La vrai révolution IA est celle de la donnée !

Hamilton DE ARAUJO
Délégué au numérique et à l'intelligence artificielle
Direction de la Transformation, Data & IA
Collectivité européenne d'Alsace (CeA)

Journée de la donnée en Alsace

13 octobre 2026

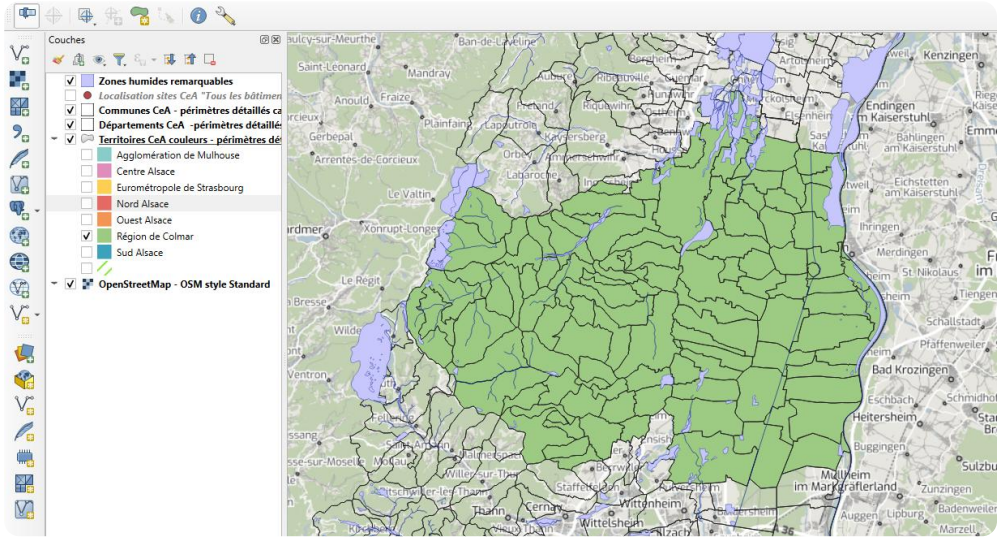


Introduction

15 interventions depuis ce matin, enfin on change de sujet. **Eh bien, non !**



	A	B	C	D	E	F	G	H	I	J	K	L
1	cle_interop	commune_in	commune_n	commune_d	commune_d	voie_nom	lieudit_comp	numero	suffixe	position	x	y
2	68015_0032_00004	68015				Allée Arthur Rimbaud		4		bâtiment	1028031.95	6753843.74
3	68015_0032_00010	68015	Baldersheim			Allée Arthur Rimbaud		10		bâtiment	1028012.41	6753966.32
4	68015_0032_00017_b	68015	Baldersheim			Allée Arthur Rimbaud		17	b	bâtiment	1028052.53	6753895.76
5	68015_0032_00023	68015	Baldersheim			Allée Arthur Rimbaud		23		bâtiment	1028044.77	6753942.34
6	68015_0032_00029	68015	Baldersheim			Allée Arthur Rimbaud		29		bat	1028036.11	6753986.82
7	68015_0032_00031	68015	Baldersheim			Allée Arthur Rimbaud		31		bat	1028032.8	6754022.61
8	68015_0533_00006	68015	Baldersheim			Allée du Quatelbach		6		bat		6753053.72
9	68015_0533_00010	68015	Baldersheim			Allée du Quatelbach		10		bat		6753063.91
10	68015_0533_00011	68015	Baldersheim			Allée du Quatelbach		11		bat		6753030.69
11	68015_0533_00012	68015	Baldersheim			Allée du Quatelbach		12		bat		6753064.78
12	68015_0533_00017	68015	Baldersheim			Allée du Quatelbach		17		bat	1027499.27	6753074.48
13	68015_0533_00019	68015	Baldersheim			Allée du Quatelbach		19		bâtiment	1027513.88	6753092.46
14	68043_0860_00049_b	68043	Bollwiller			Rue du Vieil Armand		49	b	bâtiment	1018125.46	6759266.47
15	68043_0860_00055	68043	Bollwiller							bâtiment	1018084.19	6759229.66
16	68043_0860_00057	68043	Bollwiller							bâtiment	1018021.86	6759264.7
17	68055_0069_00001	68055	Bruebach							bâtiment	1027021.11	6741643.06
18	68055_0023_00002	68055	Bruebach							bâtiment	1026711.88	6742568.25
19	68055_0017_00001	68055	Bruebach							bâtiment	1027014.94	6742332.53
20	68055_0017_00006	68055	Bruebach					6		bâtiment	1027024.75	6742361.89
21	68055_0017_00010	68055	Bruebach			Rue Basse		10		bâtiment	1027024.98	6742415.36
22	68055_0017_00011	68055	Bruebach			Rue Basse		11		bâtiment	1027005.48	6742438.51
23	68055_0017_00012	68055	Bruebach			Rue Basse		12		bâtiment	1027023.92	6742437.73



