



Fiche d'accompagnement

Parcours : DATA, stockage de la donnée

Séance 2 : Big data, Data center et BDD

Niveaux : cycles 3 et 4

Âge : 10 ans et +

Avec



Sommaire

Objectifs	3
Savoirs	3
Savoir-être	3
Savoir-faire	3
DÉROULEMENT DE LA SÉANCE	4
Présentation des objectifs	4
Étape 1 : Introduction	4
Émergence des représentations	4
Étape 2 : Stocker plus de données ?	4
Approfondissement	4
Étape 3 : Vidéo « Internet est-il dans les nuages ? »	4
Étape 4 : Décryptage et construction de la définition	5
Émergence des représentations	6
Étape 5 : C'est quoi une base de donnée ?	6
Fonctionnement d'une BDD avec DataDecode	7
Étape 6 : Introduction à la manipulation d'une BDD	7
Conclusion	10
Étape 7 : Messages-clés	10
Pour aller plus loin	10

Objectifs

- Comprendre les mécanismes liés au stockage et au partage de grande quantité de données

Savoirs

- Comprendre le principe de DataCenter
- Comprendre le fonctionnement d'une base de données

Savoir-être

- Comprendre les enjeux liés au stockage d'une grande quantité de données

Savoir-faire

- Ajouter des mots et des images dans DataDecode.et s'initier à la manipulation d'une base de données

Thèmes abordés

- Le stockage des données.
- Le data center
- La base de données

Ressources utilisées

- Vidéo «Culture Décode : Internet est-il dans les nuages ?»
- Mission : DataDecode

Durée de l'atelier

- 45 minutes

Modalité

- Activité de manipulation en binômes.

Matériel nécessaire

- Un ordinateur connecté à un vidéoprojecteur ou TNI
- Si possible, un ordinateur par enfant ou par binôme

DÉROULEMENT DE LA SÉANCE

Présentation des objectifs

Étape 1 : Introduction

Dans ce parcours on va s'intéresser aux plateformes contenant beaucoup de données et à l'organisation de ces données à leur accessibilité.

Émergence des représentations

Étape 2 : Stocker plus de données ?

Le but de cette étape est de faire émerger les représentations de chacun pour pouvoir ensuite co-construire une première définition des Datacenters avant de la confronter aux informations apportées par le média de l'étape 7.

Expliquez au groupe que les données sont stockées dans des fichiers, eux-mêmes stockés sur des disques dur. Les fichiers ont une taille qui est mesurée en octet. Les disques durs ont une capacité maximale. Ils ne peuvent pas stocker plus d'octet que cette capacité.

Pour alimenter la réflexion vous pouvez demander à la classe :

- Et si on veut stocker plus de données qu'il peut y avoir dans un ordinateur comment fait-on ?

Vous pouvez conclure en disant que parfois, on peut avoir besoin de tellement de données qu'un seul ordinateur n'est pas capable de toutes les stocker. C'est par exemple ce qui se passe quand on navigue sur le web. Tous les sites internet ne sont pas sur le même ordinateur parce que ça fait beaucoup trop de données. Dans ce cas il faut utiliser plusieurs ordinateurs.

Approfondissement

Étape 3 : Vidéo « Internet est-il dans les nuages ? »

Visionnez tous ensemble la [vidéo sur les datacenters](#). Dites aux jeunes de retenir le plus d'informations possible pour pouvoir en discuter à l'étape suivante.

Étape 4 : Décryptage et construction de la définition

Cette étape permet de construire une définition avec les jeunes. En classe entière, essayez de définir les datacenters. Avant de lire la définition, nous vous conseillons de demander aux jeunes :

- Comment les données circulent-elles ?
- Selon vous, quelles données sont stockées ? Où et comment sont-elles stockées ?

Éléments de réponse	Info + / exemple
Comment les données circulent-elles ?	
Les données numériques circulent à la vitesse de la lumière et à l'aide de tuyaux souterrains et sous-marins. 99% du trafic internet mondial passe par des câbles sous-marins. Il y en a un peu moins de 300 dans les fonds marins.	Ces câbles sont en fibre optique ; ce qui permet de faire circuler les données grâce à la lumière. Les plus longs câbles sous-marins peuvent atteindre plusieurs dizaines de milliers de kilomètres de long.
Selon vous, quelles données sont stockées ? Où et comment sont-elles stockées ?	
Les données en ligne sont stockées sur des ordinateurs auxquels on peut se connecter via internet. On les appellent des serveurs et ils tournent 24h sur 24, 7 jours sur 7. Toutes les données, qu'elles soient personnelles ou non, sont stockées sur des serveurs conservés dans des centres de données appelés "DataCenter". Ces infrastructures sont soumises à des règles (exemple : RGPD) qui permettent de fixer notamment le temps de conservation des données.	Les données sont triées et organisées par mots-clés, puis stockées sur des serveurs. Les serveurs forment ce qu'on appelle le Cloud.

