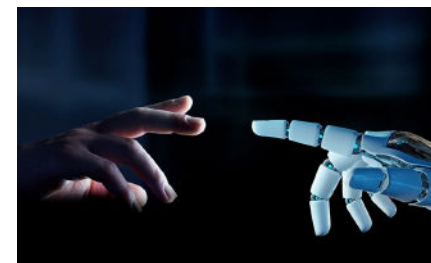


21 janvier 2021

Session de sensibilisation à l'Intelligence Artificielle



Direction régionale des entreprises, de la concurrence, de la consommation, du travail et de l'emploi (Direccte)

Action co-financée par la **Direccte**



Utilisation des Micros :

- Par défaut, les micros des participants sont inactifs et gérés par SCS;
- Entre chaque partie, il y aura une session de Q/R;
- Les participants peuvent intervenir lors des sessions de Q/R en posant leur question par écrit via le module « questions »

Utilisation des Caméras :

- Les caméras de chaque participant seront désactivées
- Seuls les intervenants pourront activer leur caméra s'ils le souhaitent

- **Mot de bienvenue et Introduction**

Bertha Besteiro - Chef Adjointe du Service économique de l'Etat en région

Stéphanie Schohn – Directrice Ecosystème SCS

- **Sensibilisation à l'Intelligence Artificielle**

Jean-Baptiste Bosc – Responsable du Développement France Euranova

Juliette Spinnato – R&D Project Manager

- L'IA, pourquoi maintenant?

- Pourquoi et comment concrétiser un projet IA?

- Parlez-vous IA?

- L'IA en action

- **Présentation de l'accompagnement Diagnostic Flash IA**

Stéphanie Schohn – Directrice Ecosystème SCS

L'entreprise Eura Nova

Une organisation tournée vers l'innovation

CHAMPS D'EXPERTISE

INTELLIGENCE ARTIFICIELLE

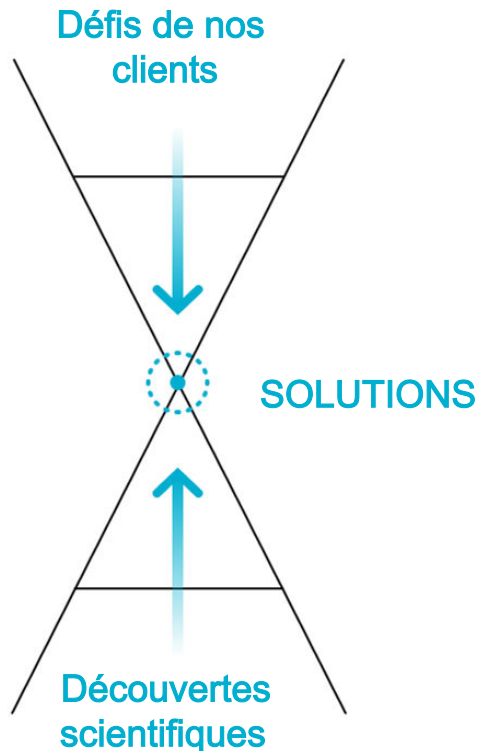
Nous créons des modèles de data science pour créer de la valeur à partir des données

ARCHITECTURE DE DONNÉES

Nous organisons les flux de données et construisons des infrastructures centrées autour des données

INGÉNIERIE LOGICIELLE

Nous développons des logiciels faciles à utiliser



NOS SERVICES

HISTORIQUE

12 ans d'expérience
135 experts
4 produits open source
1 solution commercialisée

CLIENTS

Média, Télécom, Banque, Assurance, Commerce, Pharma, RH
Transport, Aviation, ...

RECHERCHE

2.2 M€ d'investissement par an au niveau du groupe (BE, FR, TN)

L'IA, pourquoi maintenant?

Entre réalité et fiction!

A votre avis...

...est-ce de l'intelligence artificielle?



Un radar automatique?



Hey Siri!

L'assistant vocal Siri?

Les clients ayant acheté cet article ont également acheté



Le moteur de recommandation Amazon?



Des bons de réduction personnalisés?

Les 3 niveaux de l'IA



Qu'y a-t-il sur cette image ? (Est-ce un chat ?)

Pourquoi est-ce un chat ?

Est-ce qu'il monte ou il descend ?

Intelligence
spécifique

Intelligence
générale

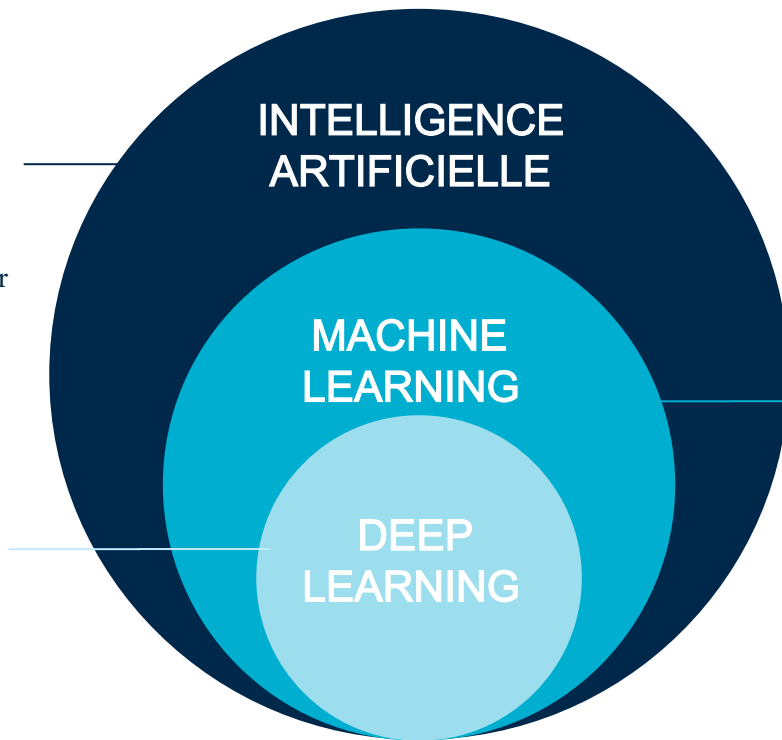
Super
intelligence

Une définition

d'ingénieur!

**INTELLIGENCE
ARTIFICIELLE**
Programmes capables
d'apprendre et de raisonner
comme des humains

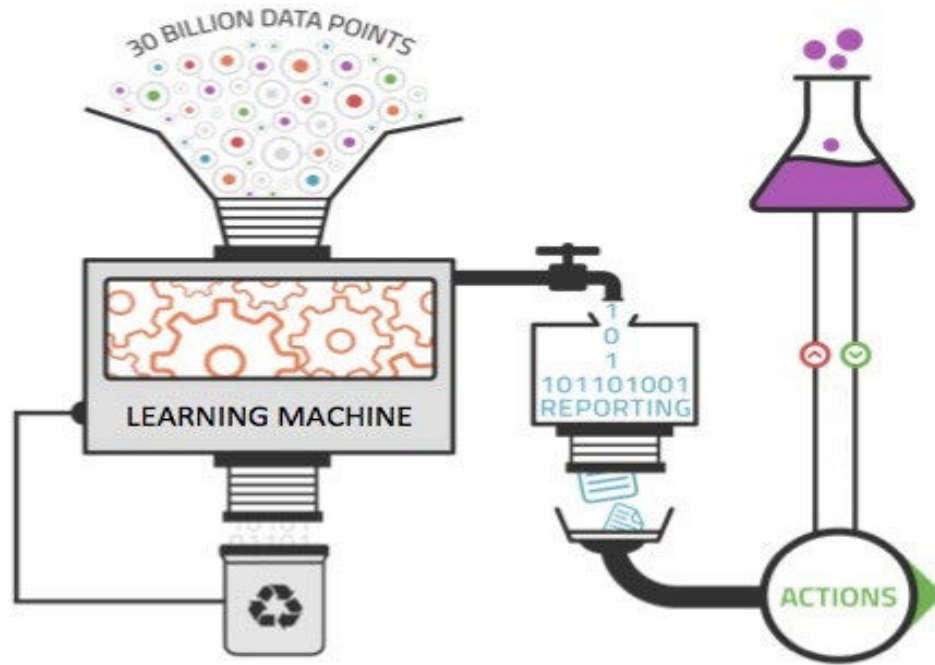
DEEP LEARNING
Sous-ensemble du ML
pour lequel les réseaux
neuraux artificiels
s'adaptent et apprennent
de grandes quantités de
données



**MACHINE
LEARNING**

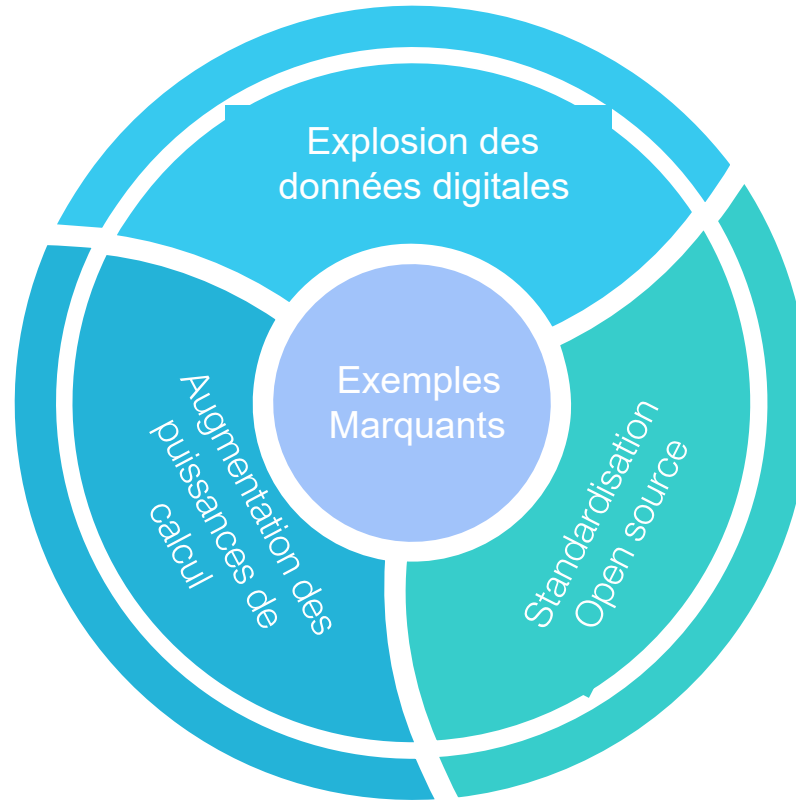
Algorithmes pouvant
apprendre sans être
explicitement
programmés

Donc au final, c'est quoi l'IA?