Ce que l'IA sait faire

l'IA « classique », pas l'IA générative, celle qui produit aujourd'hui l'essentiel de la valeur en collectivité, tient en cinq verbes :

01 CLASSER

À quelle catégorie appartient cet élément ?

Orienter une demande vers le bon service

02 PRÉDIRE

Que risque-t-il de se passer ?

Anticiper une sollicitation ou une défaillance

03 ESTIMER

Quelle valeur peut-on approcher ?

Estimer un coût, un délai ou une fréquentation

04 DÉTECTER

Y a-t-il quelque chose d'inhabituel ?

Repérer une anomalie ou une rupture

05 REGROUPER

Quels éléments se ressemblent ?

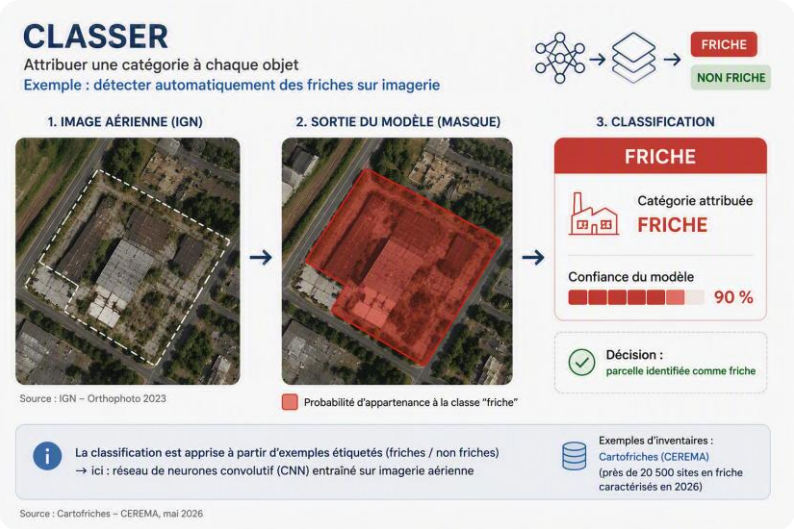
Identifier des profils similaires

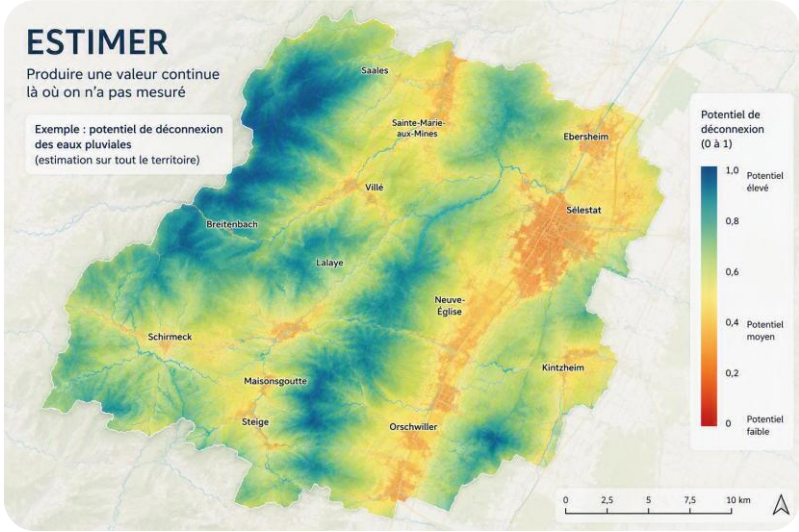
TECHNIQUES À MOBILISER SELON LE BESOIN

