



Universitat Autònoma de Barcelona


TextCounter Manual d'Usuari

Enrique Gómez Becerra

Borja Arcos Comas

Versió: 1.0

Data: 07/05/2022

 Universitat Autònoma de Barcelona	TextCounter
	Manual d'Usuari

Full de control de modificacions

Títol	TextCounter, Manual d'usuari
Versió	1.0
Autor/a	Enrique Gómez Becerra Borja Arcos Comas
Data	07/05/2022

Control de versions		
Versió	Descripció / Motiu	Data
1.0	Document inicial	07/05/2022

 Universitat Autònoma de Barcelona	TextCounter
	Manual d'Usuari

Índex

1.	Descripció del sistema	Pàg 3
2.	Objectius	Pàg 3
3.	Manual d'usuari	Pàg 3
3.1.	Preparació de l'entorn	Pàg 3
3.2.	Execució de l'aplicació	Pàg 3
3.3.	Obtenció dels resultats	Pàg 4
3.4.	Exemple	Pàg 4

 Universitat Autònoma de Barcelona	TextCounter
	Manual d'Usuari

1. Descripció del sistema

El sistema TextCounter és un aplicatiu destinat al càlcul de percentatges d'aparició de paraules per lletra en un fitxer de text. Aquest software fa ús del paradigma Map-Reduce i, a partir d'un o múltiples fitxers de text, calcula el percentatge d'aparició de paraules per cada lletra dins de cadascun dels fitxers introduïts. Com a resultat s'obté imprès per pantalla una llista de lletres acompanyades del percentatge de paraules del fitxer que les contenen.

2. Objectius

L'objectiu d'aquest document és explicar amb claredat el funcionament del software TextCounter per facilitar-ne el seu ús a partir d'una explicació de la preparació de l'entorn i tot el necessari per fer funcionar l'aplicatiu correctament.

3. Manual d'usuari

3.1. Preparació de l'entorn

Per fer funcionar el software TextCounter inicialment caldrà tenir instal·lat Docker, ja que és l'eina a la qual es basa el funcionament de l'aplicació. Una vegada instal·lat, és necessari crear un directori amb aquells fitxers de text sobre els quals es volen realitzar els càlculs.

La creació d'aquest directori és imprescindible, ja que en el moment de fer córrer el contenidor amb l'aplicació es crearà un volum en aquest directori per fer accessibles els arxius de dades.


També és necessari tenir descarregats els fitxers amb el codi font (MapReduce.py, Main.py i TextCounter) i el fitxer Dockerfile per la creació de la imatge Docker.

Una vegada s'han realitzat les accions esmentades, l'entorn està preparat per fer funcionar el software TextCounter amb èxit.

3.2. Execució de l'aplicació

Per fer funcionar correctament el software, només cal fer ús de 2 comandes Docker dins del directori amb els arxius esmentats a l'apartat anterior.

La primera comanda a utilitzar és *"docker build -t mapreduce ."*. Aquesta comanda crea la imatge de docker a partir del Dockerfile present al directori amb el nom *mapreduce*.

 Universitat Autònoma de Barcelona	TextCounter
	Manual d'Usuari

Amb la imatge creada, ja podem fer ús de la segona comanda “*docker run -it -v <path-directori-fitxers>:/data mapreduce ./TextCounter fitxer1.txt fitxer2.txt*” on *<path-directori-fitxers>* s’ha de substituir per el path absolut del directori amb els fitxers de dades; i *fitxer1.txt* i *fitxer2.txt* s’han de substituir pels noms dels fitxers a utilitzar presents dins d’aquell directori. Es poden afegir més fitxers al final de la comanda.

3.3. Obtenció dels resultats


Una vegada executades les dues comandes explicades a l’apartat anterior, el software entrarà en funcionament i realitza els càlculs sobre els fitxers especificats.

El resultat es mostrarà per pantalla i estarà format pel títol del fitxer seguit de les lletres que apareixen en aquest acompanyades dels percentatges de paraules en el fitxer que les contenen.

3.4. Exemple

A continuació mostrem un exemple d’execució del software Textcounter executant la comanda “*docker run -it -v \${PWD}\data:/data mapreduce ./TextCounter example.txt quijote.txt*”. El resultat mostrat per pantalla és el següent:

```
example.txt:
t : 20.37%
o : 20.99%
p : 12.35%
a : 52.47%
n : 25.31%
d : 12.96%
e : 46.3%
c : 17.9%
u : 12.96%
i : 17.9%
l : 30.25%
r : 21.6%
s : 29.63%
v : 6.17%
ã : 6.17%
m : 15.43%
g : 8.02%
b : 8.64%
h : 1.23%
f : 3.7%
y : 1.85%
```

 Universitat Autònoma de Barcelona	TextCounter
	Manual d'Usuari

q : 1.23%
z : 0.62%
x : 1.23%
j : 0.62%

quijote.txt:
p : 9.13%
r : 23.77%
i : 18.02%
m : 11.28%
e : 50.3%
a : 39.95%
t : 15.3%
d : 21.03%
l : 20.76%
n : 26.04%
g : 4.46%
o : 35.48%
s : 29.81%
h : 5.09%
q : 8.54%
u : 20.17%
j : 2.77%
c : 14.83%
í : 3.28%
ó : 2.41%
y : 6.6%
f : 1.97%
b : 6.18%
v : 4.44%
z : 1.7%
á : 1.87%
ú : 0.33%
ñ : 1.12%
x : 0.08%
é : 1.89%
ü : 0.02%
ï : 0.0%
w : 0.0%
à : 0.0%
ù : 0.0%