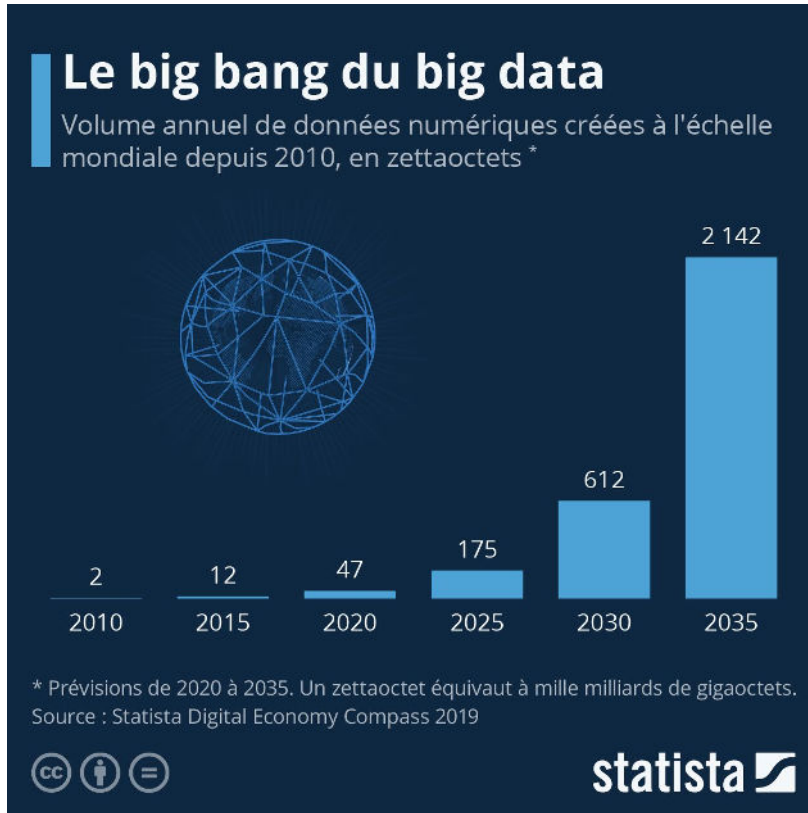


Programmation basée sur des exemples

Pourquoi maintenant?



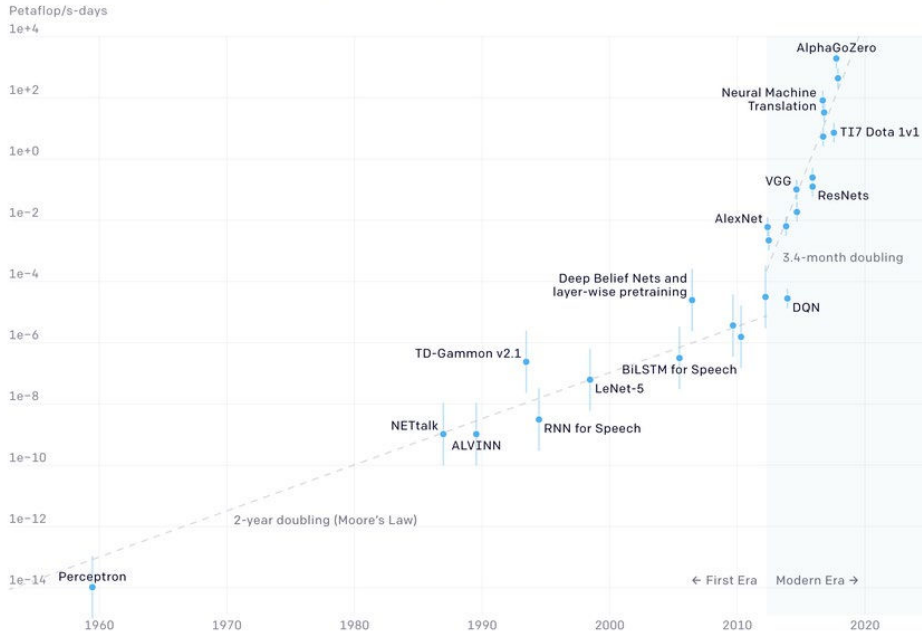
La matière première de l'IA : la data



L'or noir du 21eme siècle!

L'augmentation des puissances de calcul

Two Distinct Eras of Compute Usage in Training AI Systems



- Mais aussi de nouveaux outils pour augmenter les possibilités de calcul :
 - L'utilisation de Machine Virtuelle via le cloud
 - Calcul distribué outil type Spark, Hadoop
 - Et plus généralement des offres de service type cloud computing (AWS, AZURE, OVH, etc.)

OPEN SOURCE

STAT TOOLS



ScalaLab



AI / MACHINE LEARNING / DEEP LEARNING



theano



Caffe



Apache SINGA

FeatureFu



VELES



neon™



DIMSUM

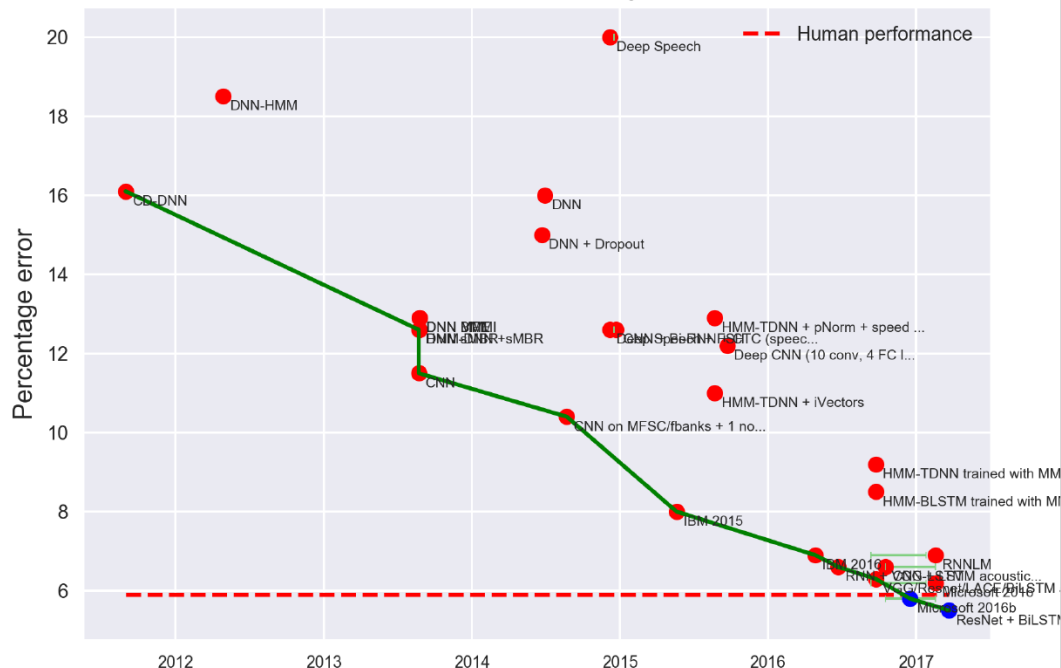
DSSTNE

Aerosolve

Des exemples marquants

L'IA pour la parole (entraînée sur 2000 heures)

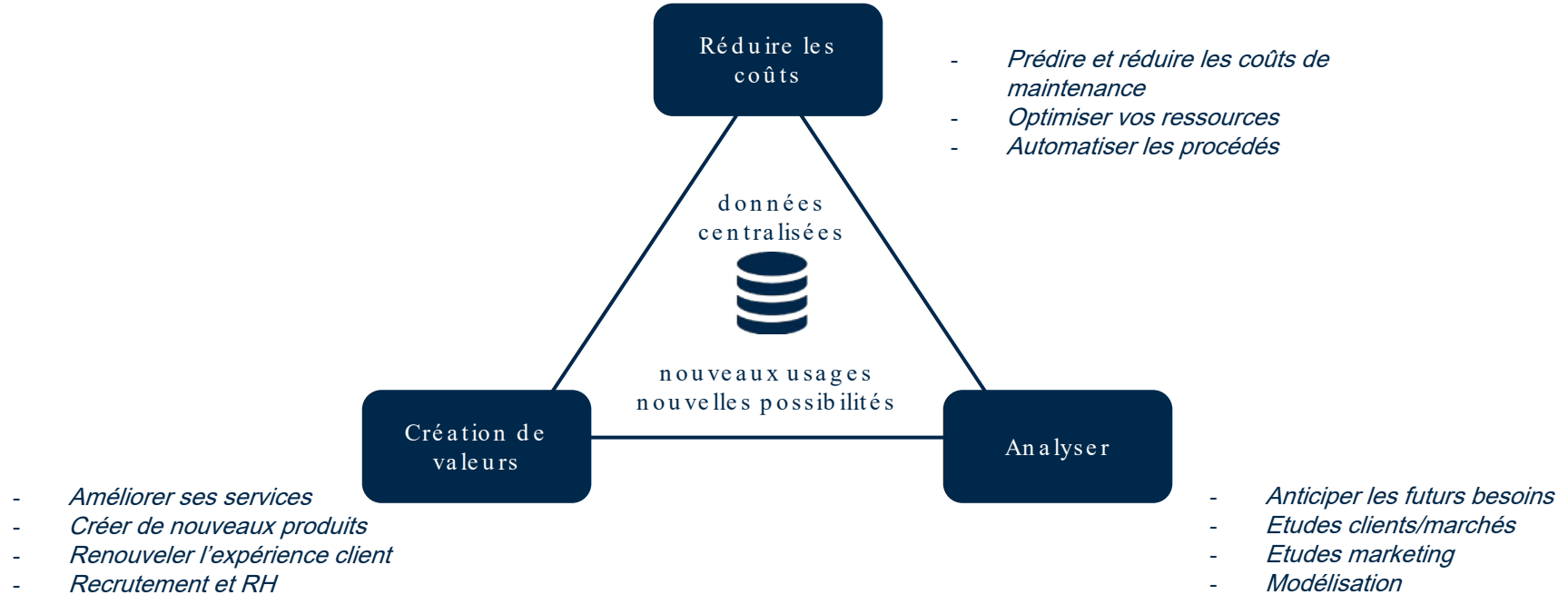
Word error rate on Switchboard trained against the Hub5'00 dataset



Pourquoi et Comment concrétiser un projet IA ?