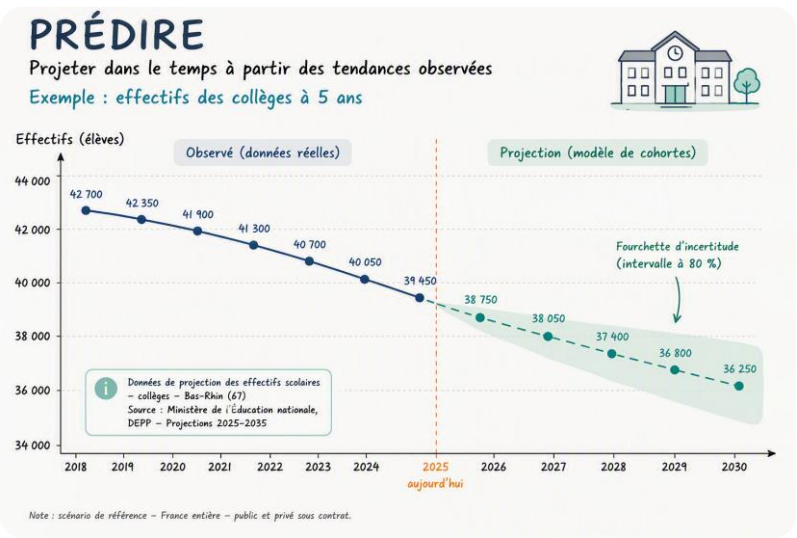
- Classification
- Régression
- Séries temporelles
- Détection d'anomalies
- Clustering
- NLP
- Vision

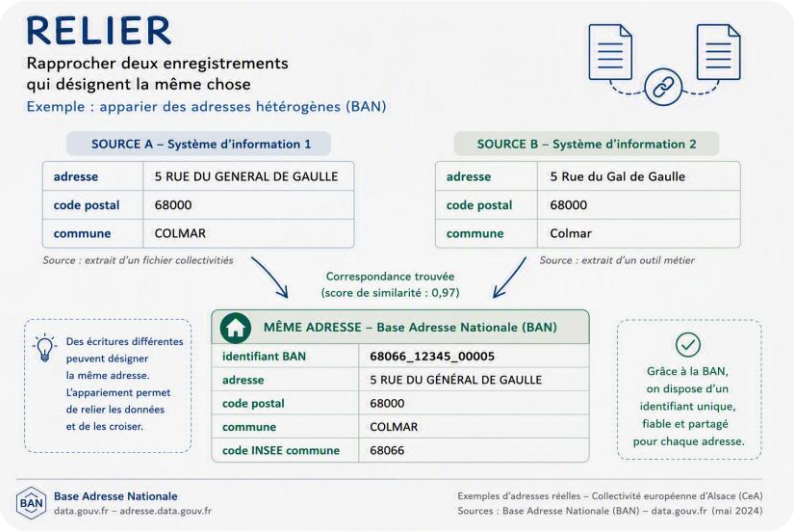
Le choix de la technique vient après la formulation du besoin et l'examen des données disponibles

Ce que l'IA sait faire









Une précision qui va peut-être vous surprendre : sur des données tabulaires — c'est-à-dire l'immense majorité de vos données métier — les méthodes les plus performantes ne sont pas les réseaux de neurones

Ce sont des méthodes à base d'arbres de décision et/ou de régression, qui ont trente ans, qui tournent sur un ordinateur portable et qui restent aujourd'hui suivant devant le Deep Learning.

