



5^e SYMPOSIUM
SUR LE TRANSFERT
DE CONNAISSANCES
EN ÉDUCATION

S'ADAPTER POUR MIEUX RÉUSSIR

L'intelligence artificielle au service de la persévérance scolaire

Une présentation de : Eric Racine, directeur général du CSS du Val-des-Cerfs

19 avril 2023

Un événement organisé par :



Et financé par :



La genèse de l'initiative

Centre
de services scolaire
du Val-des-Cerfs

Québec 

Objectifs initiaux



Intervenir le plus tôt possible

La proactivité est **cruciale**. Les signes avant-coureurs du décrochage doivent être décelés le plus **tôt** possible dans le cheminement scolaire.

Intervenir le plus précisément

Garantir une allocation **efficace** et **pertinente** des ressources du centre de service implique une détection **efficace** des étudiants susceptibles de décrocher.



Intervenir en optimisant les ressources disponibles

Afin d'être en mesure d'utiliser d'une manière **efficace** les modèles, il est crucial d'en tirer un maximum de connaissances et d'informations lors de la prise de **décision**.

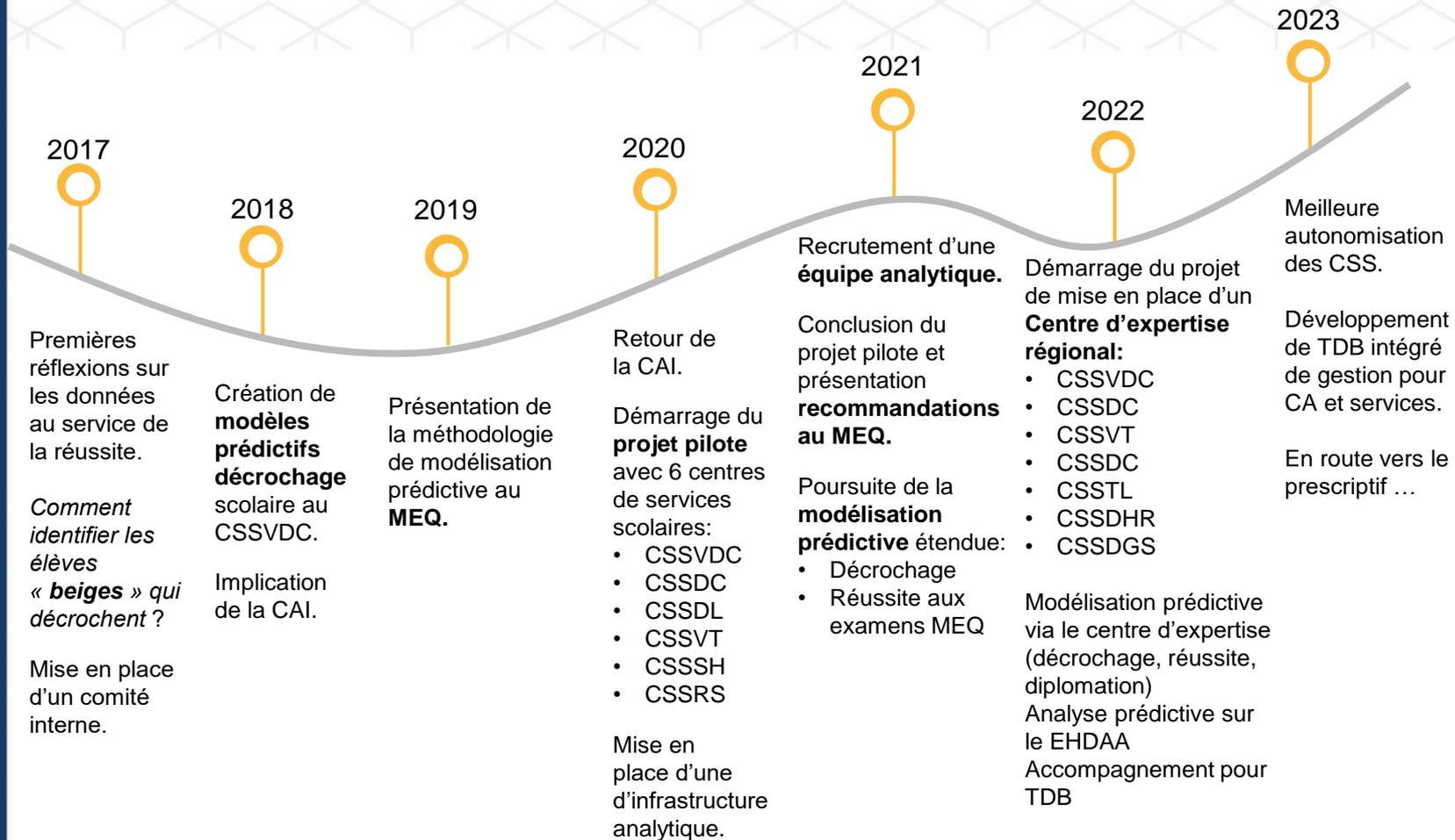
Produire des mesures quantitatives basées sur les données

Ceci dans le but de maximiser et **valoriser** les **données** recueillies à travers l'ensemble des systèmes et le temps en utilisant des techniques de pointe **en intelligence artificielle** et en **analytique avancée**.