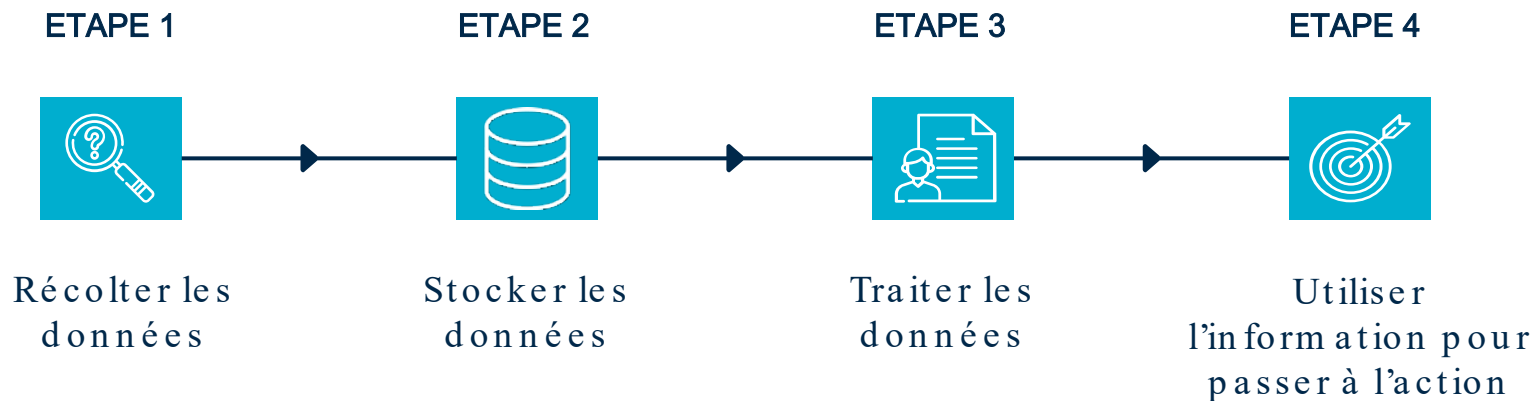
Les usages de l'IA

Les trois vecteurs de progression



Comment

utiliser l'intelligence artificielle ?



L'intelligence artificielle n'est qu'un moyen d'atteindre un objectif.

La connaissance métier est le meilleur atout.

1^{er} Challenge : Automatisation

Intégration dans l'infrastructure existante



Comment puis-je intégrer mon PoC dans mes systèmes ?



Automatiser la collection des données



Ré-entraîner les modèles automatiquement

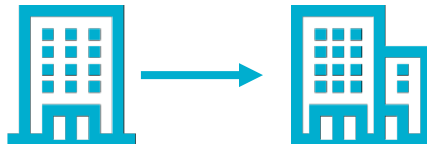


Appliquer et monitorer le modèle



Intégrer la pile technologique

Consommer les prédictions via une architecture orientée service



“Votre PoC a été intégré avec succès.”

2ème Challenge : Organiser les données

Tout s'em mêle et c'est normal n'est-ce pas ?

Questions fréquentes :

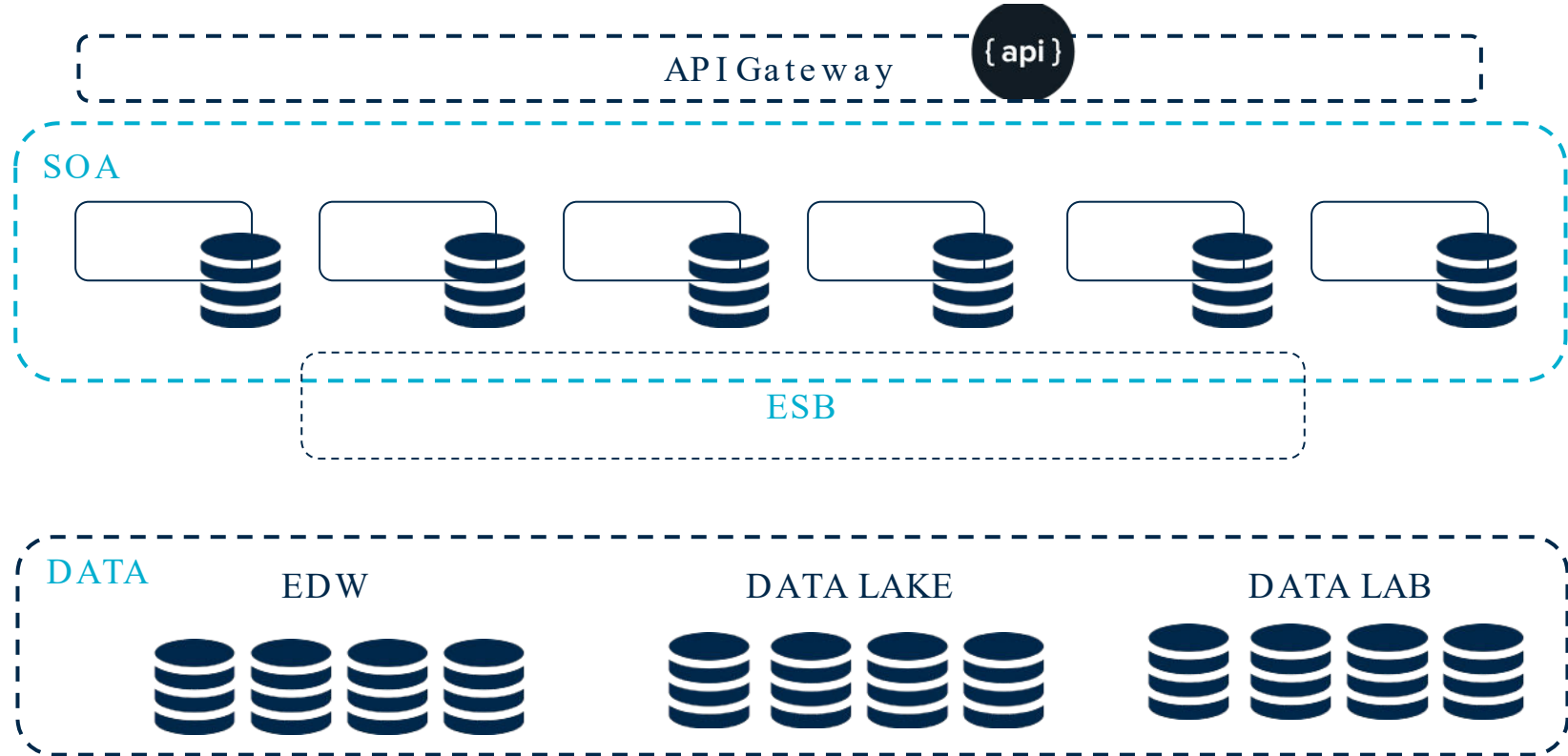
1. Où se trouve la donnée de confiance ?
2. Comment propager les mises à jour, comment s'assurer que les données restent cohérentes ?
3. Comment utiliser mes données dans mes applications ? Comment construire un modèle prédictif ?
4. Comment concilier "BI" et "Analytics" ?
5. Qu'en est-il de mes processus de gouvernance ?

Questions clés à se poser afin de réduire le "time to market" lors de l'intégration de mon PoC.



3^{ème} Challenge : des données et des services

SOA (ou architecture par μ -service) VS data architecture



4^{ème} Challenge : l'attitude Big Data / ML

Compétences et état d'esprit



Afin de construire avec succès des PoCs :

- Les équipes techniques doivent avoir conscience des objectifs business.



- Elles doivent être en mesure de transformer des besoins métiers en solutions techniques.



- La méthodologie Agile est une bonne manière de délivrer des PoCs étape par étape en se focalisant dès le début sur la valeur créée et en s'assurant de l'alignement constant entre besoin et solution.



- La proximité avec le business est fondamentale.



- Les PoCs introduisent souvent de nouvelles technologies et paradigmes qui doivent être maîtrisés.

