

NLP

5IABD

January 28th 2020

Génération de biographie

Les Quiromanciers

Valentin POIRIER - Olivier QUERET

Manitra RANAIVOCHARISON - Yasmine TOUZENE - Bernard VONG

Summary

1. Les Quiromanciers

- Notre concept
- Notre cible

2. Data processing

- Instagram Factory & Model
- Wiki Factory
- Model GPT2

3. Overview

- Architecture
- Streamlit UI
- Diagramme de séquence

4. Démo

Notre concept

Les histoires



Notre concept

Les histoires personnalisées

-
La chiromancie



Notre concept

“

Romancer une biographie personnalisée
selon les goûts et expériences de chacun

”

Les Quiromanciers

Notre cible



Persona #1

Paula



À 26 ans, Paula aime **voyager** et accorder du temps pour lire



Elle est **influenceuse** et en profite pour faire le tour du monde



“J’adore lire, c’est pour ça que plonger dans l’**univers des livres** me passionne”

Notre cible



Persona #2

Jean



À 31 ans, Jean se passionne pour les **plats du monde** et partage sur **Instagram** ses créations



Il est **chef cuisinier** à New York dans son restaurant étoilé “Encore”

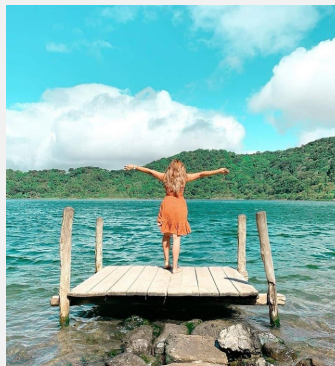


“Je me demande toujours ce que je serais devenu, si je n’étais pas chef”

Instagram Factory



API Instaloader



Location : géo-localisation de la photo

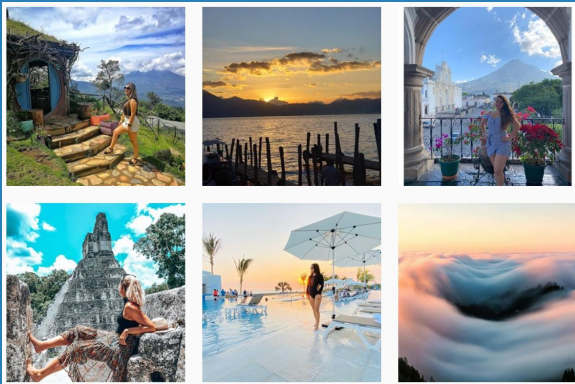
Caption : légende générée automatiquement par Instagram

	User_Name	Location	Caption
400	guatemala_magica	antigua-guatemala-la-ciudad-colonial	one or more people, people standing, shoes and...
531	guatemala_magica	sababa-san-pedro	sky, cloud, mountain, outdoor, nature and water
636	guatemala_magica	casa-blanca-glamping	outdoor
722	guatemala_magica	barriletes-gigantes-de-santiago-sacatepequez-g...	one or more people, people standing and outdoor

Instagram Classification

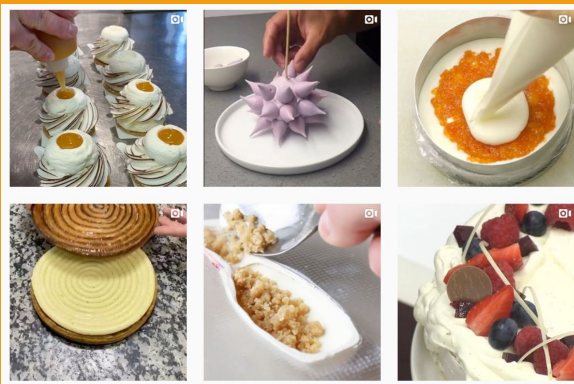
Explorer

mountain outdoor
sky cloud



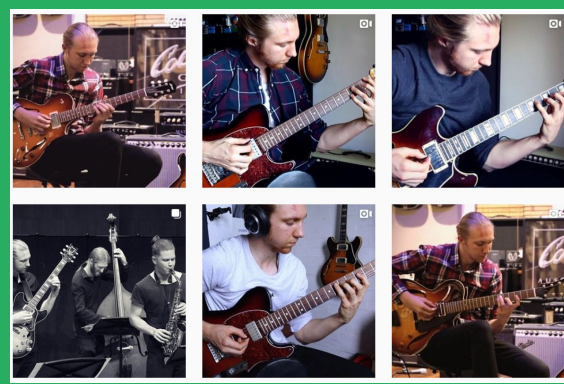
Foody

food drink
coffee eating



Musician

music musical
playing instruments



Instagram Classification

Bag of words

hat	hola	house	indoor	mountain	nature	night	ocean
7	1	3	2	584	275	29	159

Analyse de similarité

	hat	hola	house	indoor	mountain	nature	night	ocean
Mountain	0.05	0.09	0.01	0.02	1.00	0.43	0.15	0.17