Construction de la définition :

Si on veut rassembler au même endroit et stocker un grand nombre de données, un seul disque dur d'un seul ordinateur ne suffira pas, il faut beaucoup d'ordinateurs branchés les uns aux autres, il faut aussi que ces données soient accessibles aux autres ordinateurs qui se connectent via internet, appelés serveurs. Et dans ce cas, l'endroit où sont entreposés tous ces serveurs qui hébergent les données s'appelle un datacenter.

Émergence des représentations

Étape 5 : C'est quoi une base de donnée ?

Éléments de réponse	Info + / exemple
Sur mon bureau il y a plein de données, est-ce que c'est une base de données ?	
<p>On pourrait croire qu'un dossier qui comporte beaucoup de données est une base de données. Mais si toi tu te retrouves dans toutes ces données, qu'en est-il pour celui qui s'y connecte à l'autre bout du monde pour trouver un article particulier, y arrivera-t-il facilement ?</p> <p>Pour que ces données soient accessibles, à distance, facilement et pour obtenir un résultat rapide malgré la quantité de données stockées on utilise une base de données (BDD) qui permet de trier, ordonner et filtrer l'information. On pourra contacter cette base depuis son ordinateur en envoyant une question précise faite de mots-clés, on appelle cela une requête. (Équivalent de recherche dans DataDecode).</p> <p>Une BDD c'est une structure qui va permettre de ranger toutes les données dans des compartiments étiquetés de façon à ce qu'on puisse la contacter avec de simples mots-clés pour obtenir rapidement un ou plusieurs résultats pertinents.</p>	<p>Une base de données c'est plutôt comme une grande commode avec plein de tiroirs. Chaque tiroir contient une seule catégorie de vêtements.</p> <p>Les vêtements rangés dans la commode, ce sont les données stockées dans la base de données. Quand on cherche des chaussettes dans la commode on sait qu'elles seront dans le tiroir à chaussettes et pas autre part. C'est pareil avec les données elles sont ordonnées dans des catégories bien précises.</p> <p>À l'inverse, quand on veut ranger un t-shirt dans la commode on va le mettre dans le tiroir à T-shirt. Et quand on veut rajouter une donnée on va l'insérer dans la base de données à la bonne catégorie.</p> <p>Et on peut imaginer plein de petits compartiments dans ces tiroirs : les t-shirts à bretelles, les t-shirts sans manches, les t-shirts à manches longues, Plus on est précis, plus on sera efficace pour ranger et extraire les habits. Comme pour les données.</p>

Fonctionnement d'une BDD avec DataDecode

Étape 6 : Introduction à la manipulation d'une BDD

Par binôme, les élèves vont pouvoir tester une création DataDecode et la compléter. Le but est d'enrichir un texte en remplaçant certains mots par des images.

Consignes affichées sur l'écran pendant l'activité :

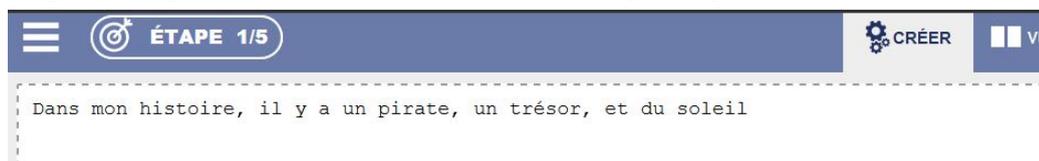
- **Étape 1/5** : Pour commencer jette un coup d'oeil dans la base de données.



On y retrouve plusieurs catégories: des mots, des images et des textes. Intéressons nous aux images, tu peux voir qu'il y en a déjà trois.

Tu peux en ajouter en cliquant sur ce bouton  mais pour l'instant ça devrait suffire.

Regarde à gauche de la base de donnée, il y a déjà une phrase construite :



Le bouton  sert à jouer la création, clique dessus et observe bien.

Quand tu as fini, clique sur  pour revenir aux consignes.

