

Comment tester l'Intelligence Artificielle ?

Sami MALLEK – Simon DEVARADJA



Les intervenants



Sami MALLEK
Expert Méthode et Outils
Responsable Automatisation



Simon DEVARADJA
Expert IA
Responsable Industrialisation



Tests Fonctionnels
(Cycle en V / Agile)



Anonymisation



Chatbot



Service d'exploitation
des images



Automatisation / RPA



Test de performance



Service de gestion
des emails



Service d'apprentissage



Audit de code



Outillage et GDE
Devops / Environnement



Service d'exploitation
des données textuelles



Service de robot
mécatronique



Comment tester l'Intelligence Artificielle ?

1.

Contexte

2.

Etude Outillage

3.

Expérimentation

4.

Prochaines étapes



« La conscience est-elle un algorithme ? »

Pourquoi ce sujet ?.....



Le nombre de projets en IA en forte augmentation. (50 en 2018 hors R&D)



Répondre à la problématique de qualité et garantir des outils fiables aux clients.



Industrialiser les tests dans un domaine probabiliste.

L'objectif est l'industrialisation de la pratique de test afin de fluidifier la mise en production des solutions dans le SI et suivre son comportement.

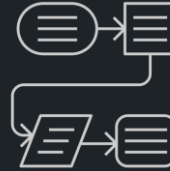


Sur un projet IA, peut-on appliquer ...



... la stratégie de test

... la méthodologie de test « classique » ?



... les outils de tests

Peut-on parler encore du rôle du testeur ? Sa place dans un projet IA ou est-il mené à disparaître ? Et à l'inverse quelle est la maturité d'un data scientist dans le domaine du test ?





Stratégie de test appliqué à l'IA

La couverture et les objectifs de tests, les risques projet, les environnements de test, l'organisation, les moyens, le planning,

Rajouter des types de tests adaptés (data, apprentissage...)

Définir un taux de confiance sur les résultats attendus.
Exemple : Atteindre + de 80%

Déterminer les risques éthiques

Identifier les moyens à mettre en place afin que la solution ne dérive pas en Build et MCO

L'aspect dynamique de la donnée apporte une nouvelle problématique dans les tests

Notre focus sur ...

1

Le « Code Testing » : Tester le code opérationnel

Le « Data Testing » : Tester la donnée de demain

2

3

Le « Model Testing » : Tester la validité du modèle IA

Le Test de feedback des utilisateurs

4

5

Le Test de monitoring : indicateurs de dérive



Le « Code Testing »: Tester le code opérationnel

Test unitaire

?

Test intégration

?

Test fonctionnel

?



Le data scientist

Si la donnée change demain, quels tests me permettraient de le remarquer ?!

Génie logiciel et Data science : deux visions à concilier



« Le Data Testing » : Tester la donnée de demain

On ne maîtrise pas la donnée de demain. Elle peut causer des situations de bugs, mais aussi l'obsolescence du modèle.

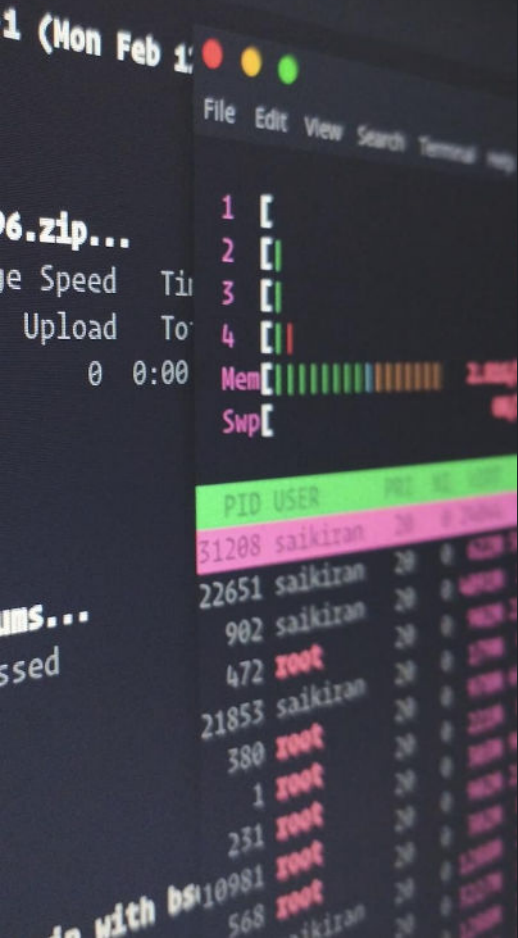
Test du format de la donnée

Vérifier que le format de la nouvelle donnée respecte celui utilisée lors du *build*.