Instagram Classification

Un score pour chaque label

User_Name	Location	traveler	food	musician	travel	sky	nature	mountain	guitar	music	instruments
chef.etchebest	10	True	2.500208	1.006909	1.000000	1.000000	1.977207	1.000000	1.000000	1.006845	1.000000
guatemala_magica	4	False	1.375565	1.000000	1.000454	3.419936	3.552092	2.607680	1.000000	1.000000	1.000000
jessieware	14	True	1.661514	1.005115	1.004866	1.050339	1.996483	1.028490	1.000000	1.027132	1.000000
joshmeader22	2	False	1.651556	1.649767	1.000000	1.000000	1.937537	1.000000	1.991029	2.193565	2.859408
nicolescherzinger	48	True	1.513237	1.000000	1.004110	1.432264	2.336702	1.152795	1.000000	1.013752	1.000000
reneredzepinoma	19	True	3.424338	1.000000	1.062420	1.017242	1.826068	1.028519	1.000000	1.025080	1.000000

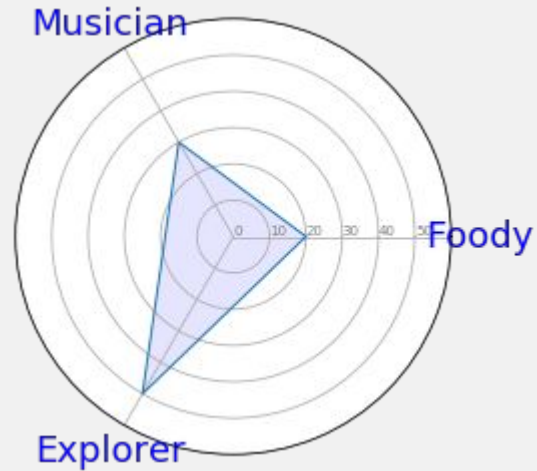
Instagram Classification

Résultat de la classification

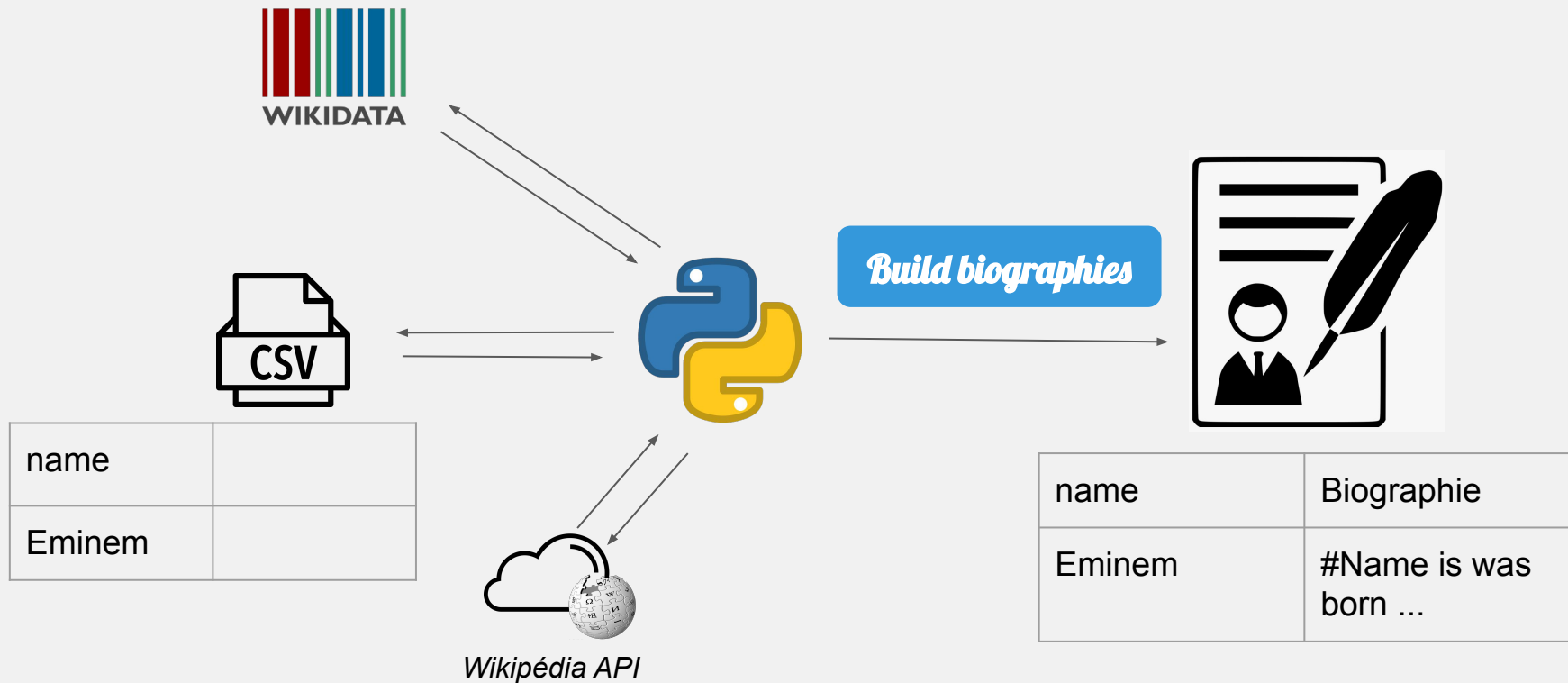
Explorer

Foody

Musician



Wiki Factory



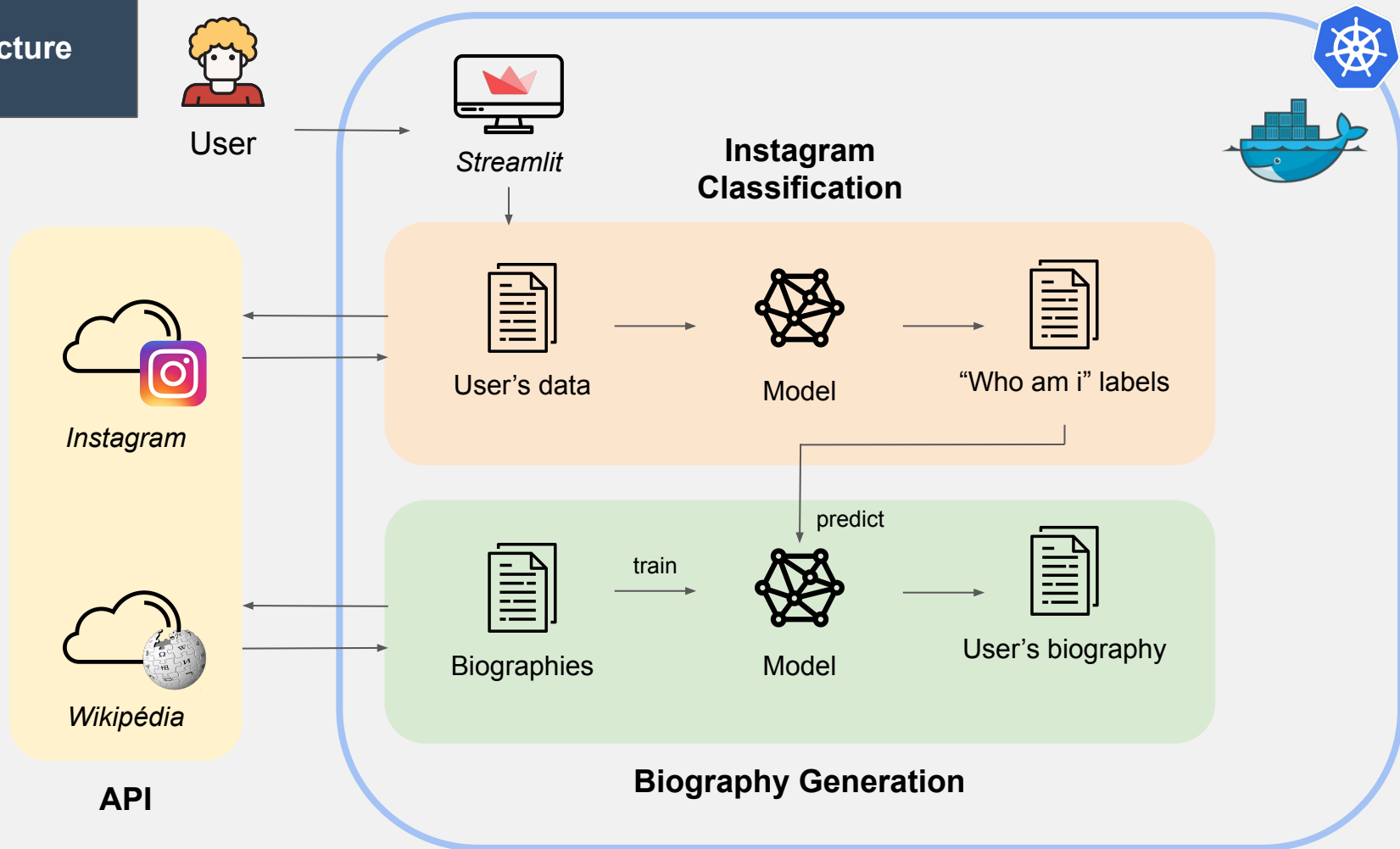
get person name

searchPages