- **Etape 2/5 :** On va maintenant ajouter du code pour qu'au lancement du programme certains mots de la phrases se transforment en image.

On voit que dans la phrase affichée, les différents mots sont parcourus un à un et que le mot "pirate" est remplacé par l'image de pirate présente dans la base de données. Pour comprendre ce qui se passe, nous allons observer

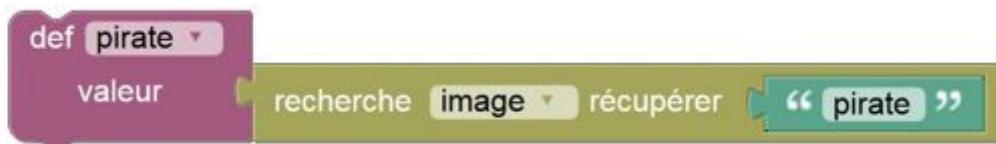
le code de la création, accessible en cliquant sur ce bouton .

Le bloc  sert à faire défiler une tête de lecture sur chaque mot et peut

déclencher un événement  avec par exemple le mot-clé "PIRATE" associé. On pourra donc déclencher un fonctionnement à l'aide de ce bloc.

Ici, lorsque le mot clé est capté on demande de remplacer le mot "pirate" par ce

qui est contenu dans la variable  définie tout en haut du programme à l'aide du bloc :



Ce bloc permet de rechercher une image qui a un certain nom dans la base de données et de la stocker dans une variable.

On a donc récupéré l'image d'un pirate que l'on stocke dans cette variable. Ensuite quand l'événement est déclenché par le mot pirate, on remplace ce mot par le contenu de la variable. Comme il y a bien une image dans cette variable, elle apparaît.

- **Etape 3/5 :** On observe aussi l'ensemble de blocs suivant :



Qui permet de chercher l'image de soleil et de la stocker dans la variable "soleil". Saurais-tu construire un événement qui transforme le mot "soleil" quand il est lu en image de soleil?

Inspire toi du premier événement construit pour ajouter le nouveau.

- **Aide** : Tu vas avoir besoin de ces blocs que tu peux trouver dans la Codothèque:

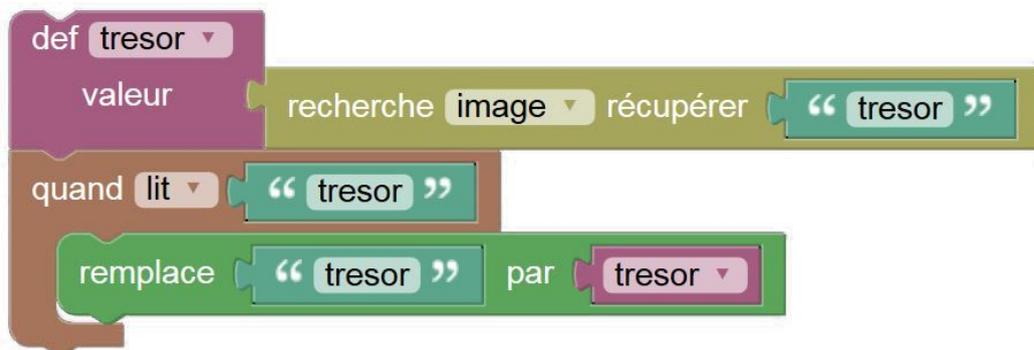


- **Etape 4/5** : Tu as réussi ?

Pour montrer que tu as compris tu peux aussi essayer de transformer le mot trésor en image.

Attention, cette fois-ci aucun bloc n'a été prévu, mais tu pourras trouver la solution dans l'aide.

- **AIDE** :



- **Etape 5/5** : Bien joué, on passe à la suite du parcours.



Clique sur  pour accéder à l'étape suivante.

Conclusion

Étape 7 : Messages-clés

Pour clore cette séance, demandez à tous les participants ce qu'ils ont appris et retenu, en notant leurs idées au tableau puis affichez l'écran contenant les messages-clés suivants :

- Plus il y a de données plus il faudra d'ordinateurs pour les stocker
- Pour héberger beaucoup de données certaines entreprises ont créés des emplacements avec beaucoup de serveurs connectés les uns aux autres, on appelle ces lieux Datacenter.
- Ces données sont ordonnées et accessibles grâce à des bases de données que l'on peut consulter à l'aide de mots-clés.

Pour aller plus loin

- Vidéo 360 sur les data centers de [google](#)