ex: vérifier le bon encodage d'un texte, la taille d'une image

Test de représentativité

Vérifier que la nouvelle donnée est statistiquement bien représentée par la donnée d'apprentissage



Le « Model Testing » : tester la validité du modèle IA



Tester si le modèle a sur-appris ou non (cross-validation...)



Tests de performance de l'algorithme et du programme : comparer le nouveau score du modèle au critère métier



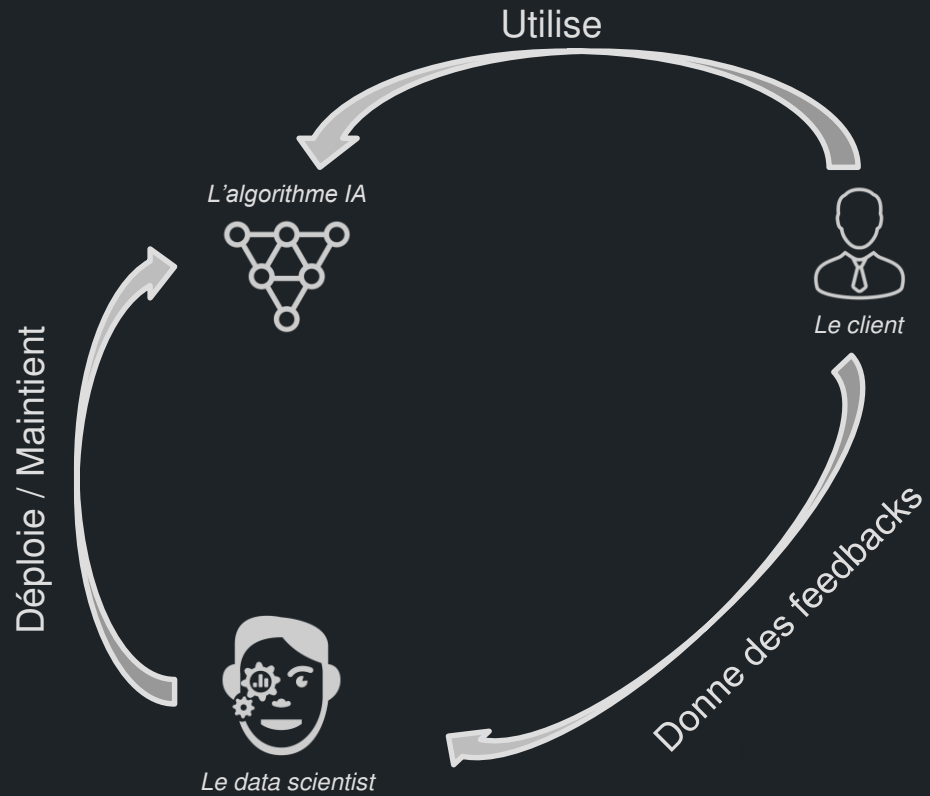
Tests de non régression : vérifier la sortie de cas simples

L'automatisation de ces tests permettent le déploiement continu d'un modèle de qualité





Test de feedback



Automatiser la récolte de feedback pour garder la pertinence du modèle.