Exemple 1

1 La question métier

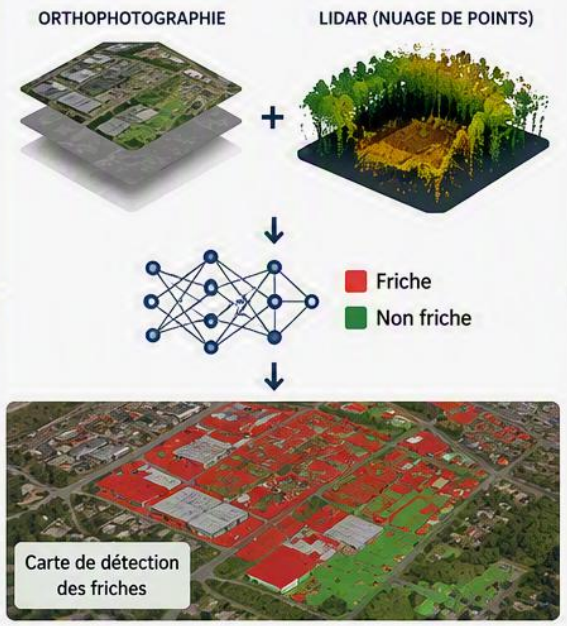
Où sont les friches du territoire ?



Objectif métier
Identifier automatiquement les friches pour orienter les politiques publiques de réhabilitation et de sobriété foncière.

2 La technique

Classification supervisée — ortho + LIDAR



Méthode
Le modèle apprend à partir d'exemples étiquetés (friches / non friches) pour prédire la catégorie de chaque parcelle.

3 La donnée qui a tout conditionné

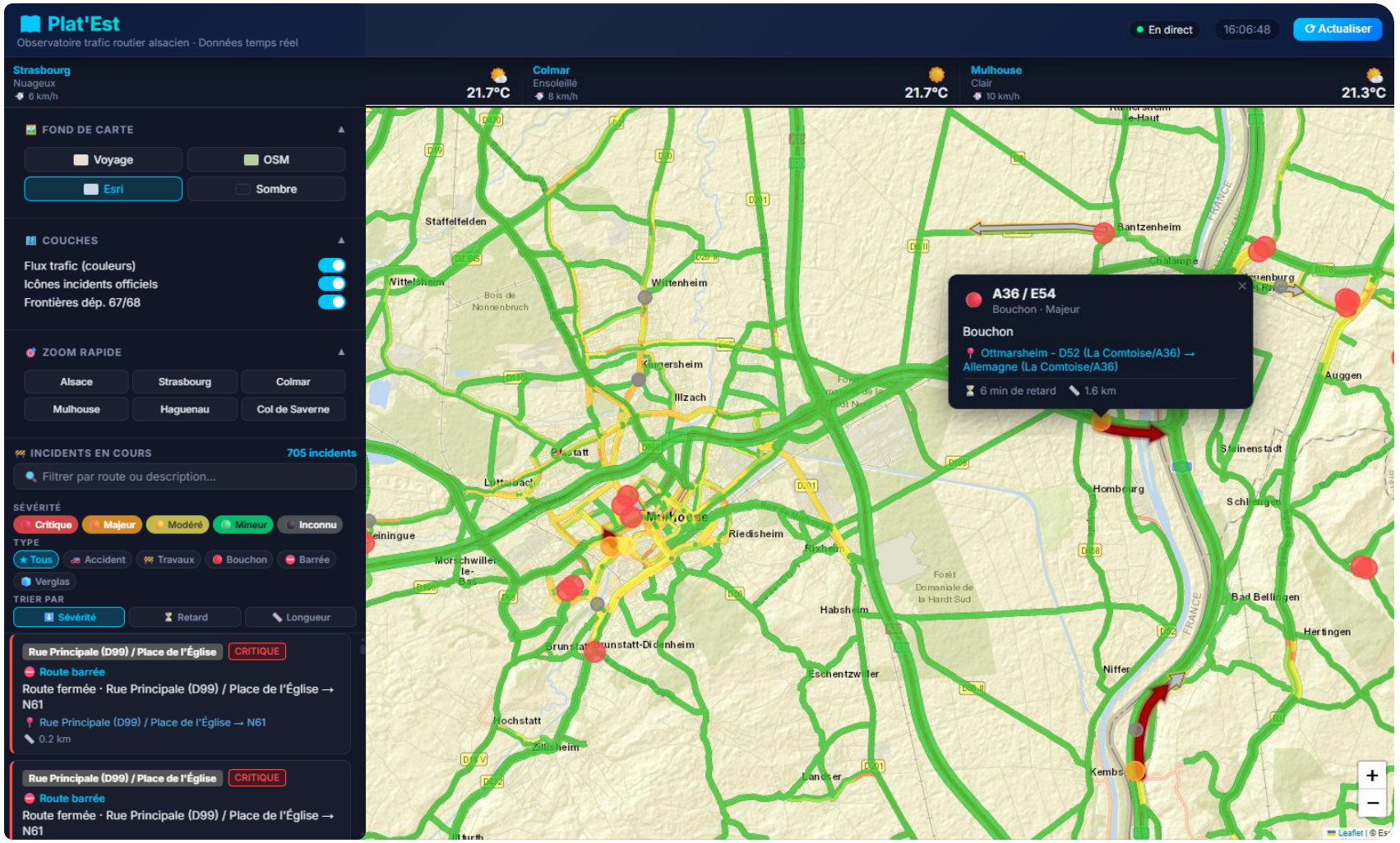
Il faut des exemples étiquetés. Beaucoup.