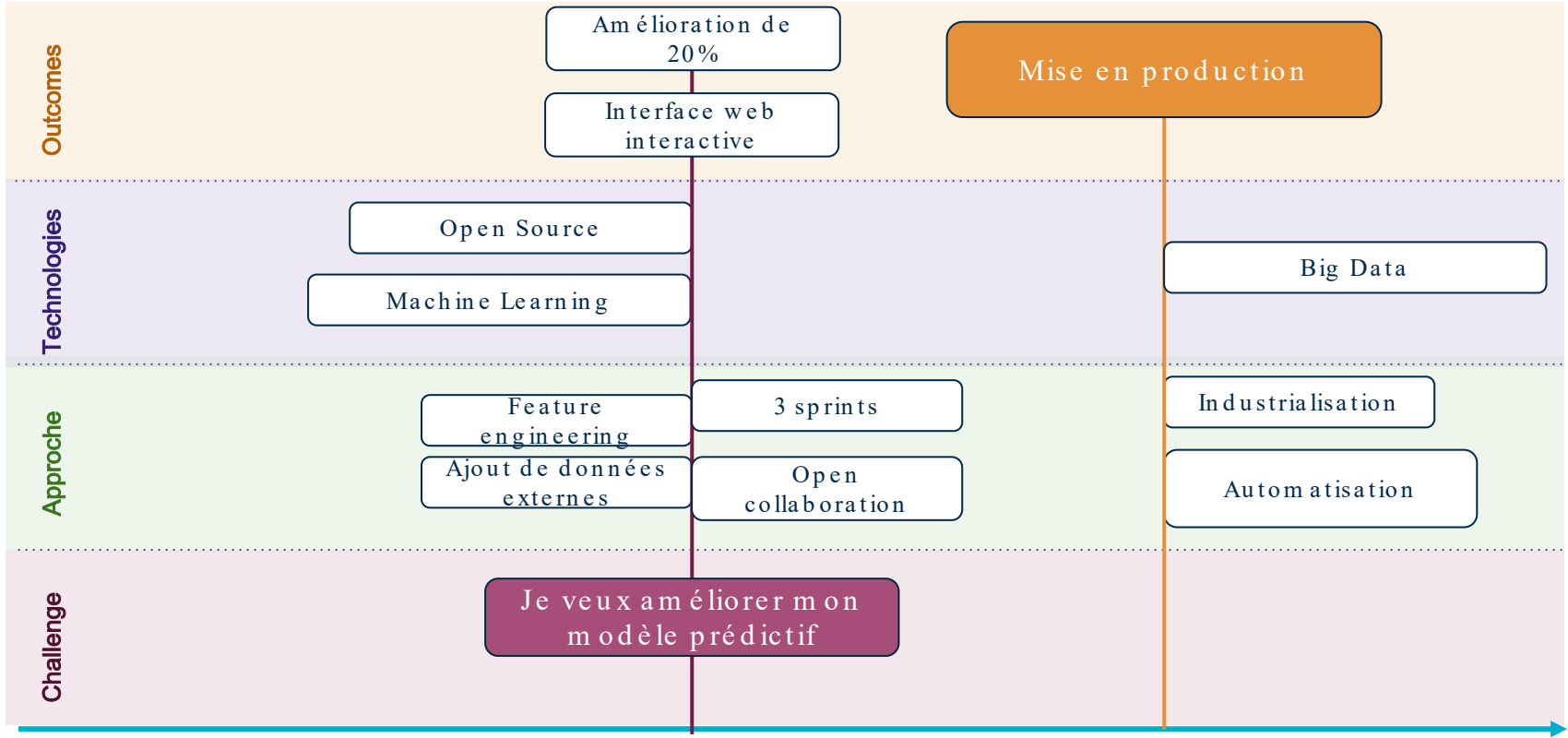
5^{ème} Challenge : gouvernance de la donnée

A l'ère du RGPD



Afin de positionner la donnée au coeur du business, une organisation forte est nécessaire :

- Passer d'une organisation centrée sur les applications à une organisation centrée sur la donnée.
- Les modèles de sécurité doivent s'adapter tout au long du parcours de la donnée à travers l'entreprise.
- Des lignes directrices communes doivent être imposées quant à l'utilisation des données par les différentes équipes.
- Il est crucial de disposer de rôles tels que "Data steward", "Data owner" et surtout CDO/DPO.
- Les concepts et vocabulaires doivent être les mêmes à travers toute l'entreprise.



Timeline

Step 1

Step 2

Parlez -vous IA ?

Des données...

Type de données?



STREAMING



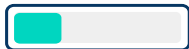
COLD



STRUCTURED



SEMI STRUCTURED



UNSTRUCTURED

Volume de données?



Accès au données?



Labellées?



Supervised learning

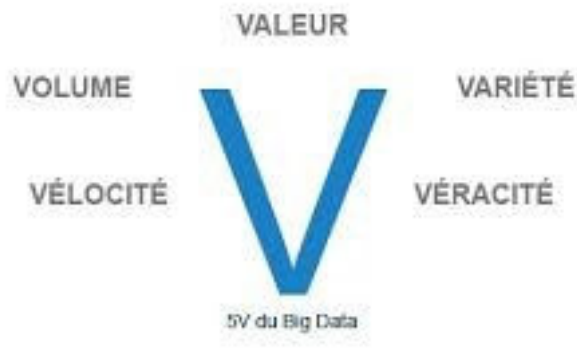


Unsupervised learning



La qualité des données

Premier facteur pour l'utilisation IA



Possibilité et qualité
d'usage



Valeur
Volume
Variété
Véracité
Vélocité



Difficulté de développement et coût
d'utilisation



Valeur
Véracité



Volume
Variété
Vélocité

Où sont vos données?



Document texte



Tableau
Excel, CSV



Photos et vidéo



Capteurs et signaux



data
warehouse



Appareils
connectés



Son et voix

Où sont vos données?

- **Dans l'entreprise**

- Site internet /données utilisateur
- Outils d'entreprise : CRM, outils de gestions, GED, etc.
- Documents numériques
- Base de données
- Capteurs / IOT

- **Acquisition active**

- Développement d'outils d'indexation
- Campagne : Test clinique, sondage, etc.
- Formulaires clients/prospects

- **Open source**

- Données publiques
- Plateforme Spécialisée (Kaggle)
- Données open-sources

- **Démarche type Open Data**

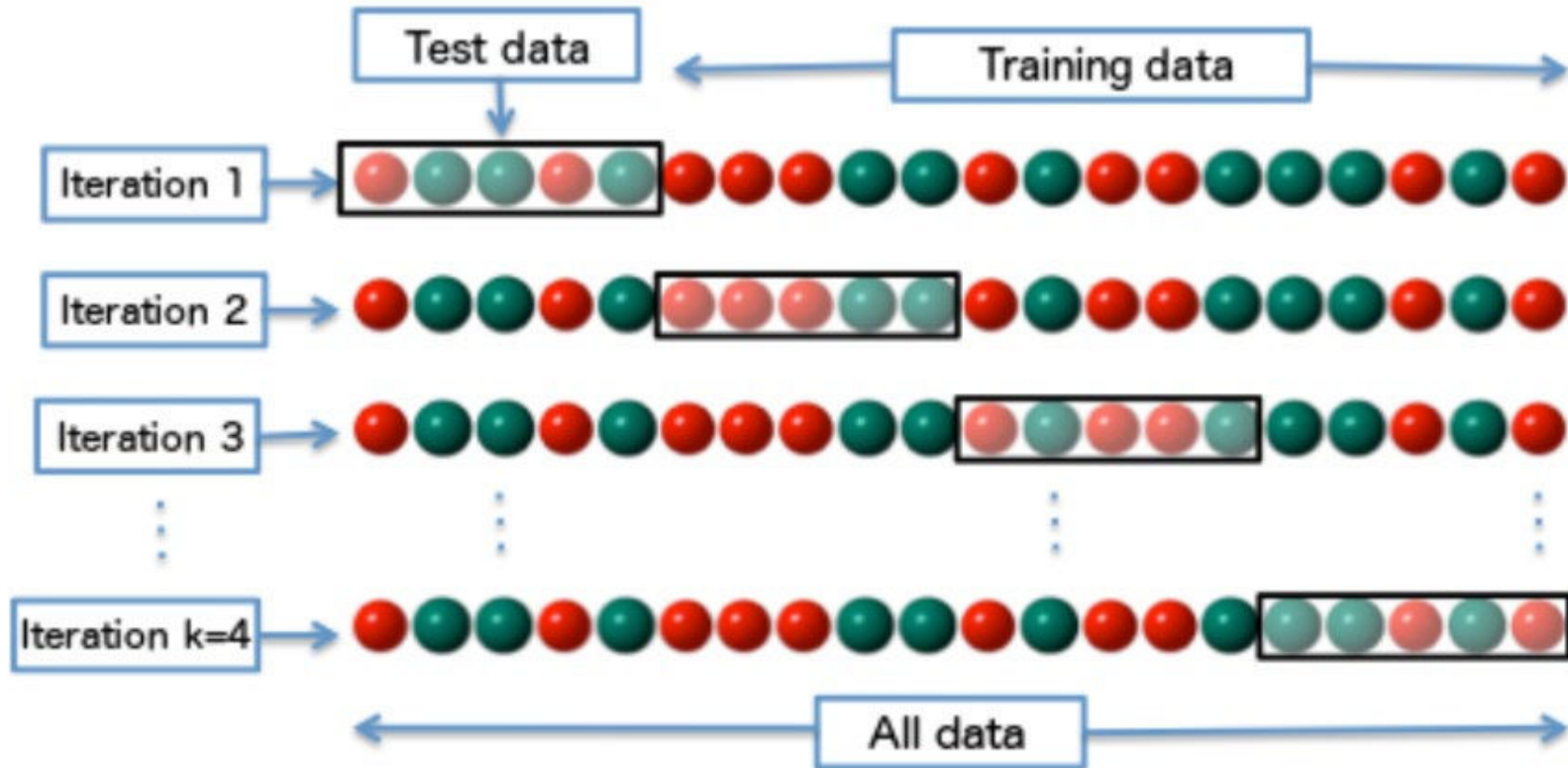
- Partage de données avec des entreprises partenaires
- Voir même concurrentes...
- Données entreprises clientes
- Entreprises spécialisées ex : Dataactivist

Créer de la donnée?

Exemple programme développé pour Bpost sur les panneaux routiers.