Test de monitoring

Consommation
CPU / RAM

Temps de
réponse moyen
de l'API

Temps de mise
à jour du
modèle

Informations
stockage

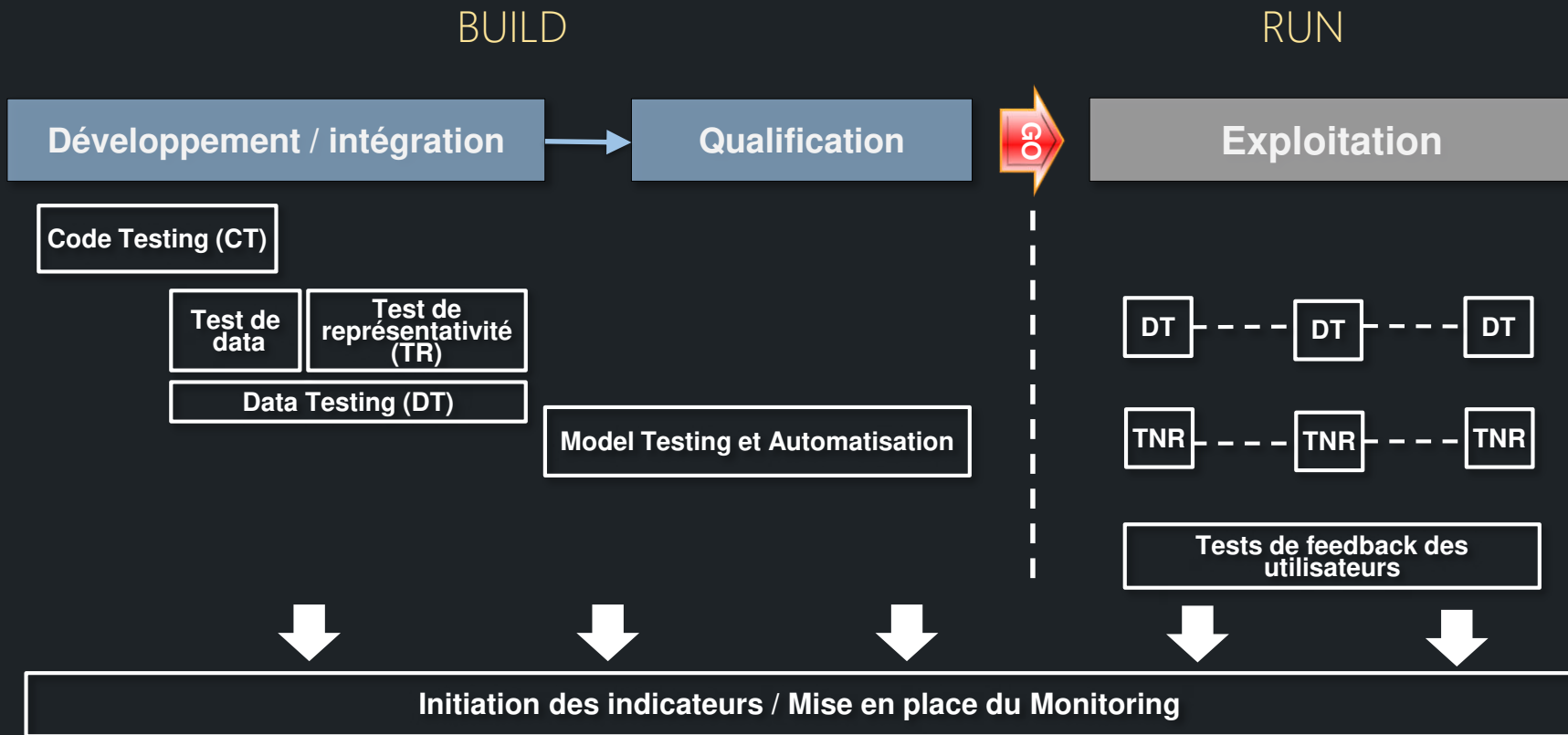
Nombre de
requêtes vers
l'API

Temps de
récupération de
la donnée

La surveillance de chacune de ces métriques permettra de noter un comportement anormal de la solution



Positionnement des **types de tests** dans un projet IA



Positionnement du « **Testeur** » dans un projet IA

Projet composé au minimum de 3 développeurs/data scientists



Ce sont les data scientists qui font les tests.

Projet composé de plus de 3 développeurs/data scientists



Un testeur est nécessaire tout le long du projet.



Evolution des compétences



Connaissance et compréhension :

- Langages de type R, Python, C++, Java
- Machine Learning / Clustering / Algorithmes
- Gestion et analyse des données structurées ou non structurées



IA



Architect



Métier



Concepteur



Dev



Equipe Qualité



Automaticien



**Comment tester l'Intelligence Artificielle ?
Etude Outillage**

A quoi doit répondre les outils ?

- Faciliter et améliorer la communication entre les data scientists et les testeurs
- Industrialiser la partie data testing afin de récupérer les métriques nécessaire en MCO.
- Tester l'IA en « production » ou plutôt comment garantir que le modèle ne dérive pas.