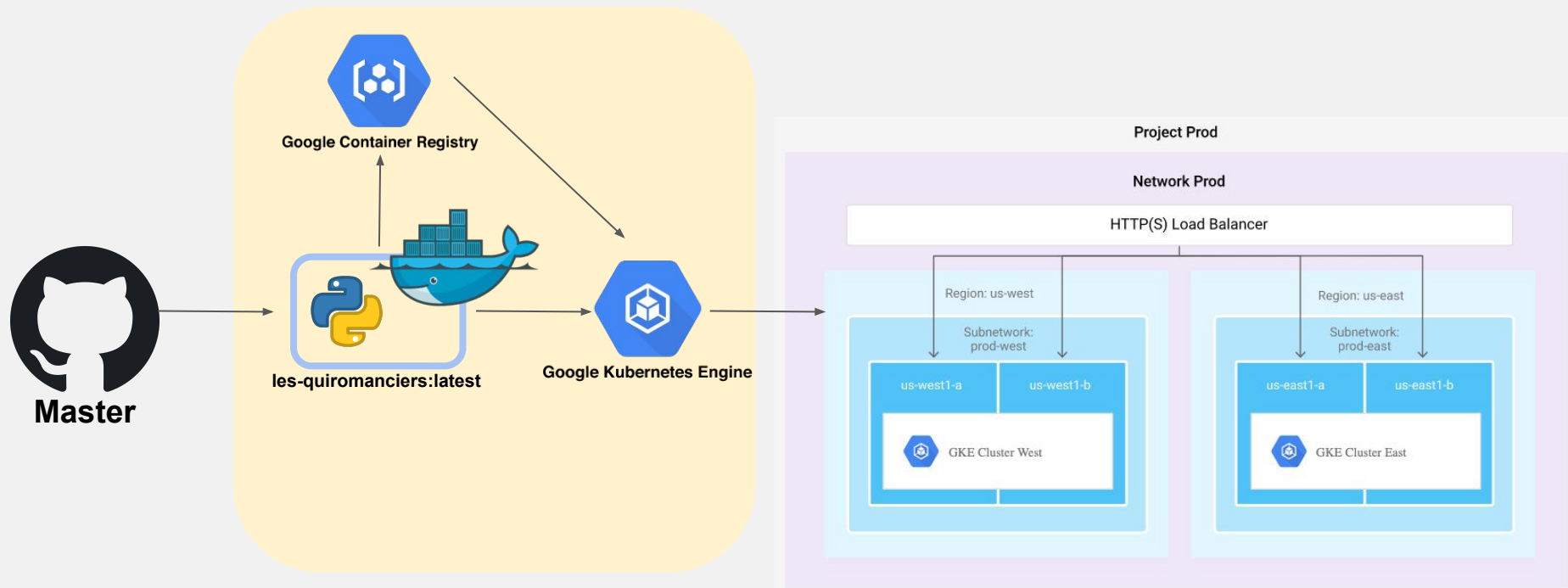
Get Biographies

<https://linkedwiki.com/query/lesQuiromanciers>

Architecture

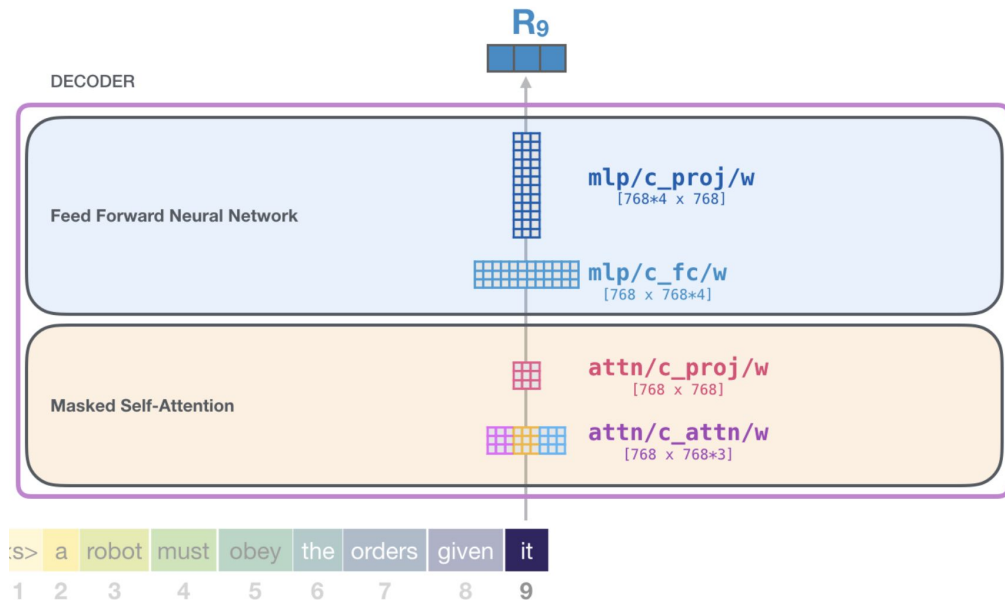
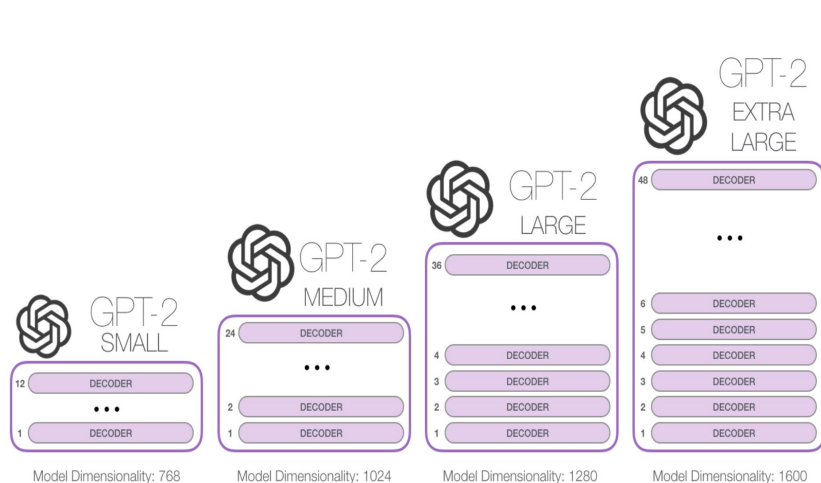


Build - Publish - Deploy



https://github.com/Harisonm/Quiromanciers/runs/409380357?check_suite_focus=true

GPT2 Architecture



GPT2

Features

Labels

	position: 1	2	3	4
Example:				
1	robot	must	obey	orders
2	robot	must	obey	orders
3	robot	must	obey	orders
4	robot	must	obey	orders

must
obey
orders
<eos>

Queries

robot	must	obey	orders
-------	------	------	--------

X

Keys

robot	must	obey	orders
robot	must	obey	orders
robot	must	obey	orders
robot	must	obey	orders

=

Scores (before softmax)

0.11	0.00	0.81	0.79
0.19	0.50	0.30	0.48
0.53	0.98	0.95	0.14
0.81	0.86	0.38	0.90

Scores (before softmax)

0.11	0.00	0.81	0.79
0.19	0.50	0.30	0.48
0.53	0.98	0.95	0.14
0.81	0.86	0.38	0.90

Apply Attention Mask

Masked Scores (before softmax)

0.11	-inf	-inf	-inf
0.19	0.50	-inf	-inf
0.53	0.98	0.95	-inf
0.81	0.86	0.38	0.90

Masked Scores (before softmax)

