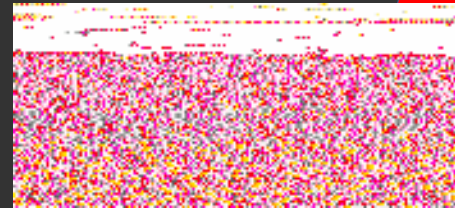


Big Data Open Data

MMI 2 – TP#1 S4



The background features a dark gray field with two large, abstract geometric shapes. A bright yellow triangle is positioned on the left side, pointing towards the bottom-left corner. A light red triangle is located on the right side, pointing towards the top-right corner.

Danielo **JEAN-LOUIS**
Développeur front-end

But du cours

- Sensibilisation aux concepts de Big Data
- Exploitation de données ouvertes
- Initiation au langage Python + Outil Anaconda
- Data visualisation : Python & Javascript (?)

Qu'est-ce que le Big Data ?

Big Data

- Mégadonnées en français
- Concept pensé dans les années 50
- Exploitation massive de données
 - Au-delà du Tb (Terabytes. 1 Tb = 1000 Gb)
 - Facebook génère ~100 Tb/jour / Twitter ~8
- 3-5-7 V

3-5-7 V

- Volume – Echelle des données
- Variété – Diversité des données
- Vitesse – Vitesse de données
- Véracité – Exactitude des données
- Variabilité
- Visualisation – Visualiser les données pour mieux les comprendre
- Valeur – Tirer du profit de ces données

Sources :

- <https://impact.com/marketing-intelligence/7-vs-big-data/> - anglais
- <https://le-datascientist.fr/les-10-v-du-big-data>
- <https://www.oracle.com/fr/big-data/what-is-big-data>

Big Data - Source de données

- Navigation web (clics, actions...)
- E-mails
- Images / vidéos / audios
- IoT (Internet of Things, objets connectés)
- Messages
- [...]
- Vous



Sources :

- <https://www.smartinsights.com/internet-marketing-statistics/happens-online-60-seconds/>

Big Data – Modèle de données

- Structurée
 - BDR (ex : SQL) / Tableur
- Semi-structurée
 - JSON / XML / Fichier de log
- Non structurée (pas de schéma)
 - Vidéo / Image...

Sources :

- <https://www.astera.com/fr/type/blog/structured-semi-structured-and-unstructured-data/>
- https://en.wikipedia.org/wiki/Semi-structured_data – anglais
- https://fr.wikipedia.org/wiki/Informations_non_structur%C3%A9es

Big Data – Le nouvel or noir

- Priorité des entreprises
 - Source de revenus potentielle conséquentes
- Priorité au sein de l'État – E. Macron 2018
- Secteur très porteur dans l'industrie
 - Data-engineer / Data-analyste / Gestionnaire de serveurs...

Sources :

- <https://www.lebigdata.fr/emmanuel-macron-big-data-priorite-europe>

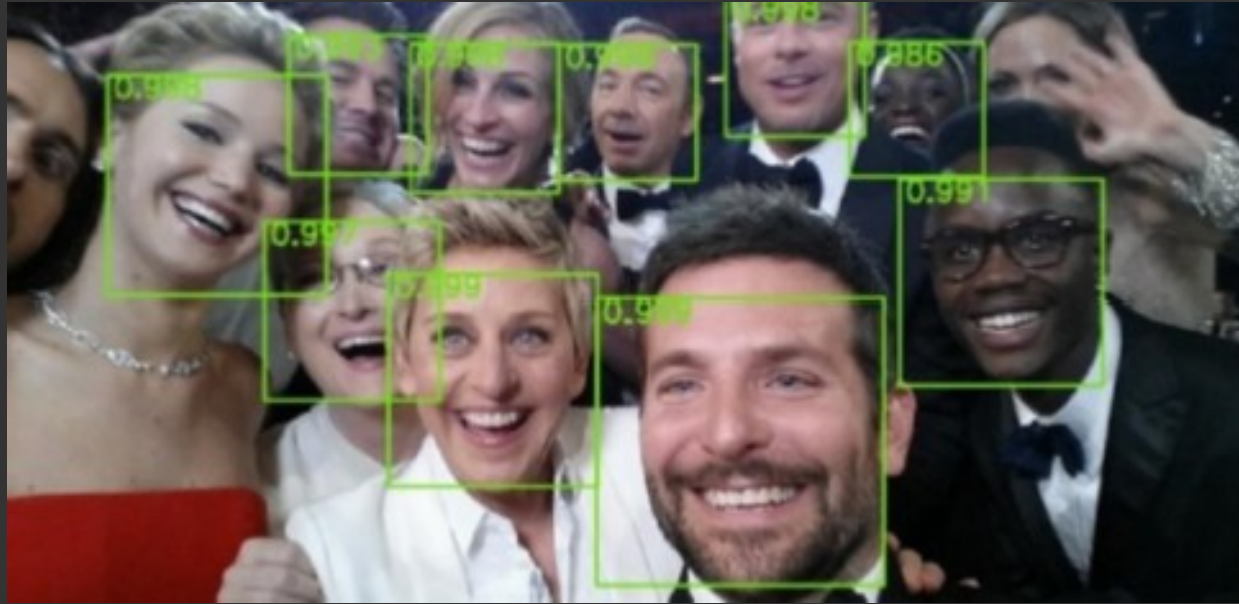
Big Data – Applications

- Gestion des stocks optimisée
- Prix dynamique sur un site d'e-commerce
- Recommandations
- Météorologie*
- Détection de fraudes
- Médecine
- Sport
- ...

Sources :

- <https://www.lebigdata.fr/emmanuel-macron-big-data-priorite-europe>

Big Data – Applications



La reconnaissance faciale : Un exemple parmi tant d'autres mais le plus connu du big data

- Volume
- Variété
- Visualisation

Big Data – Applications



Assistant vocal : une application du big data
(ici Siri d'Apple)

- Volume
- Variété
- Vitesse

Big Data – Applications

Enter both words below, separated by a space.

LYNN *nexttime*

Provided by reCAPTCHA™

Submit

Captcha : une application du big data

Big Data – Applications

**Et vous connaissez-vous des
application du Big Data ?**

Big Data – Enjeux / Défis

- Optimiser la logistique
- Stocker les données
- Récolter des données de qualités
 - Big Data : 80% nettoyage / 20 % d'actions
- Éthique de la donnée
- Sécurité de la donnée
- Sensibiliser les employés (surtout les chefs)



Le Big Data, c'est l'avenir

Questions ?