Donnée clé
Le modèle ne peut apprendre que si l'on dispose de nombreux exemples fiables, validés sur le terrain.

Sans exemples étiquetés en quantité suffisante : pas de modèle fiable, pas de résultats exploitables.

Exemple 2

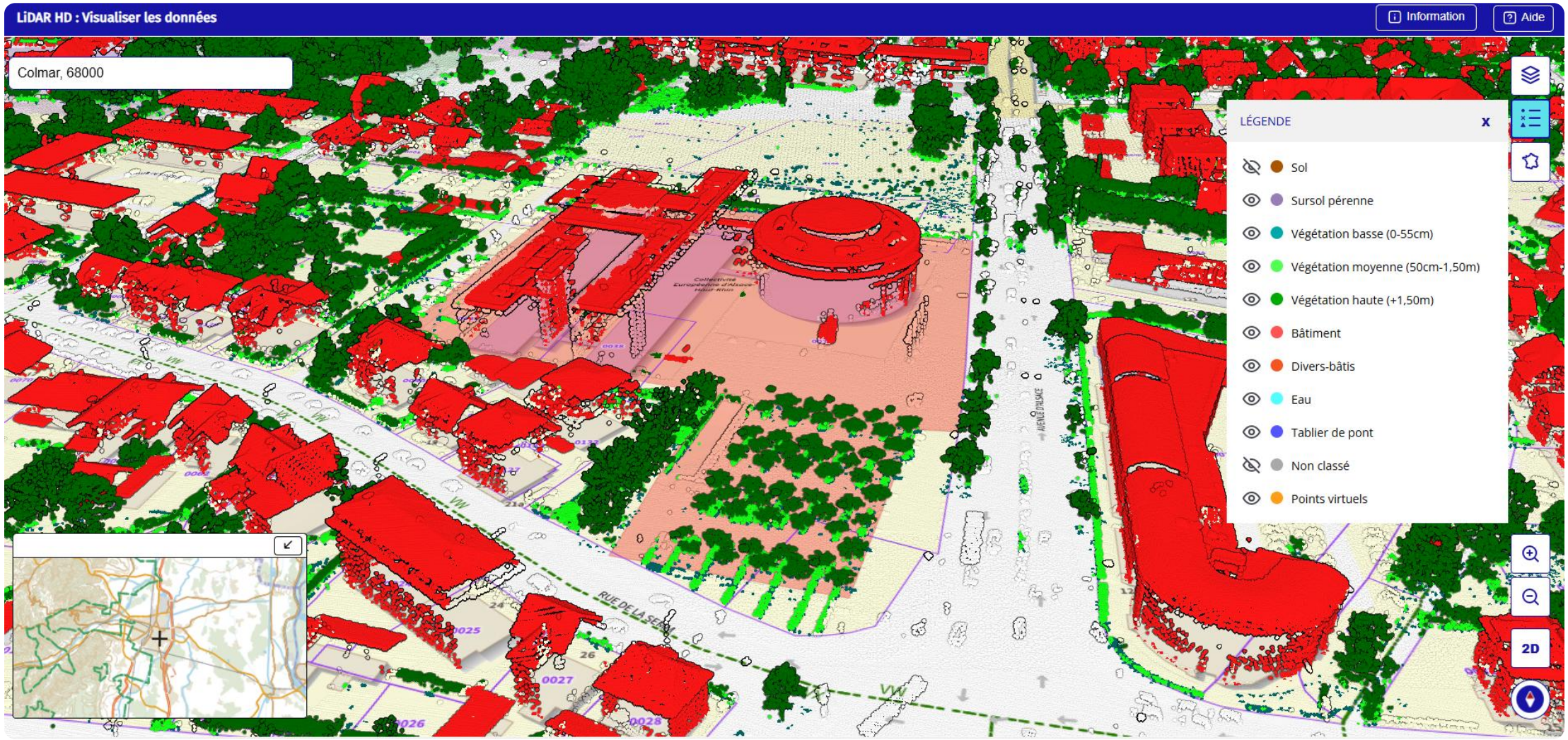


Un outil de suivi et de gestion du réseau routier ?

Ce qui bloque, ce n'est pas la modélisation.

Par exemple, dix ans d'historique d'interventions ont été saisis en texte libre, sans nomenclature partagée. « Verglas », « chaussée glissante », « salage préventif », « RAS » — un humain comprend. Un modèle, non.