Synthèse de l'étude



EXPLORATION



BUILD



QUALIFICATION



RUN



**Comment tester l'Intelligence Artificielle ?
Expérimentation**

Présentation de Gecko

Moteur de recommandation de connaissance sur les activités d'une centrale nucléaire : Comparaison d'un rapport de l'agent de maintenance avec la base des rapports afin d'identifier les activités identiques à la maintenance décrite dans le rapport.



Type de problématique : Non Supervisé







Algorithmes : Word2Vec (Embedding)
WMD (Distance)




Technologies : Python, Gensim, NLTK



REX en chiffres.....

	<i>Le problème</i>	<i>La solution</i>	<i>Le gain</i>
	<p>Pas de tests unitaires</p> <p>Pas d'arborescence</p> <p>Pas d'audit de code</p>	<ul style="list-style-type: none">- Unittest, logging : chaque brique est testée en production.- Normalisation du repository Git- Utilisation de SonarCube	<p>Gain en temps : 35 jours de data scientist en moins sur une période de 6 mois.</p>
	<p>Pas de data testing</p>	<p>Avec unittest : la donnée est testée en temps réel</p>	<p>Gouvernance optimisée + gain en temps</p>
	<p>Récupération de 4 à 5 feedback par semaine</p>	<p>MEP d'un système de feedback</p>	<p>3 fois plus de feedback par jour : une amélioration continue du modèle</p>
	<p>Infrastructure mal dimensionnée</p>	<p>Test de monitoring</p>	<p>Anticipation des problèmes d'infrastructure : gain financier et en temps</p>





**Comment tester l'Intelligence Artificielle ?
Prochaines étapes...**

Système de feedback automatisé

1^{er} étape : Mise en place de feedback intégré dans l'interface de restitution des résultats de l'algorithme IA

■	10/01/2019 18:04:21	CHASTH	G01	AL01	Ecart < entre Température air chaud alternateur et Température air froid alternateur 2	unknown	5
---	---------------------	--------	-----	------	--	---------	---

il y a 22 alertes

Alerte

Date de début : **11/01/2019 - 01:05:44**

Code usine : **MTEZIH**

Predicted_T4 : **T4DR**

EGF : **G02**

SEF : **TU02**

Traitement des alertes

Priorisation : **1** Libellé : **Dépassement seuil > Débit fuite joint arbre**

Bien prédit : oui non T4 réel :

Commentaire:

■ Traité

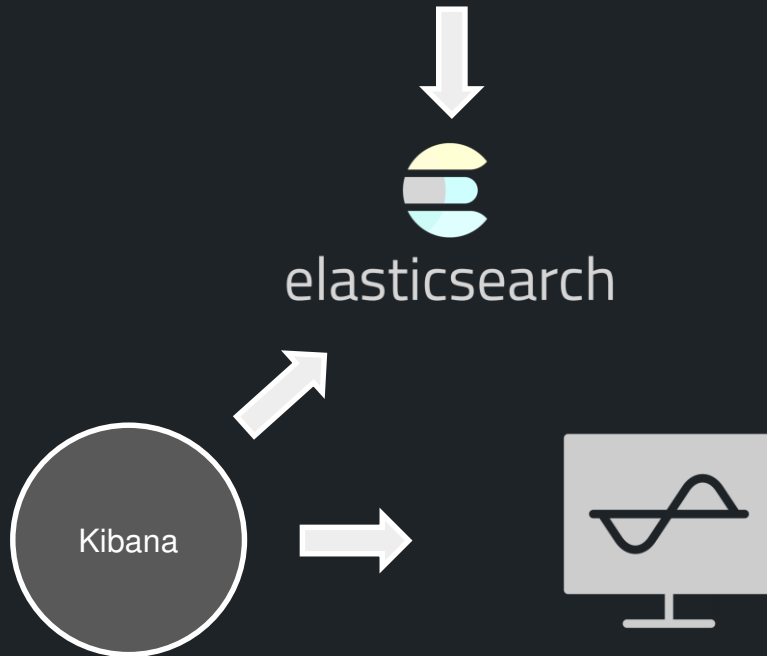
2^{ème} Etape à prévoir: Mettre en place une automatisation de la récolte de feedbacks (via un système de mail avec demande d'annotations)



