



ANALYSER LES IMAGES DE TWITTER AVEC L'API VISAGE D'AZURE

GLOBAL AI NIGHT NANTES — 05SEP2019

BONSOIR À TOUS !



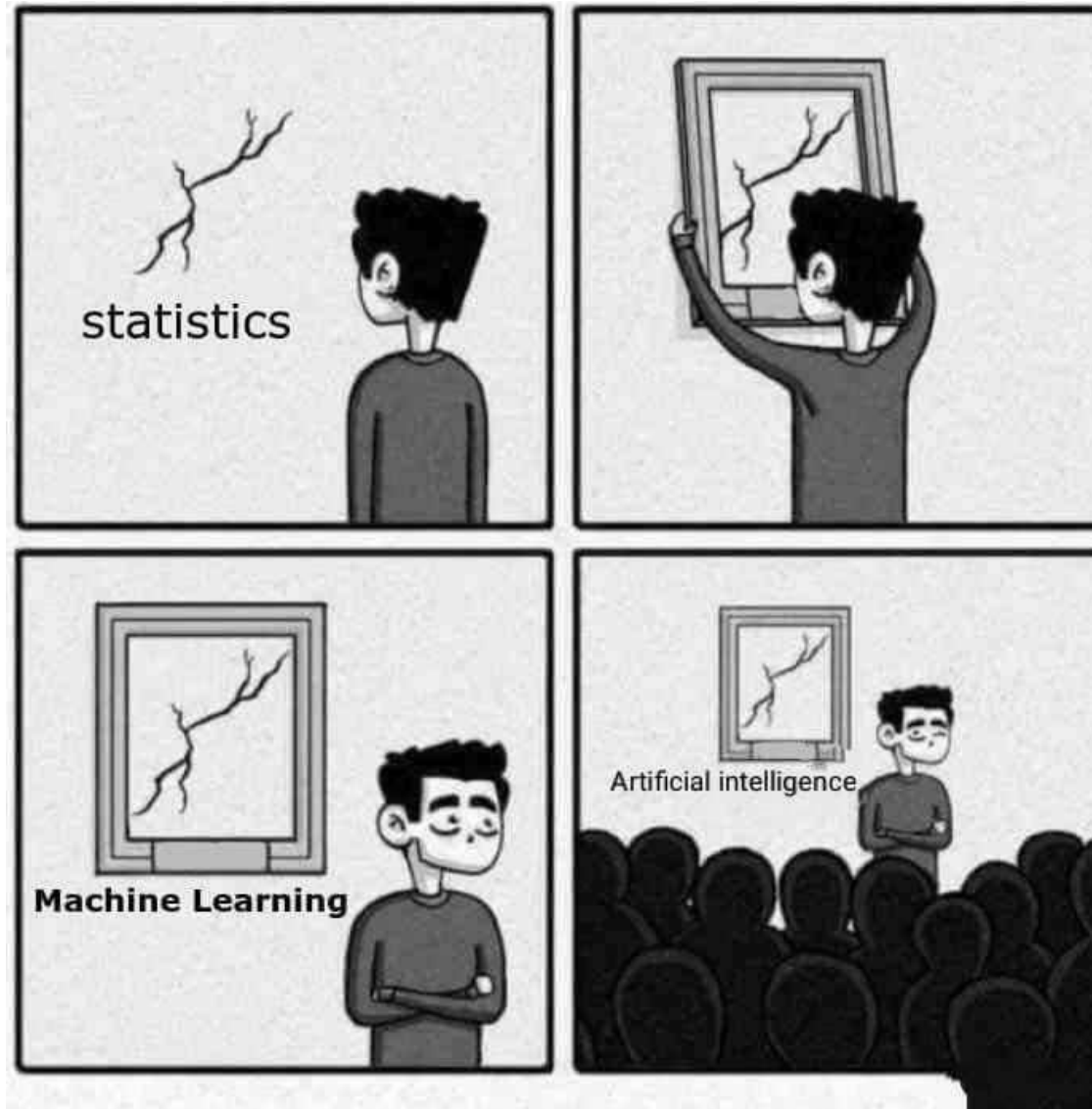
SERGE BLANC

MANAGER DATA SCIENCE

BUSINESS&DECISION NORD-OUEST

- Manager Data Science chez Business&Decision
- Lancement du meetup #DataScienceNantes
- 20 ans dans la Data
- Des projets dans 30 pays
- Vient de poser ses valises à Nantes 😊