Exemple 3


LiDAR HD : Visualiser les données
Information
Aide
Colmar, 68000
LÉGENDE
Sol
Sursol pérenne
Végétation basse (0-55cm)
Végétation moyenne (50cm-1,50m)
Végétation haute (+1,50m)
Bâtiment
Divers-bâtis
Eau
Tablier de pont
Non classé
Points virtuels
2D

Dans aucun de ces trois projets, l'algorithme n'a été le facteur limitant.

Un projet IA vaut ce que vaut sa donnée

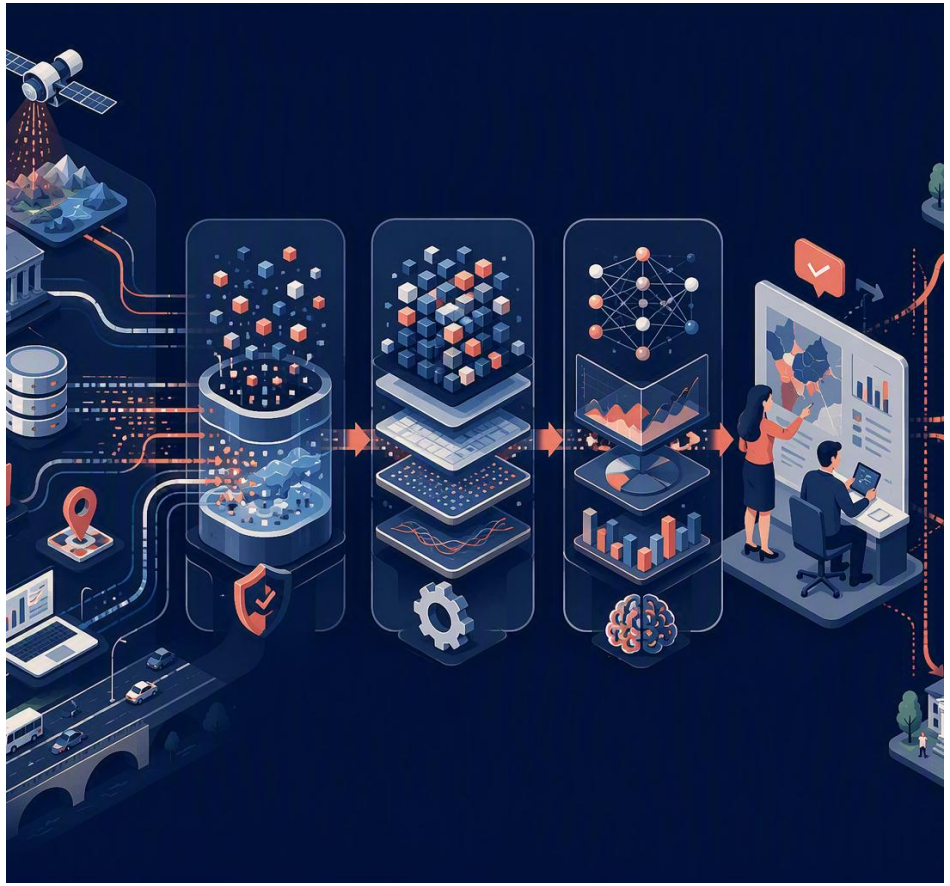
Le modèle n’est qu’un maillon du projet IA

