

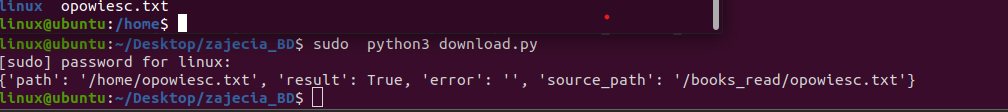
* 1. read.py – który zapisze do pliku ‘intro.txt’ pierwsze 5 linijek z pliku opowieść.txt na HDFS (wskazówka, metody Client zwracają generatory)



* 1. mv.py – który przeniesie plik opowiesc.txt z foldery detective\_stories, do books\_read n HDFS



* 1. download.py – który sciagnie plik opowiesc.txt do twojego home katalogu



1. Wykorzystując przykład z zajęć napisz program, który policzy występowanie słów w tekście opowiesc.txt. Można to zrobić w Jupyternotebook. Należy zastosować technikę map reduce i multiprocessing.