STATISTICS → MACHINE LEARNING → AI





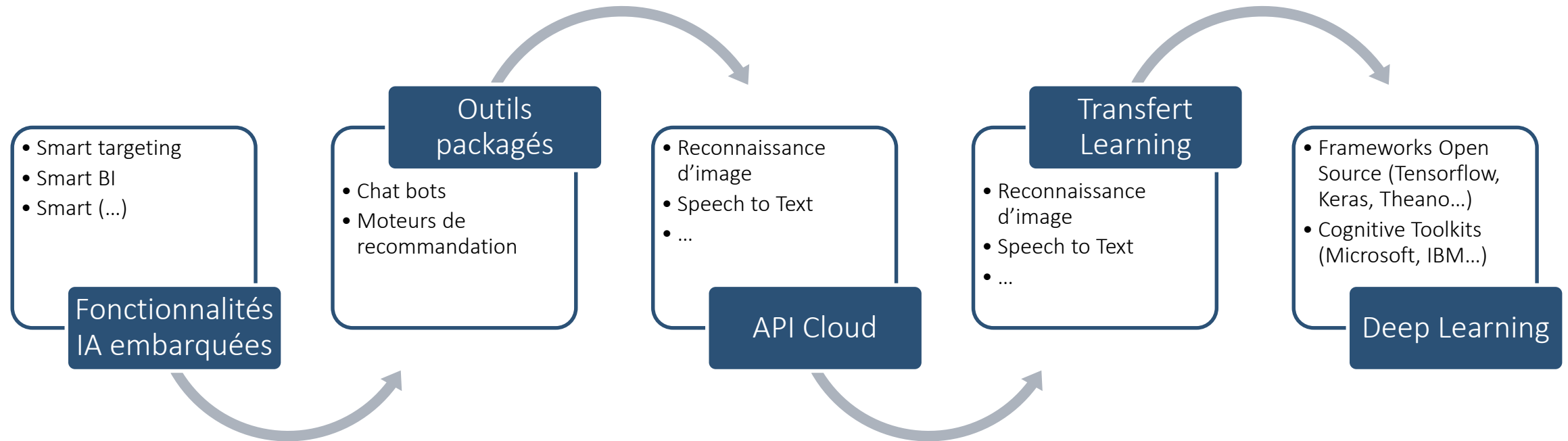
Any sufficiently
advanced technology
is indistinguishable from
Magic...

Arthur C. Clarke

IA GÉNÉRALE VS. IA RESTREINTE



QUELLES TECHNOLOGIES IA METTRE EN ŒUVRE ?



ANALYSE DES RÉSEAUX SOCIAUX

Fréquence d'apparition

Text Mining

Analyse de sentiment

Reconnaissance des
marques

Analyse des produits

Reconnaissance des
visages / célébrités



Analyse de la #ligue1

1 - Identification

#Joueurs



#Club



#Match



2 - Récupération



Programme R & python
tous les lundi matin

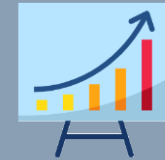


Stockage des résultats
(Texte + Image)

3 - Création de KPI



Nombre de tweet par #



Evolution par journée / Joueur
/ Club

4 - Textmining



Lemmatisation / Nettoyage



Mot influent

5 - Analyse d'image



Webscrapping



Reconnaissance Facial

DÉMO!



PROCESS D'ANALYSE



Base de données de footballeurs



Image du tweet



Détection du visage et reconnaissance faciale

CONSTRUCTION DE LA BASE D'APPRENTISSAGE

WEBCRAPPING SUR GOOGLE IMAGES

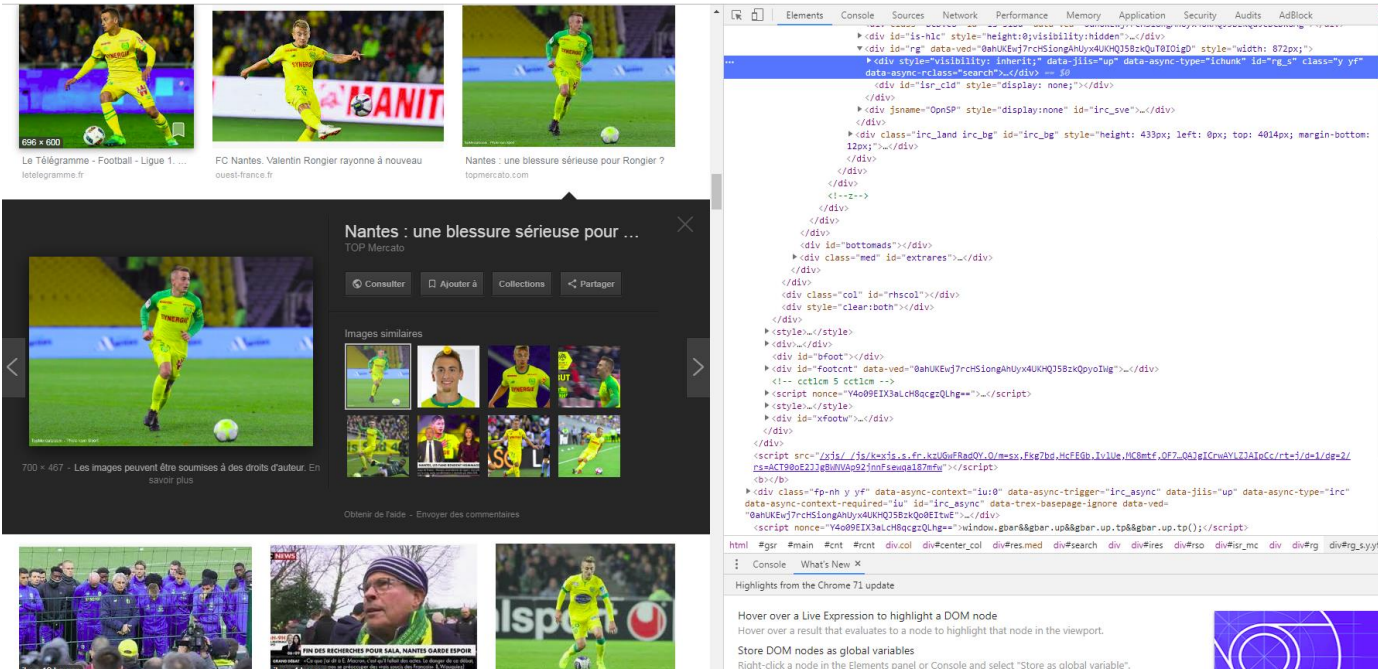
Recherche google images via Python sur un footballeur précis