« Comment détecter les différents profils de décrocheurs discrets de façon quantitative et systématique ? »

Évolution du projet



Centre
de services scolaire
du Val-des-Cerfs

Québec



Projet pilote persévérance scolaire et intelligence artificielle

Centre
de services scolaire
du Val-des-Cerfs

Québec 

Sommaire de la démarche

Six phases composent la démarche du projet

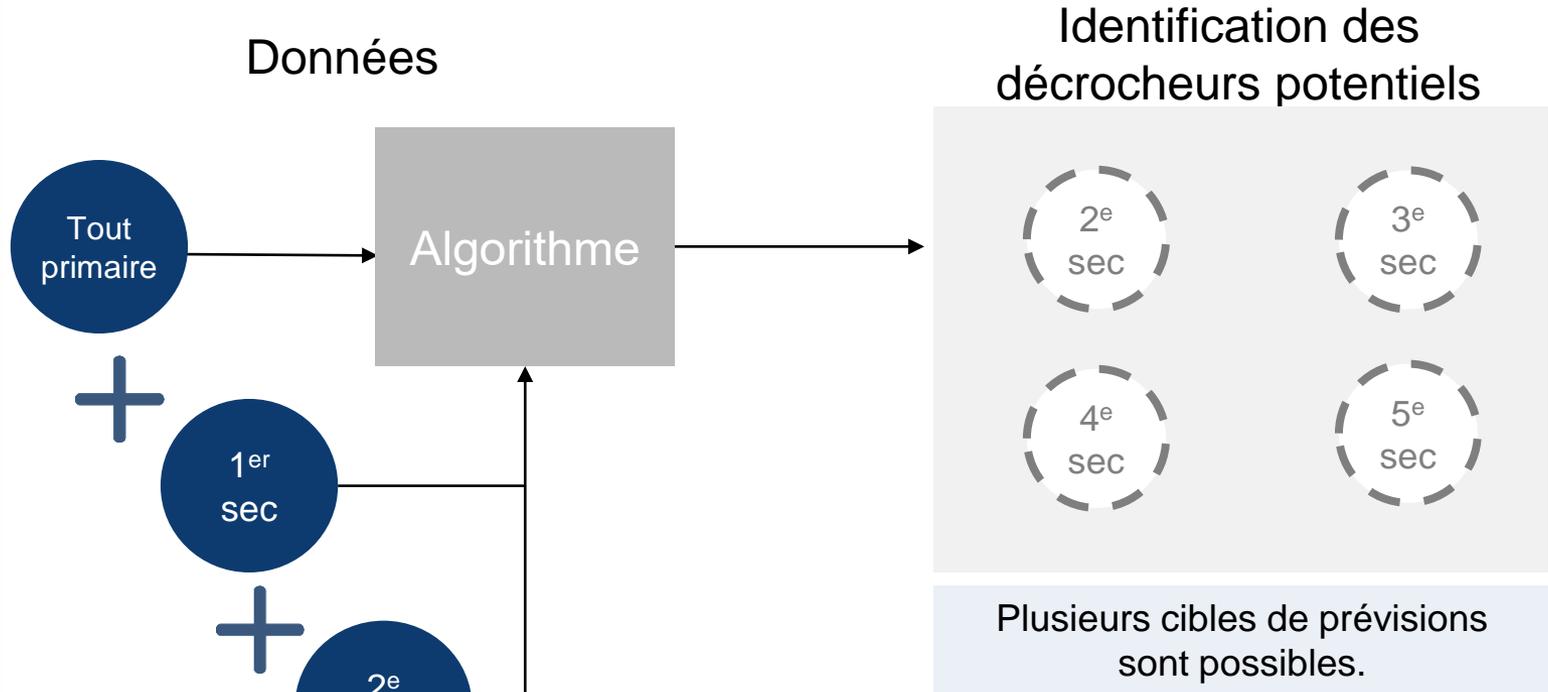


Certaines phases se sont réalisées en parallèle pour optimiser la démarche, notamment les phases 5 et 6 qui ont impliqué seulement le CSSVDC.