Cross validation



Les 3 principales familles

d'apprentissage automatique



L'IA en action

Dans les entreprises

DES CAS D'USAGE DANS DES DOMAINES VARIÉS, UNE MÉTHODOLOGIE SIMILAIRE



Les types d'algorithmes

Machine learning algorithms

Classification

k-Nearest Neighbor →

Support vector machine →

Naive Bayes →

✓ Decision tree →

Random forest →

Regression

✓ Linear regression →

✓ Multivariate linear regression →

✓ Logistic regression →

✓ Cox regression →

Clustering

k-Means →

k-Medians →

Expectation maximization →

✓ Hierarchical →

Self-organizing maps →

Deep learning

Multi-layer perceptron →

Recurrent neural networks →

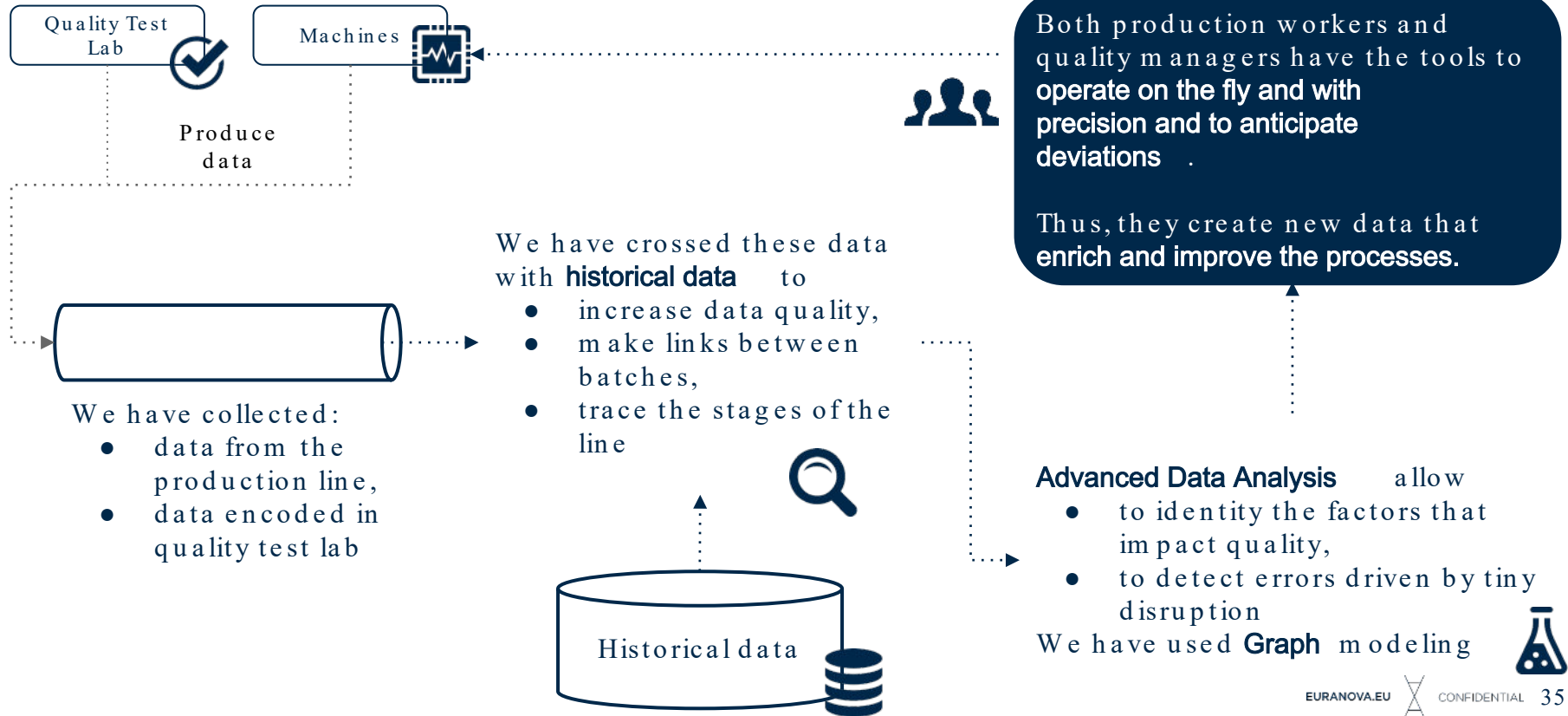
Convolutional neural networks →

Boltzmann machine →

Multi-layer autoencoder →

4.0 Industry: détection des anomalies

Comprendre pourquoi une chaîne de production donne des lots défectueux



Détection instantanée des fraudes

Détecter les comportements frauduleux en temps réel

Our client: an important Belgian bank

Their goal: detect fraudulent behaviour in real time

The project: tap into machine learning techniques and real-time streaming to enhance **fraud detection**

Challenges:

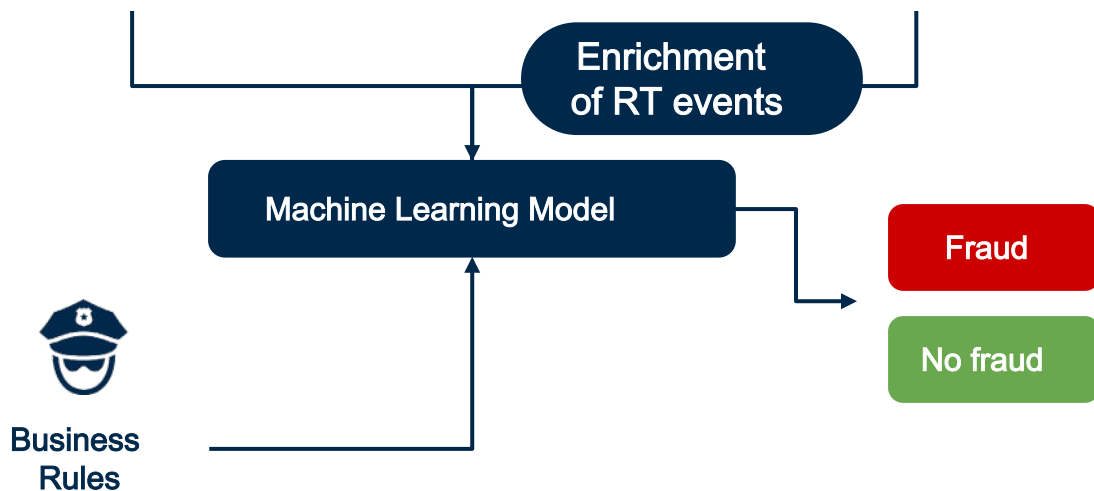
1. cross real-time data from transactions and historical data (former transactions and customers' information)
2. use a machine learning model, and enrich real-time events
3. use business rules from the bank to decide if there is a fraud or not



Real-time data on customers' transactions



Internal data on customers' identity & former transactions



Prévoir les trajectoires des avions

Augmentez vos revenus en faisant plus

Our client: an international air traffic management organisation

Their question: how to safely increase traffic fluidity in airports during adverse meteorological conditions?

The approach: a three-year project split into four tracks to increase airplane traffic by predicting the whole trajectory of the plane at landing and takeoff

The project was carried out in collaboration with a company specialised in wake-related issues for aeronautical applications

Track n °1



Increase the number of landings by predicting the landing times

Track n °3



Increase the traffic fluidity by predicting the crosswinds during the trajectory

Before

A safety distance was imposed
BUT
it was **not adapted** to all circumstances

After

The **buffer is reduced**
AND
there is a **safe separation time** between the planes



Increased traffic

Track n °2



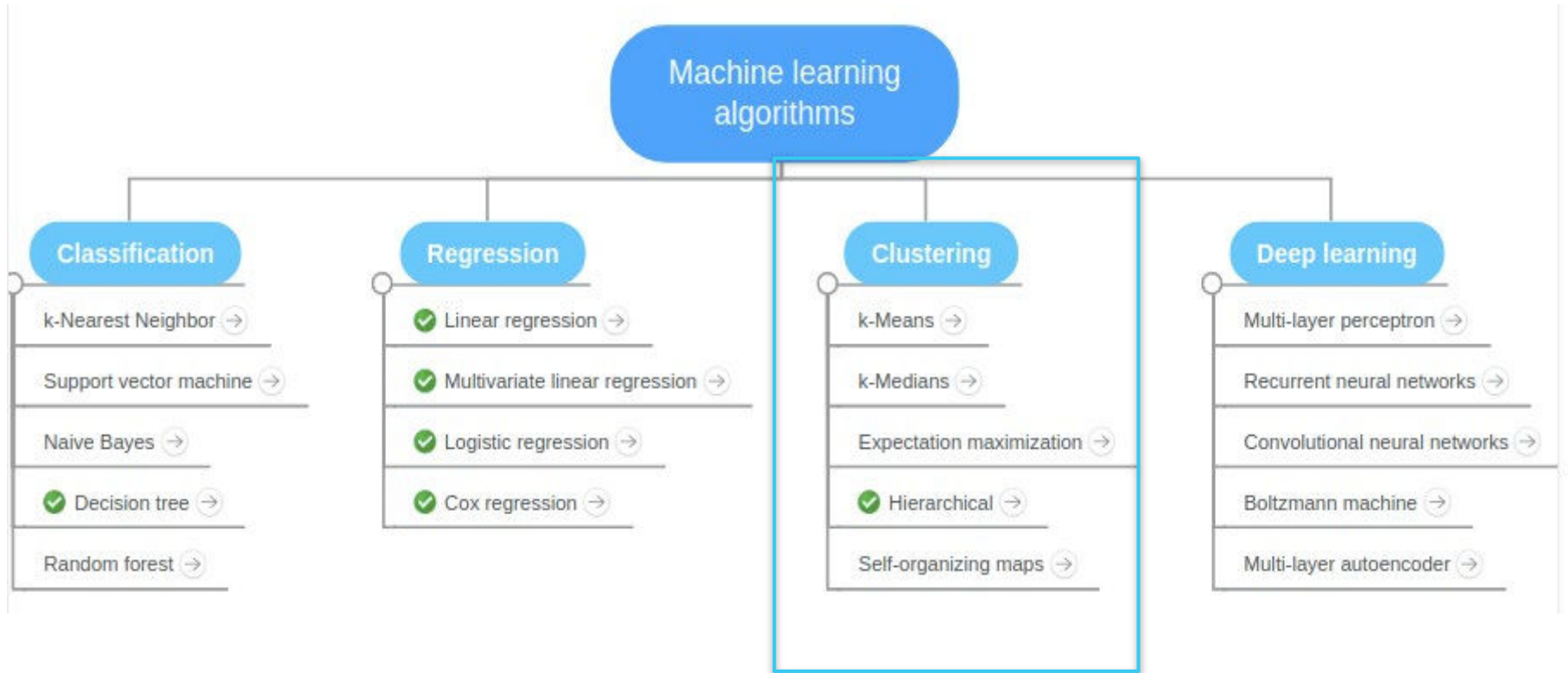
Raise the number of take offs by predicting take off times

