



DS n° 3 PHP Site web basique pour visualiser une base de données d'image captioning

Durée : 1h30min.

An: 2020-2021 Niveau: YM2 Semestre: S4

L'accès à Internet via ordinateur est autorisé. Sont interdites à l'usage toutes les applications de tchats – messagerie instantanée, mailing, ou ce qui permet de communiquer avec des personnes. Seule la lecture dans les forums est autorisée ; y poser des questions ne l'est pas. Les surveillants ont le droit de scruter toutes les applications lancées sur votre ordinateur pendant la période de l'évaluation afin de s'assurer du bon respect des consignes.

Vous allez créer un site web dynamique basique sur votre localhost. Pour cette finalité, clonez depuis github, via VScode, sur votre ordinateur, le repo préparé « DS3PHPYM22021prenomnom ». Créez le virtuelhost pour pouvoir visualiser le résultat des étapes de développement de votre application sur votre navigateur.

Liens de téléchargements : (lancez les téléchargements et placez adéquatement la BD dans votre projet)

- Les phrases: https://www.kaggle.com/shtvkumar/karpathy-splits?select=dataset_flickr8k.json
- Les images: https://www.kaggle.com/aladdinpersson/flickr8kimagescaptions

Description succincte de la base de données :

- La base de données est composée de deux éléments :
 - Dossier des images : « /flickr8kimagescaptions »
 - Fichier d'annotations JSON : « /dataset_flickr8k.json »
- Il y a environ 8000 images dans la base de données. Ces images appartiennent à 3 classes (splits):
 « train », « val » pour validation et « test ». Ces classes sont à retrouver dans le fichier JSON (dataset_flickr8k.json).
- À côté des images, il y a un fichier CSV (captions.csv). C'est un fichier à ignorer pour ce projet.
- À chaque image sont associées 5 phrases (chaînes de caractères) dites « captions » ou, descriptions en français. L'origine des descriptions vient d'agents humains qui ont alimenté cette base de données en respectant un ensemble de protocoles et instructions qu'il n'est pas intéressant de rappeler pour réaliser ce projet.
- Le fichier d'annotations (dataset_flickr8k.json) permet de voir comment les captions sont liées à leurs images (voir aussi l'annexe). Il y a des champs à ignorer, vu les besoins de ce projet.

Objectif du projet :

Créer un site web basique pour visualiser des <u>images avec leur descriptions</u> sous un système de pagination adéquat, et par classe.

Étapes de réalisation :

1. Téléchargement de bootstrap (à ce stade, la base de données est déjà téléchargée et est ajoutée au répertoire. Voir l'arborescence dans l'annexe) et préparation du gabarit (html/css vide). Le script principal « index.php » est mis en place avec liaison au bootstrap et un seul contenu « h1 » pour dire Hello Yncrea! – faire le premier commit: BASE DE L'APPLICATION.





DS n° 3 PHP Site web basique pour visualiser une base de données d'image captioning

. .

Durée : 1h30min.

- 2. Parsing de la base de données et affichage du contenu en fonction de trois onglets d'un menu horizontal fixe (voir annexe menu horizontal) :
 - Train images (listing)
 - Validation images (listing)
 - Test images (listing)

Le contenu à afficher à chaque fois est constitué d'une image (occupant une position à gauche), et ses 5 captions, affichées à droite de l'image (voir annexe – contenu image).

On peut penser aux onglets comme des liens vers de nouveaux scripts à créer, ou mieux, un lien vers le même script « index.php » mais avec des informations d'URL adéquates, afin de permettre l'affichage (listing) des images d'une classe (train, val ou test) uniquement. — faire le deuxième commit : MENU HORIZONTAL.

3. Mise en place d'un système de pagination avec possibilité de naviguer vers : (quand c'est possible) les deux premières pages, les deux dernières pages, la suivante et la précédente (voir annexe – boutons de pagination). Le nombre d'images par page est fixé à 4. – faire le troisième commit : PAGINATION 4pp.