Feedback

Un monitoring

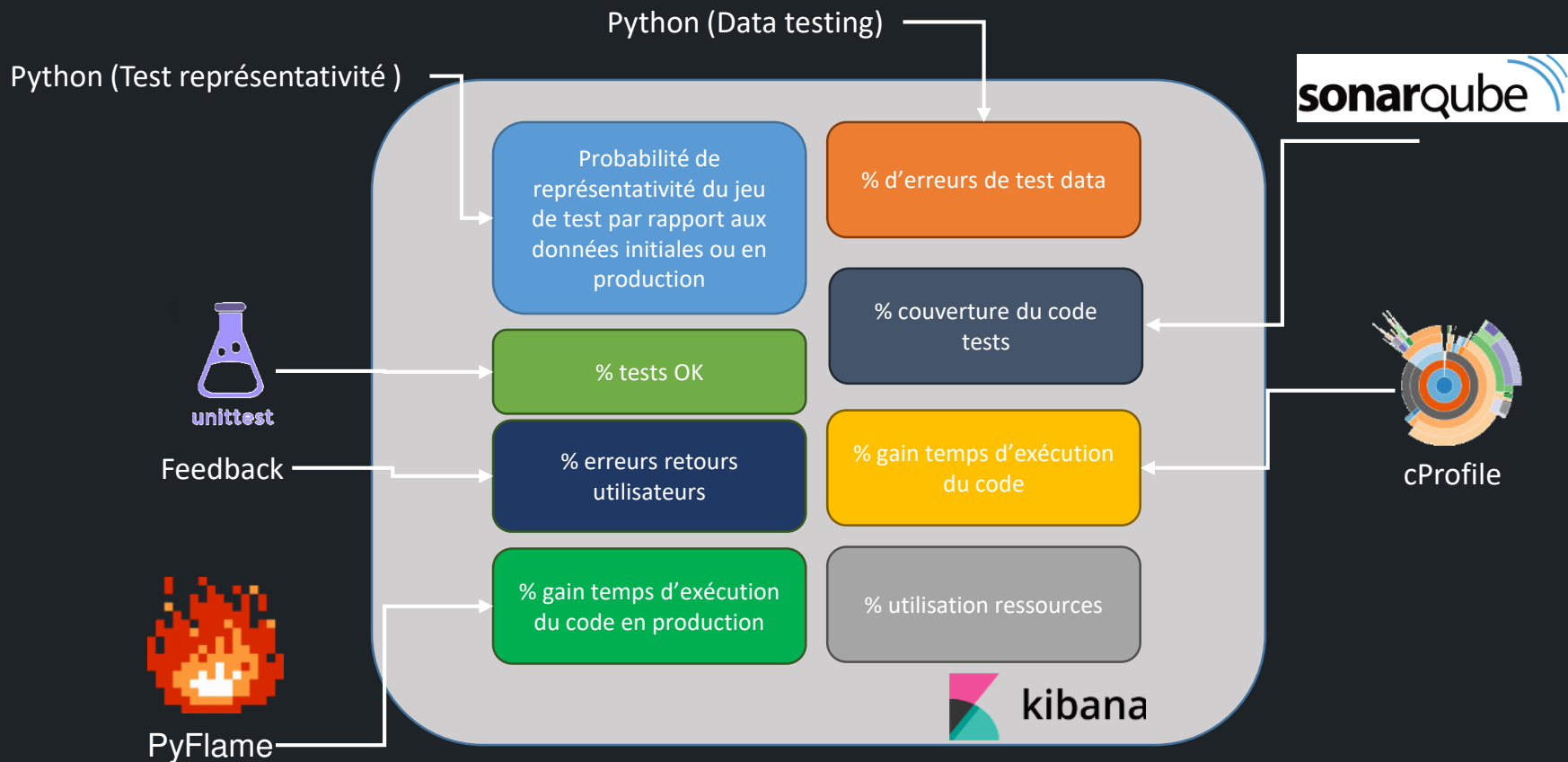
Normaliser l'obtention de toutes les métriques possibles pour surveiller une IA



Ceci permettrait de personnaliser une interface de monitoring pour toute solution IA



Maquette IHM du Monitoring





Le monitoring de l'IA



Avez-vous encore des doutes.....?



Comment tester l'Intelligence Artificielle ?

Des questions ?