Track n °4



Raise the number of landings by predicting the occupancy times of the runway

Les types d'algorithmes



Marketing ciblé dans le retail

Comprenez mieux les comportements d'achat de vos clients

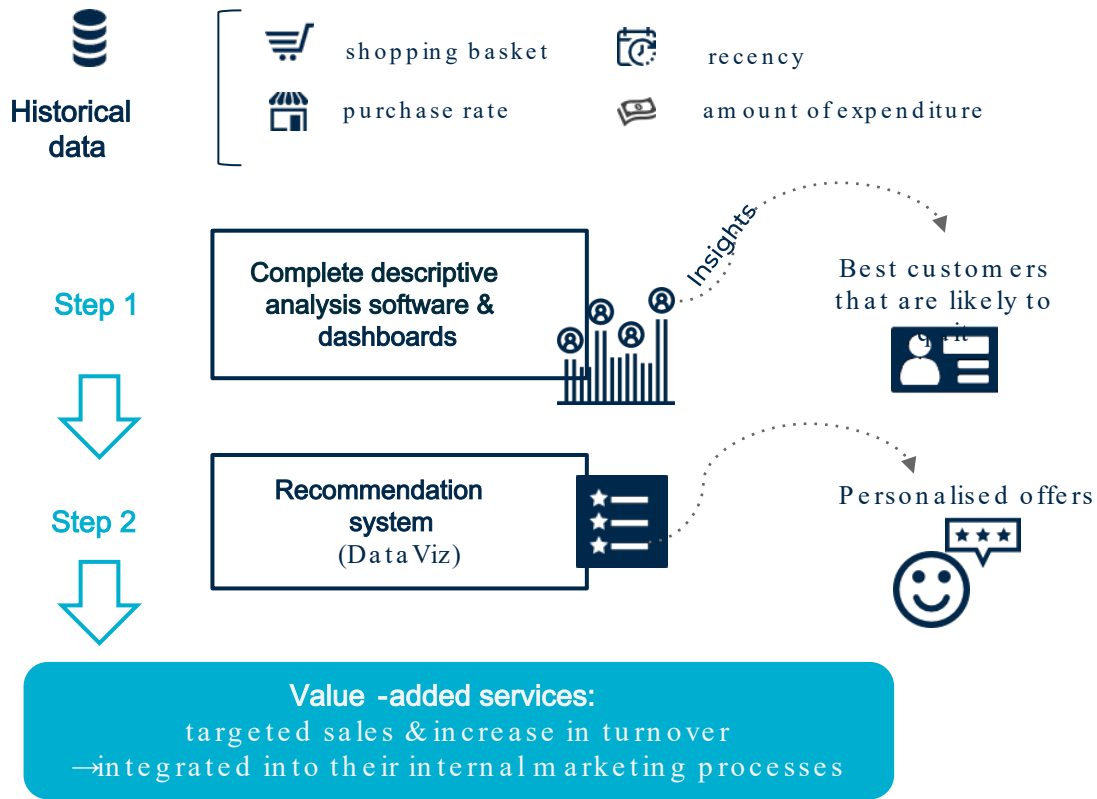
Our client: one of the largest retailers in Europe

Their question: how to enhance our customer segmentation and retain our best customer?

The result: an individual marketing automation software to maximise the profits generated by the “best customers”.

Challenges

- requires an accurate understanding of the specific business model
- upgrade statistical methods with machine learning techniques



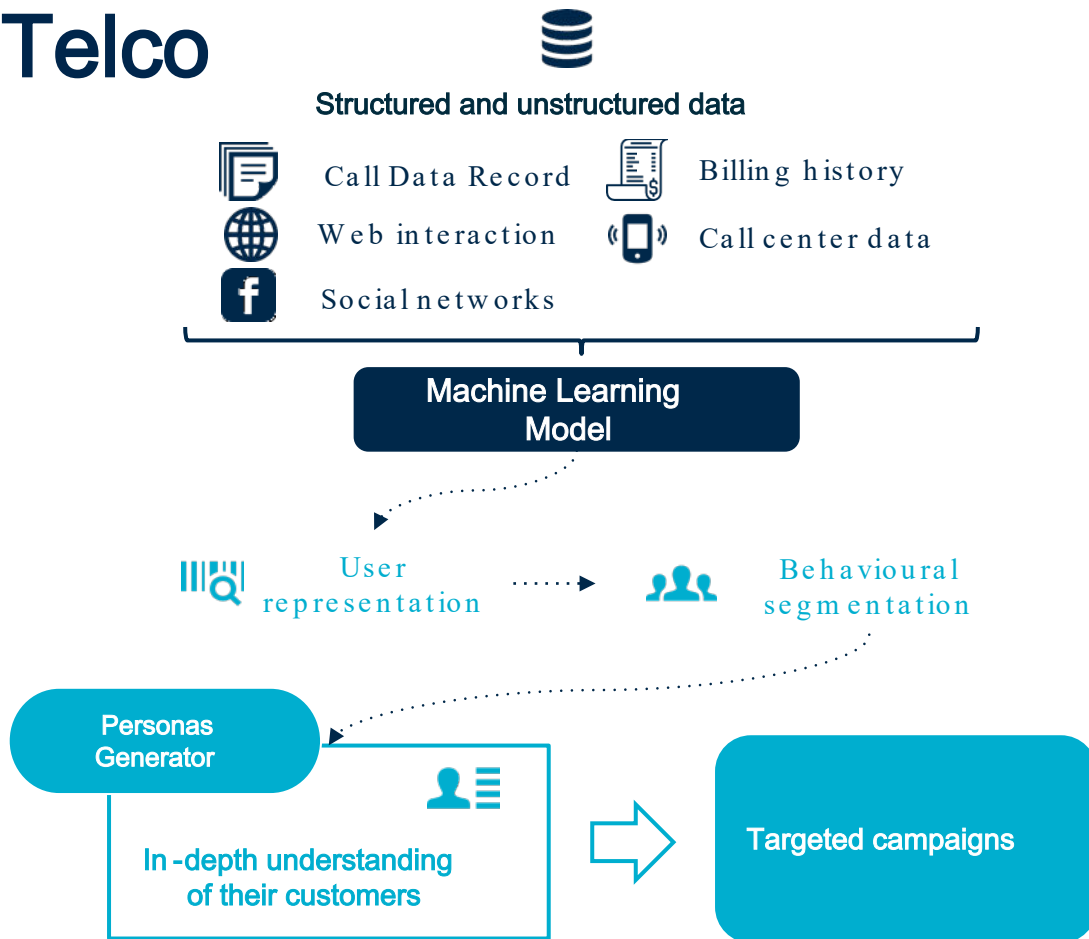
Personas pour le Telco

Our client: an international telco manufacturer

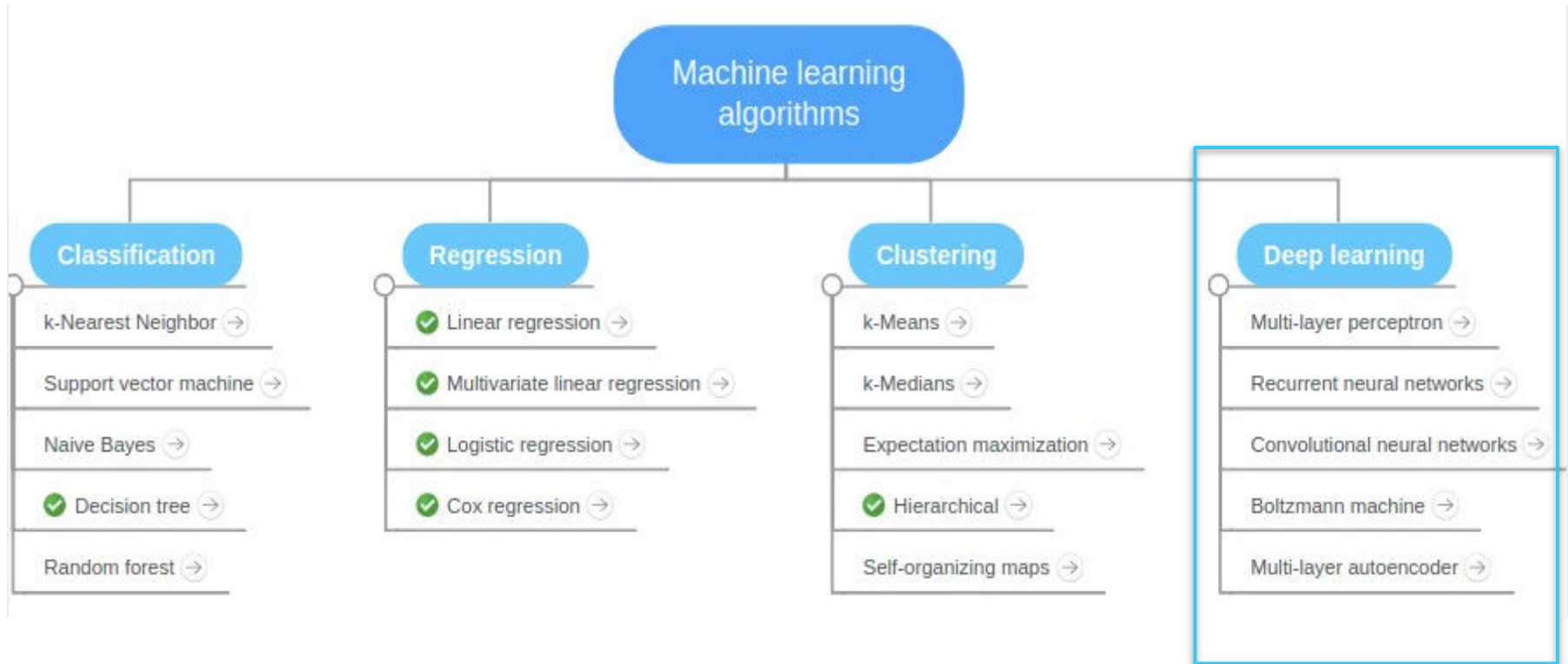
Their question: how can we get new customers' insights: human-centric, highly user-friendly, and empathetic

Our approach: combining heterogeneous data to design an algorithm predicting personas automatically

The result: a personas predicting model, generating segments' representative customers, describing their goals, lifestyles, activities, influencers, etc. to help answer questions like why are my customers churning? why does this subgroup use this product?