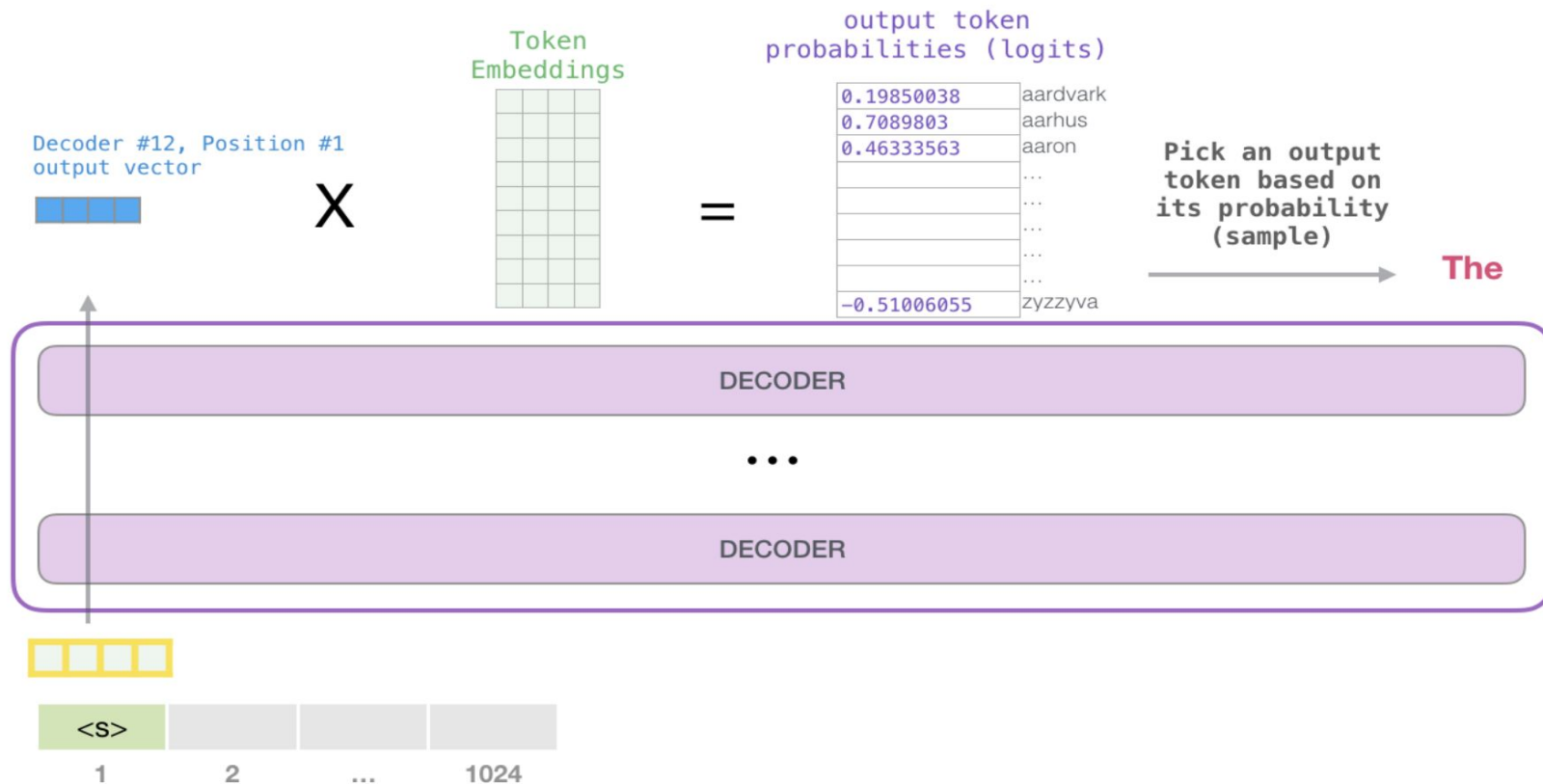
0.11	-inf	-inf	-inf
0.19	0.50	-inf	-inf
0.53	0.98	0.95	-inf
0.81	0.86	0.38	0.90

Softmax
(along rows)