Le travail décisif se déroule avant et après l’entraînement.

01 Besoin métier	06 Contrôler la qualité
02 Localiser les données	**07 Construire le modèle**
03 Obtenir les accès	08 Intégrer au processus
04 Comprendre et documenter	09 Suivre dans le temps
05 Nettoyer et rapprocher	

Le modèle : 1 étape sur 9

→ La valeur vient de toute la chaîne



Actu: la course de la donnée

Le Monde

ÉCONOMIE • PIXELS



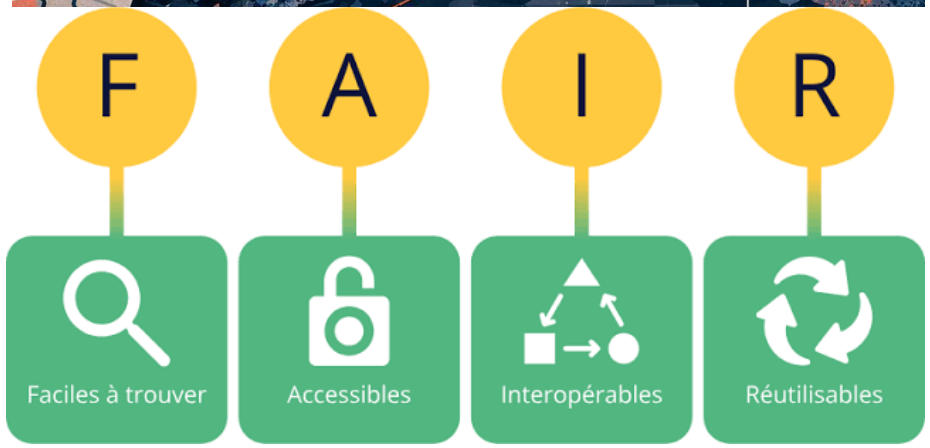
SpaceX rachète Cursor, éditeur de logiciel de code assisté par IA valorisé à 60 milliards de dollars

Le rachat annoncé de la société survient alors qu'une course est engagée dans la Silicon Valley pour capter les développeurs professionnels, devenus le principal moteur de croissance de l'industrie de l'IA.

CURSOR Grok

Donnée disponible ≠ donnée utilisable



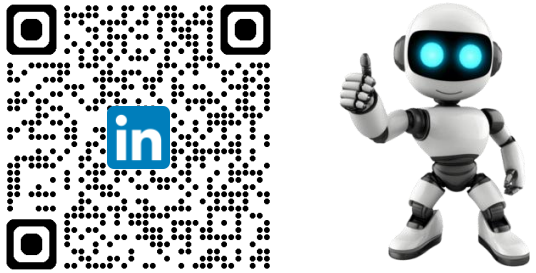


01	**ACCESSIBLE**	Trouvable, disponible et utilisable
02	**STRUCTURÉE**	Formats, champs, unités et définitions connus
03	**PARTAGEABLE**	Réutilisable et interopérable
04	**FIABLE**	Complète, cohérente, actuelle, représentative
05	**GOUVERNÉE**	Responsable, droits, sécurité et règles claires

Le format et le mode de partage ne sont pas des détails techniques : ils déterminent la capacité à croiser, automatiser et maintenir.

Merci pour votre attention !

Avez-vous des questions ?



hamilton.de-araujo@alsace.eu
Délégué au numérique et à l'intelligence artificielle
Direction de la Transformation, Data & IA
Collectivité européenne d'Alsace (CeA)