Identification des URL images dans la page HTML

Téléchargement des images

Vérification de l'intérêt des images (à la main)

Recherche google sur un footballeur (Valentin Rongier, FC Nantes)



CONSTRUCTION DE LA BASE D'APPRENTISSAGE

CRÉATION DU MODÈLE



```
#Identification
SUBSCRIPTION_KEY = '475d73e9a9d14052980111d80a7de822'
BASE_URL = 'https://westeurope.api.cognitive.microsoft.com/face/v1.0'
PERSON_GROUP_ID = 'knowns-persons4'
CF.BaseUrl.set(BASE_URL)
CF.Key.set(SUBSCRIPTION_KEY)
```

```
#Création du groupe de personnes "Footballeurs"
CF.person_group.create(PERSON_GROUP_ID, 'Footballeurs')
```

```
#Création d'un footballeur et de sa base image (ici Killian Mbappe)
```

```
### ATTENTION !!!!! A compiler une seule fois sinon conflit d'identification
## et les images seront en double dans la base
```

```
name = "Mbappe"
user_data = "Joueur PSG"
response = CF.person.create(PERSON_GROUP_ID, name, user_data)
```

```
#identifiant du footballeur
person_id = response['personId']
```

```
#Ajoute toutes les photos du footballeur dans la base personne créée à l'étape précédente
```

```
os.chdir("C:\\Users\\cguichard\\Documents\\evenement_Qlik\\test_footballeurs\\mbappe")
list_fich = os.listdir('.')
```

```
for photo in list_fich:
    print(list_fich.index(photo))
    try :
        CF.person.add_face(photo, PERSON_GROUP_ID, person_id)
    except :
        print('erreur')
        continue
```

```
#entraîne l'algo
CF.person_group.train(PERSON_GROUP_ID)
```

Création d'un groupe de personnes
dans l'API Azure via Python

Création du profil du footballeur
dans l'API

Ajout des images dans le profil du
footballeur

Entraînement du modèle

RÉCUPÉRATION DE L'IMAGE DU TWEET

| Mot Clef | Tweet | Favorited | FavCount | Created | Truncated | ID | Status Source | ScreenName | RT Count | Is RT | Retweet | URL |
|-------------|---|-----------|----------|---------------------|-----------|-------------|--|------------|----------|-------|---------|--|
| @PSG_inside | RT @PSG_inside: <U+0001F1EB><U+0001F1F7><U+0001F30D> #PSGinternational Nos champions du monde @AreolaOfficiel, @kimpembe_3 et @KMbappe ... | FALSE | 0 | 08/10/2018 16:01 | FALSE | 1,04891E+18 | Twitter Web Client | 2345789cm | 257 | TRUE | FALSE | https://twitter.com/i/web/status/1047858654299525120 |



Récupération des tweets via
R et l'API Twitter

Si le tweet contient une PJ,
l'API donne l'URL du tweet

On charge la page web du
tweet dans R grâce à l'URL

On repère l'url de l'image
dans le code source

on ajoute l'url de l'image
dans une nouvelle colonne

UTILISATION DE L'API VISAGE D'AZURE

DÉTECTION DES VISAGES



Détection du visage grâce à l'API

Renvoi de la position des visages dans python sous forme de dictionnaire

Représentation des rectangles grâce à la librairie PIL de Python

UTILISATION DE L'API VISAGE D'AZURE

RECONNAISSANCE DES VISAGES



Un footballeur est candidat pour un visage lorsque le score est supérieur à 0.5.
S'il n'y a pas de candidats viables, alors on note « inconnu ».

LES LEÇONS DU PROJET

On peut adapter un modèle d'IA en quelques jours avec les API Microsoft

On peut créer un prototype à base d'IA en quelques semaines

Constituer la base d'analyse prend 80% du temps

Attention aux biais de sélection !

La courbe d'apprentissage monte vite

AVEZ-VOUS DES QUESTIONS?