Centre
de services scolaire
du Val-des-Cerfs

Québec

Entraînement des modèles



 C'est à cette phase que les modèles seront paramétrisés.

Augmentation de précision
----->

Réduction du temps d'action potentiel

Aide à l'interprétation d'un modèle prédictif

Le choix des modèles sélectionnés se fait en fonction du meilleur niveau de prédiction du risque de décrochage à partir de centaines de variables décrivant le profil complet des élèves, issues de cinq catégories :



Pour chacun des modèles présentés, deux niveaux d'interprétation des modèles sont possibles :

Niveau global



La table des variables d'importance procure la liste des 30 variables les plus importantes pour inférer le niveau de risque de l'ensemble des élèves.

Le rapport détaillé des variables présente le détail des résultats de chacune des variables utilisées dans le modèle.

Niveau individuel



Le profil résumé d'un élève indique l'importance des cinq catégories de variables pour estimer le risque de décrochage de cet élève.

Attention! De tels modèles ne permettent pas de déterminer quelles variables sont susceptibles de **causer** le décrochage!

Résultats et conclusions du projet pilote (PPPSIA 20-21)

1. **Qualité des modèles**

Les modèles se sont avérés de bonne qualité dans tous les CSS, à différents niveaux.

2. **Nouvelles ressources**

L'infrastructure déployée et le recrutement d'un scientifique de données ont été des facteurs clés de succès.

3. **Personnalisation**

Déterminer des définitions propres aux CSS et prendre en compte le contexte de ceux-ci est une étape nécessaire pour rendre les modèles pertinents et utiles.

4. **Réplication**

Légitime de penser que l'application de la méthodologie des modèles est possible à tous les types de CSS et CS au Québec.

5. **Enthousiasme**

Les CSS ont partagé leur grande volonté de poursuivre les analyses prédictives ou une démarche de valorisation des données.

6. **Covid-19**

Une approche de mitigation de l'impact du COVID-19 a été développée et permet d'obtenir des modèles prédictifs de qualité malgré les variations dans les données recueillies depuis la pandémie.

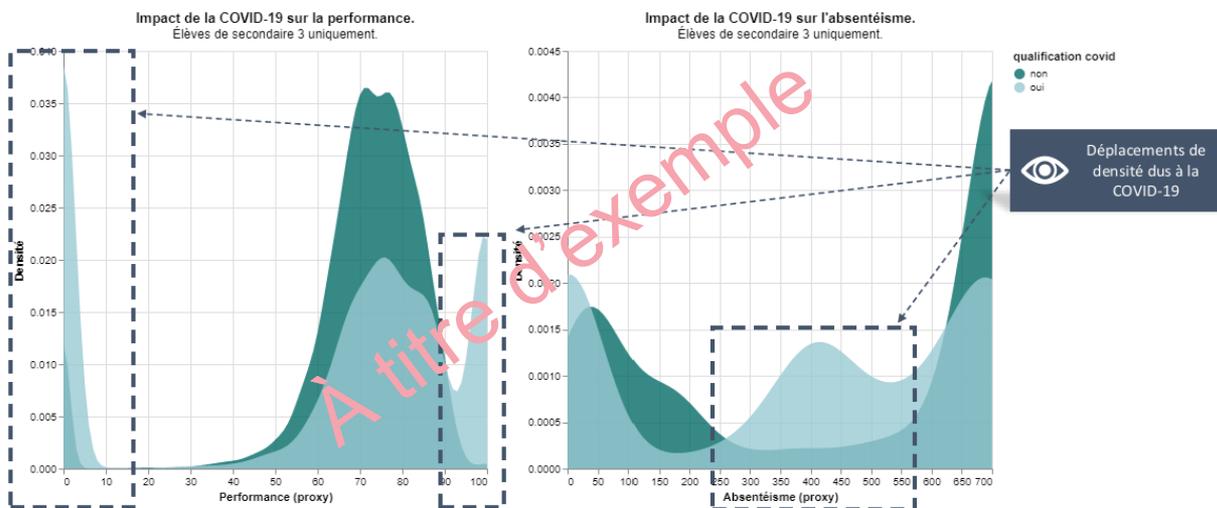
Changement de pratique : validité + valorisation + visualisation des données

Centre
de services scolaire
du Val-des-Cerfs

Québec 

Impacts de la COVID-19 sur les données sources

Au cours du projet pilote, nous avons **bonifié la robustesse de notre approche** afin de prendre en compte l'impact des données sources en retirant des modèles les variables ayant subi une trop grande influence. **La dérive conceptuelle** (illustrée ci-dessous) de la donnée est un phénomène crucial à prendre en considération en analyse prédictive et en apprentissage automatique. La statistique de **Kolmogorov-Smirnov** permet de mesurer l'écart entre deux lois, ce qui permet de filtrer les variables qui ont trop dérivé. Un **rapport technique** sur ce phénomène a été produit pour le CSS du Val-des-Cerfs.