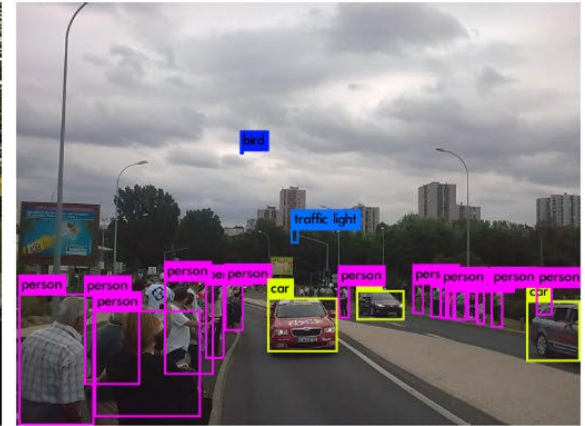
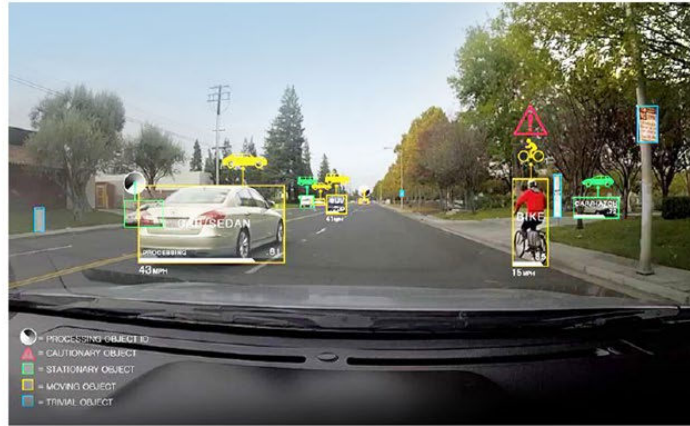
Les types d'algorithmes



L'IA pour l'image

Deux composantes

- Détection
- Reconnaissance



<https://www.nvidia.com/en-au/self-driving-cars/drive-px/>
<https://pjreddie.com/darknet/yolo/>

Reconnaissance de panneaux routier

La solution
développée



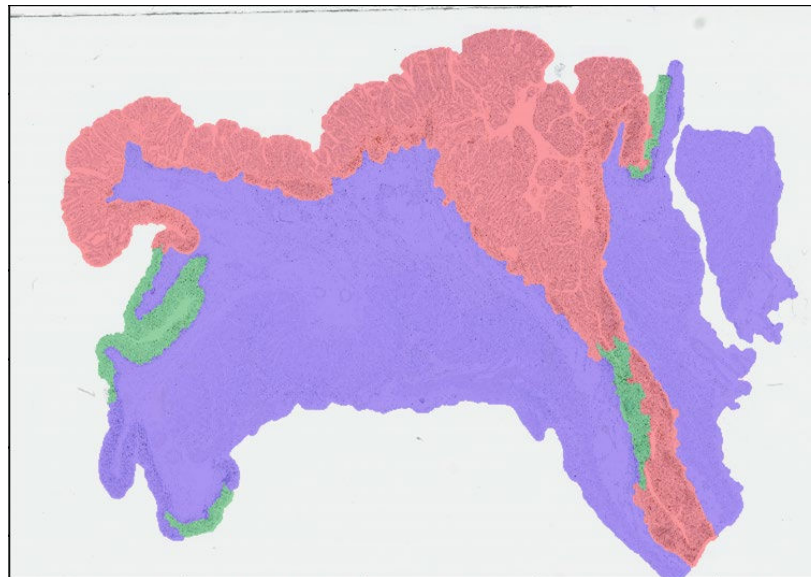
Segmentation d'image

pour les diagnostics en immuno-oncologie

Partner: a reference diagnostic company in immuno-oncology
Their goal: Perform the segmentation of tissues on histopathological data

The project: adapt state-of-the-art deep learning methods to perform the analysis of the image

Result : implementation and training of a deep learning pipeline. The model achieve a precision of 0.91. The client qualified the results as “**a major improvement**” over the previous solution

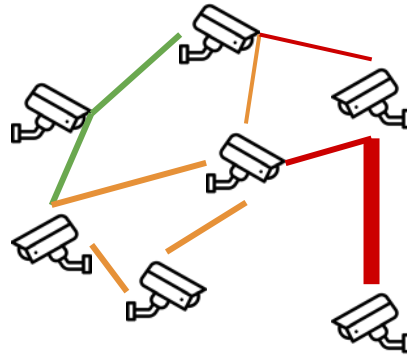


Une solution : pour GPMM

Identification des points de congestion récurrents et prédiction de l'état futur du trafic sur le port



Analyse locale du trafic
Qualification de la congestion



Analyse globale des flux et de la congestion

Découverte des relations entre les différentes zones du port

Evolution du trafic

- Définition d'une mesure du trafic (# de véhicules par minutes, densité du trafic, ..)
- Algorithmes de détection et caractérisation d'acteurs (piétons, VL, PL, trains, ...)
- Détection d'anomalies
- Réseaux de neurones pour définir l'état du trafic à un instant t

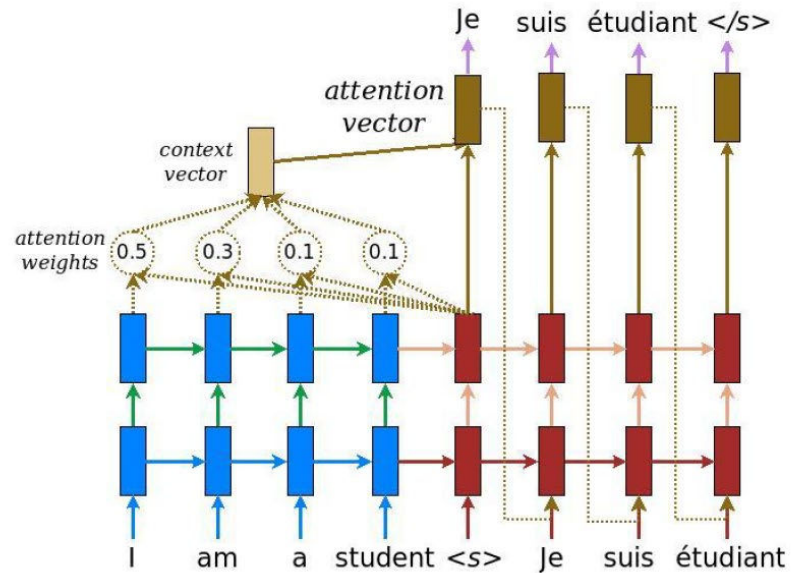
- Représentation du trafic portuaire sous la forme d'un graphe
- Pas d'information a priori sur les positions des 900 caméras
- Evolution dans le temps
 - prédiction à court terme du trafic et identification des points de congestion à venir
 - prédiction à long terme par l'identification de patterns récurrents



L'IA pour le texte

Apprendre à traduire...

Avec 36 millions de phrases

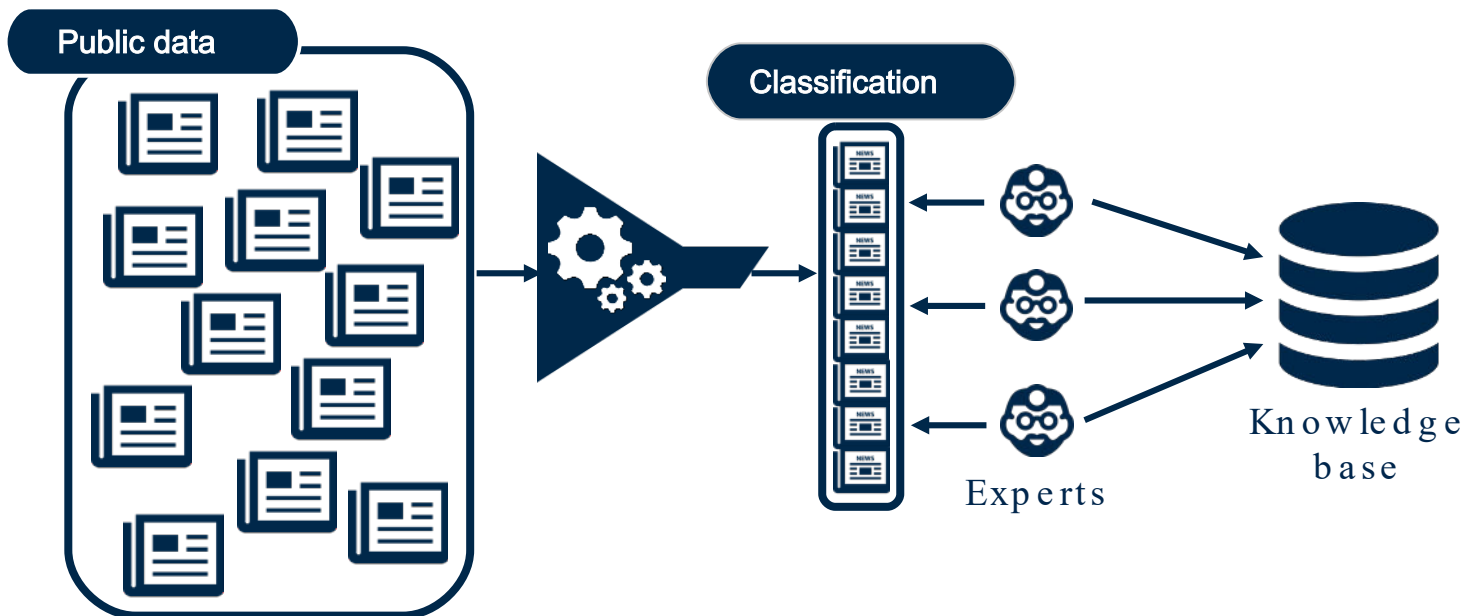


Extraction de connaissances d'un texte

pour la surveillance de la santé

Partner: CESP A - Centre d'Épidémiologie et de Santé Publique des Armées

The project: CIFRE PhD - Decision support in the French defence health monitoring system with NLP and ML.