Livrables:

Les trois commits (la description donnée en gras aux commits doit être scrupuleusement respectée dans vos projets)



An: 2020-2021

Niveau: YM2

Semestre: S4





ANNEXE

Extrait reformaté du contenu JSON :

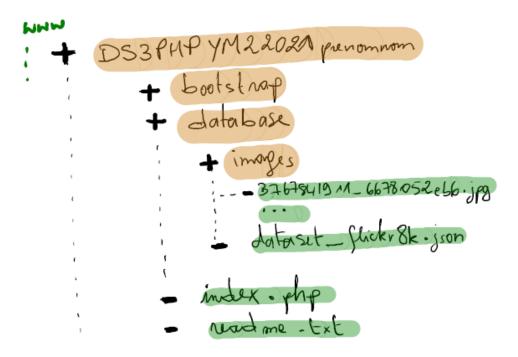
```
"images": [
           "tokens": ["a", "black", "dog", "is", "running", "after", "a", "white", "dog", "in", "the", "snow"],
           "raw": "A black dog is running after a white dog in the snow .",
           "tokens": ["black", "dog", "chasing", "brown", "dog", "through", "snow"],
            "raw": "Black dog chasing brown dog through snow",
           "tokens": ["two", "dogs", "chase", "each", "other", "across", "the", "snowy", "ground"],
           "raw": "Two dogs chase each other across the snowy ground .",
           "tokens": ["two", "dogs", "play", "together", "in", "the", "snow"],
           "raw": "Two dogs play together in the snow .",
            "tokens": ["two", "dogs", "running", "through", "a", "low", "lying", "body", "of", "water"],
           "raw": "Two dogs running through a low lying body of water .",
    "split": "train",
    "filename": "2513260012_03d33305cf.jpg"
```



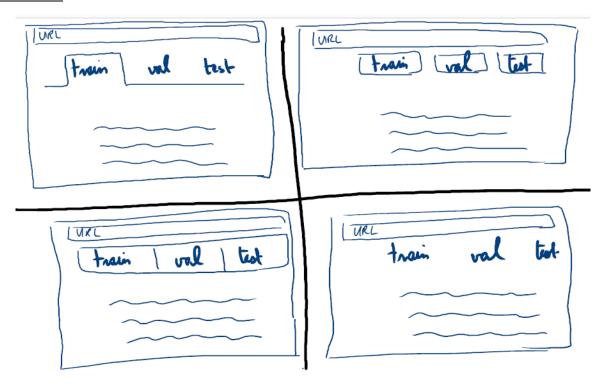


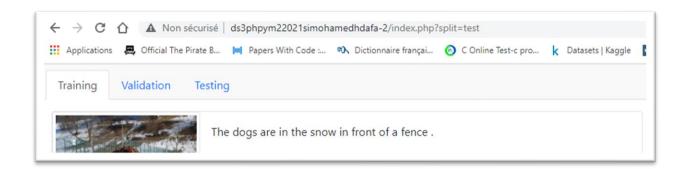
ANNEXE

Arborescence:



Menu horizontal:



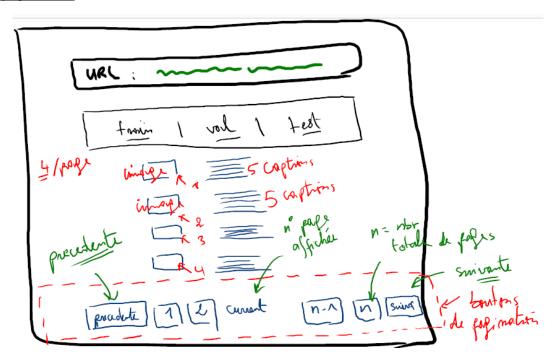






ANNEXE

Boutons de pagination:



Contenu image:



A black dog running across green grass.

A black dog running in grass .

A black dog runs in the grass, its tongue hanging out.

A black dong runs across a grassy field

Small black dog running in grass



A boy jumping off of a dock while his friend watches

Two boys are climbing a wooden platform and jumping into a river .

Two boys jumping off the pier into the water .

Two boys preparing to jump off a pier located on a large body of water .

Two children jump off the dock into the lake .

..