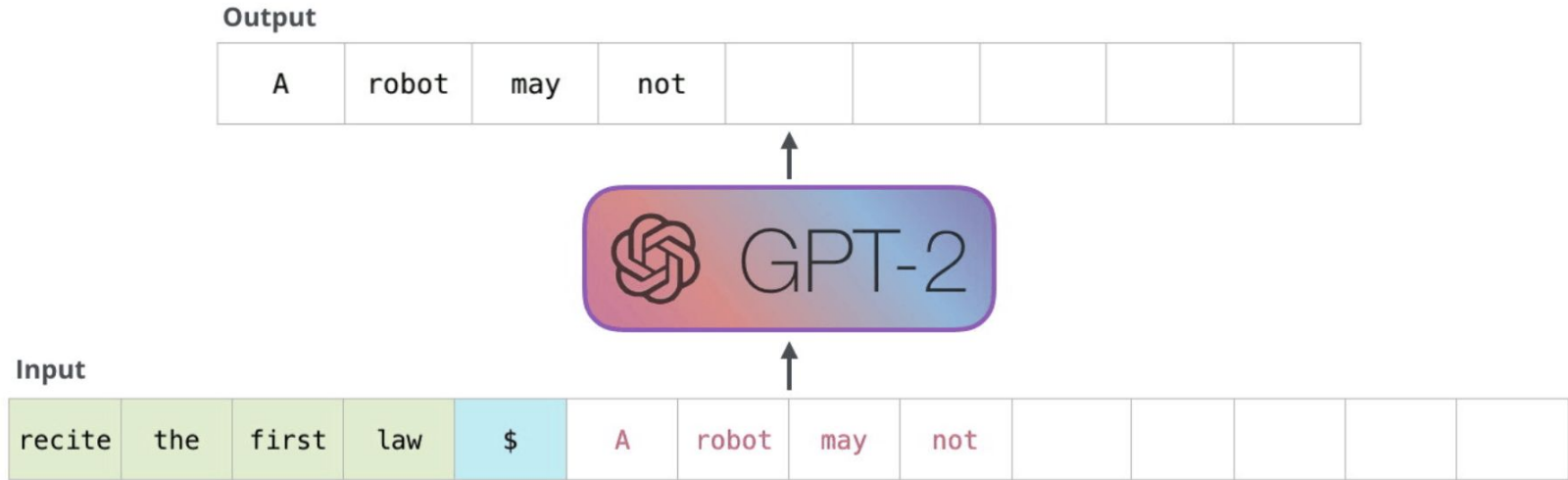
Scores

1	0	0	0
0.48	0.52	0	0
0.31	0.35	0.34	0
0.25	0.26	0.23	0.26

GPT2

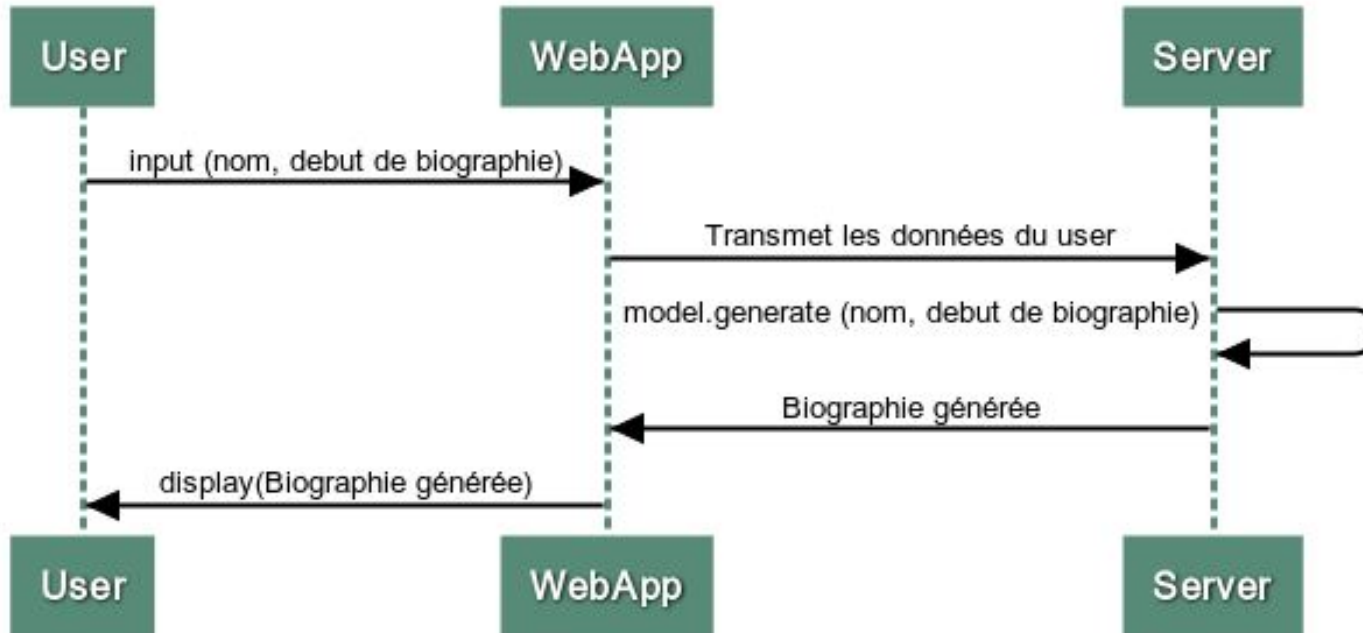


GPT2



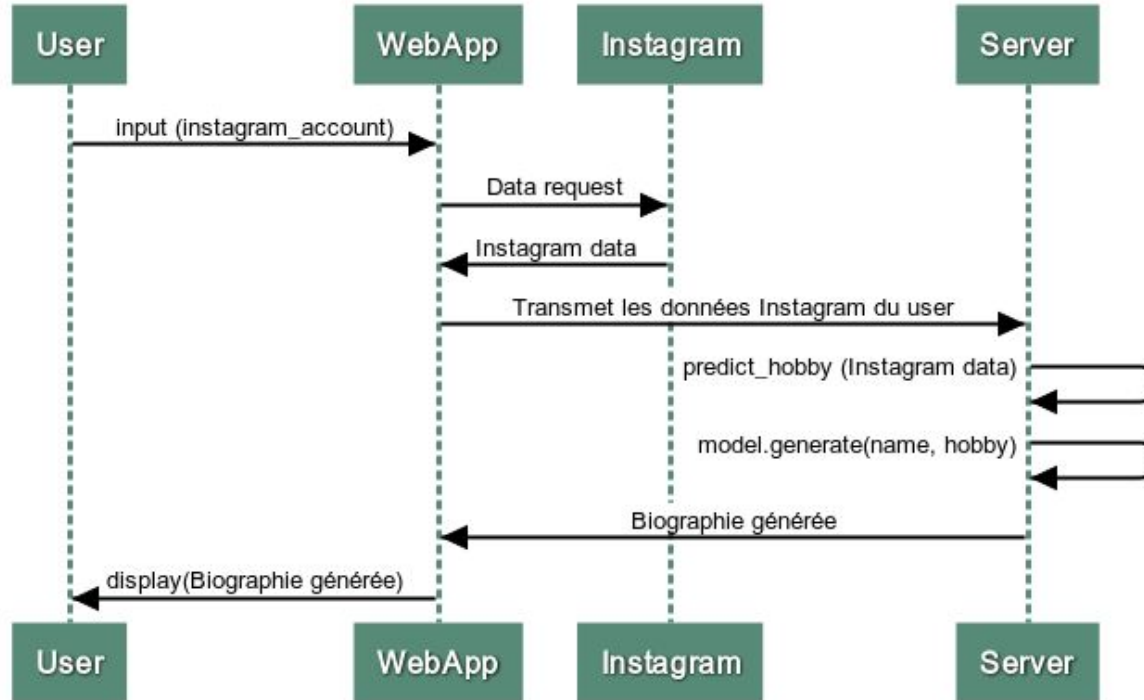
Diagrammes

Diagramme de Sequence sans Instagram

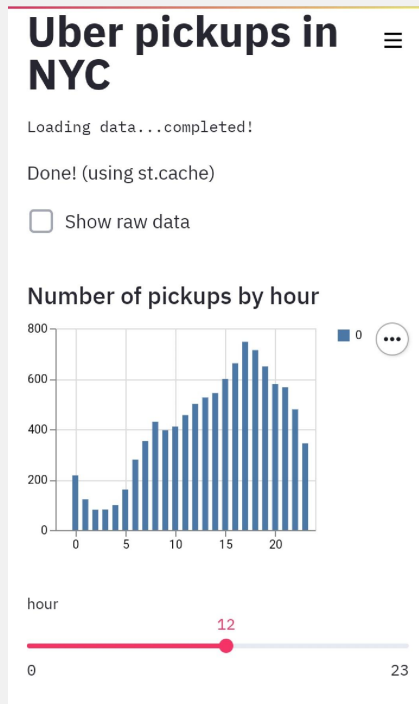


Diagrammes

Diagramme de Sequence avec Instagram



Streamlit



- Framework open source
- Développer des applications pour présenter des projets de **Data Science**
- Interface Web générée par le code python
- Permet une **interactivité** entre l'utilisateur et le code
- Assez récent, communauté en développement

Demo

<http://35.192.181.198:8501/>