Indexation des communications

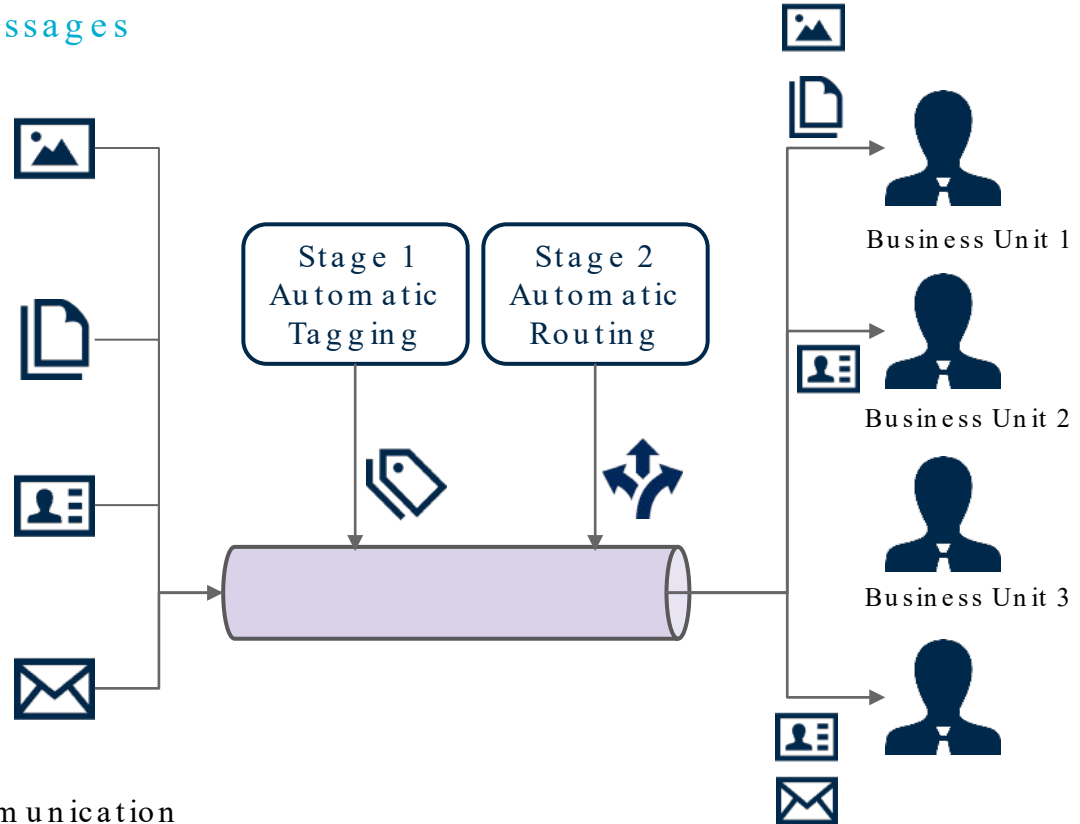
Améliorer l'acheminement de vos messages

Our client: an international insurance group

Their question: how to route the right message to the right department automatically?

The project: tap into machine learning and deep learning techniques (K-Means, DBSCAN, Word2Vec,...) to tag the message automatically

The result: a model able to route each incoming message towards the right department through automatic tagging by leveraging NLP, neural networks & clustering



Matcher les CV & descriptions de poste

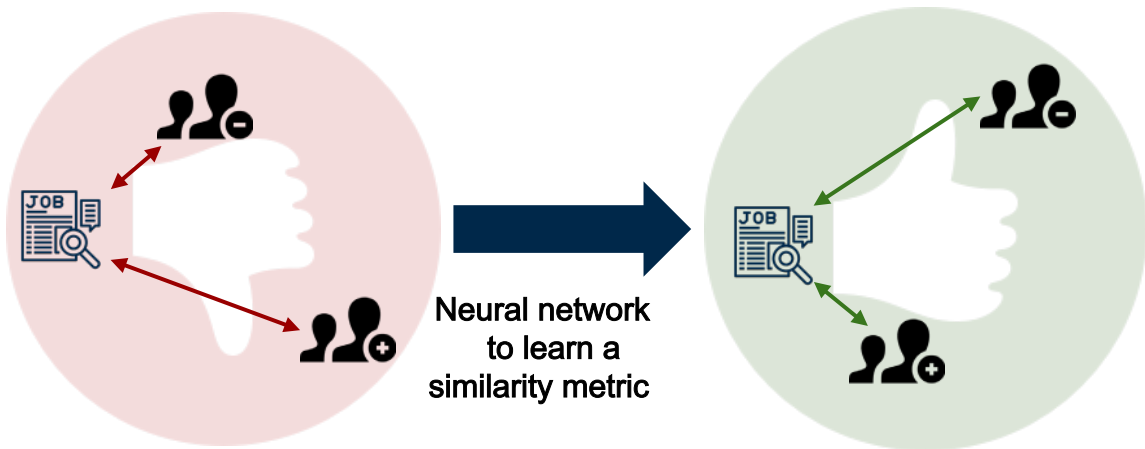
Apprendre une métrique de similarité à partir de différents types de données

Partner: HR company

Their goal: Quickly identify best profiles

The project: Adapt deep learning methods used in Face Recognition to Natural Language Processing

Result : A software able to select the best profiles among thousands of resumes. Headhunters qualified the results as “**huge time saver compare to usual HR tools**”.



Two types of data:



Job descriptions



CVs

Optimisation d'un google juridique

Pour fournir l'information dont les clients ont réellement besoin

Our client: a marketing company specialised in lawyers' communication. They designed a legal search engine

Their question: How to leverage AI to extract value from legal documents ?

Their challenges:

- scrap legal texts from websites
- train the ML model with limited data

The result:

- a bot that automatically scraps texts from the Internet
- a modelable to rank documents according to relevance using scores
- a modelable to draft legal summaries of documents

Step 1: Scrap documents



Step 2: Develop a search engine

NLP

ML

Ranking using frequency score of Elasticsearch



Search Engine

1. Document a
2. Document b
3. Document c
4. ...

Next steps

Integrate a model that generates summaries of legal texts automatically

Vous accompagner au travers d'un Diagnostic Flash IA

Répondre à vos enjeux pour

- **une meilleure offre mieux adaptée aux besoins de votre client**
 - **une offre différenciatrice de la concurrence**
 - **obtenir des avantages coûts et des gains de productivité**
-
- **Améliorer votre offre produits/services**

 - **Améliorer et/ou optimiser vos processus internes**
 - **Conception/Design**
 - **Maintenance**
 - **Production**
 - **Exploitation**
 - **Gestion et prévention des risques/sécurité**
 - **Economie d'énergie**
 - **Qualité**
 - **Réglementation**
 - **Ressources Humaines**
 - **Financiers**
 - **Marketing**
 - **Communication**
 - **Développement commercial**

- Un **programme d'accompagnement individualisé**
- Avec des **experts hautement qualifiés du domaine**
- **Inclus** dans le programme :
 - Une partie collective d'information & sensibilisation (**1/2 journée**)
 - Une partie individuelle (**3 jours en présentiel et rédactionnel hors présentiel**) avec :
 - Compréhension et audit de situation et **évaluation de la qualité des données sur un échantillon cible**
 - Identification des principales opportunités produits/services ou processus internes et des cas d'usages
 - Identification des technologies appropriées
 - Travail de recommandations avec plan d'actions et priorisation des actions Quick Wins
 - Suivi des actions mises en place
- **Livrables** :
 - Rapport synthétique de Diagnostic et plan d'actions
 - Rapport de recommandations

- Eligibilité :
 - Être une **TPE, PME** au sens européen du terme
 - Avoir au moins **un produit ou une solution disponible** susceptible d'être impacté par l'IA ou avoir des **données structurées en quantité suffisante** pour envisager le développement de nouveaux produits/services à base d'IA ou pour améliorer des processus internes par l'utilisation de l'IA
 - Être **membre d'un des Pôles partenaires**
 - Le siège social ou une partie significative des effectifs totaux doivent être **en PACA**
 - Être à jour de ses **obligations fiscales et légales**

- Faire une **demande** de participation au programme individuel **auprès de votre Pôle**

- **Coût** du programme :
 - Sensibilisation **gratuite**
 - Diagnostic individuel : **1 500€HT (1 800€TTC) au lieu de 3 000 €HT - 50 % aide ETAT**

Durée de réalisation : de 1 à 3 mois

Jade	SAEZ	Pôle Mer Méditerranée	saez@polemermediterranee.com
Victoria	DUPUIS	Optitec	victoria.dupuy@pole-optitec.com
Morgane	MILTGEN	Eurobiomed	morgane.miltgen@eurobiomed.org
Nicolas	ZAROOUR	Eurobiomed	nicolas.zarour@eurobiomed.org
Valentina	VOLOGNI	Capenergies	valentina.vologni@capenergies.fr
Magali	JAFFARD	Safe	magali.jaffard@safecluster.com
Patricia	GUIRAUDIE	Novachim	patricia.guiraudie@novachim.fr
Stéphane	MOUTARD	Novachim	Stephane.moutard@novachim.fr
Valérie	EL MERINI	Ea Ecoentreprises	valerie.elmerini@ea-ecoentreprises.com
Philippe	CLAUD	Innov'Alliance	philippe.claud@pole-innovalliance.com
Yvan	KEDAJ	Aqua Valley	y.kedaj@aqua-valley.com

MERCI!