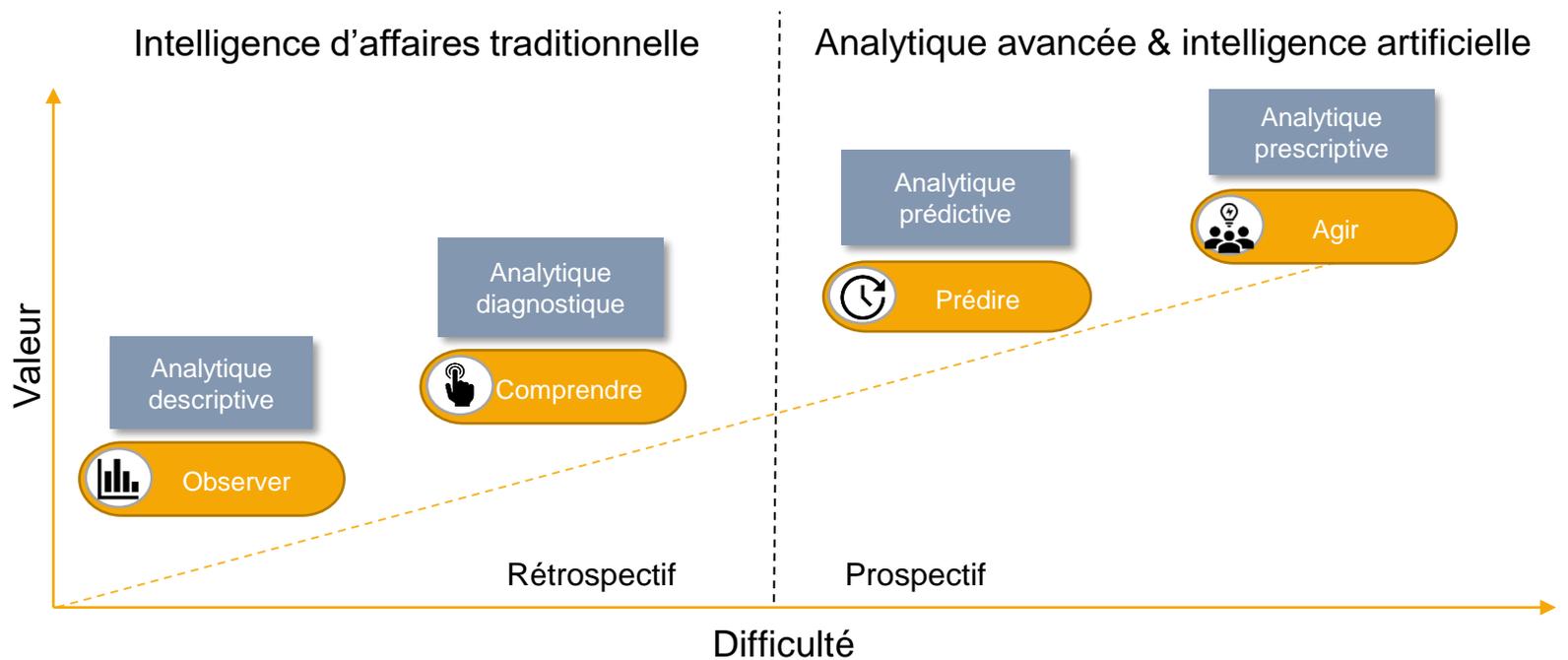
(Données du CSS du Val-des-Cerfs uniquement)

Centre
de services scolaire
du Val-des-Cerfs

Québec



L'échelle de valorisation de la donnée



3 VUES DE TABLEAU DE BORD : PERSÉVÉRANCE / RÉUSSITE / PERFORMANCE

Exemple de tableau de bord persévérance

Titre du modèle: p6_s3_to_s5

Filtres: École: Tout, Fiche: Tout

Statistiques de risque:

- Nbr. d'élèves à risque élevé: 45
- Nbr. d'élèves à risque modéré: 107
- Nbr. d'élèves à risque faible: 113

Tableau d'incidence des 5 grandes catégories par élève:

Fiche	Niveau de risque	Probabilité
1490256V	Risq	10,00%
1490256Q	Risq	10,00%
1490256T	Risq	10,00%
1490256Z	Risq	10,00%
14902564	Risq	10,00%
14902568	Risq	10,00%
1490256C	Risq	10,00%
1490256G	Risq	10,00%
1490256J	Risq	10,00%
1490256N	Risq	10,00%
1490256P	Risq	10,00%
1490256R	Risq	10,00%
1490256S	Risq	10,00%
1490256V	Risq	10,00%
1490256X	Risq	10,00%
1490256Y	Risq	10,00%
1490256Z	Risq	10,00%

Explication par catégorie (Facteur explicatif):

- performance
- assiduité
- comportement
- variablesmex
- profil scolaire

Détails des catégories:

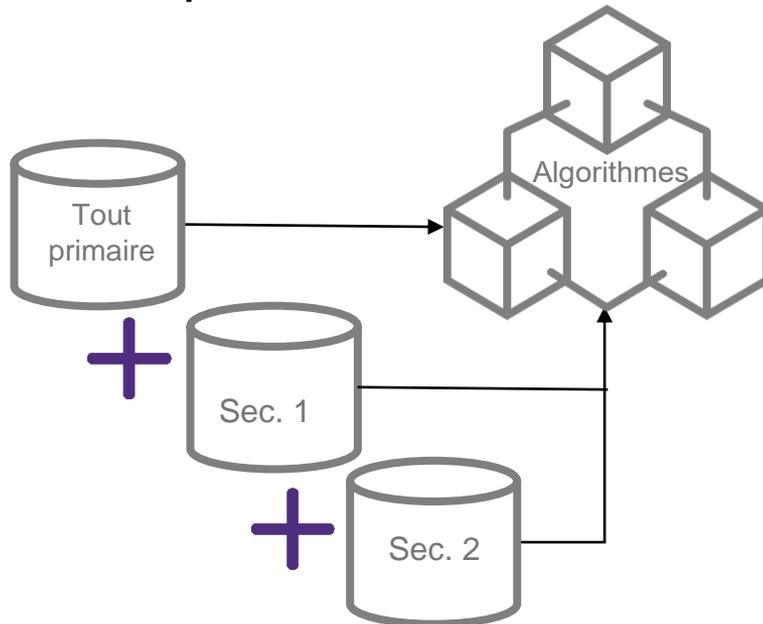
catégorie	mon_result...	mon_resu...
autres		
rid_resul_etape_globa		

PROTOTYPE

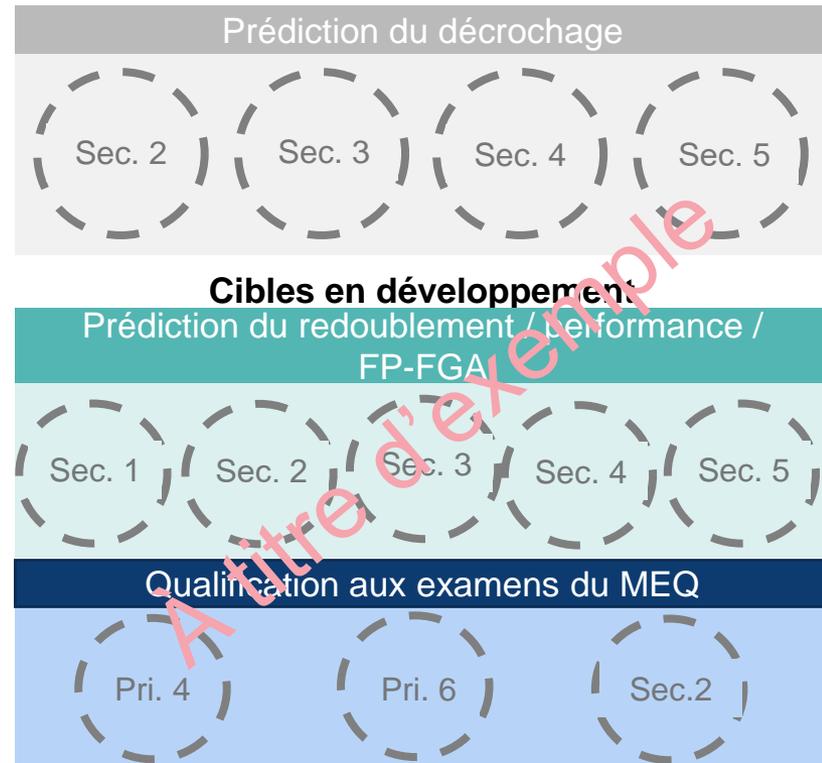
Entraînement des modèles

- **Entraînement, évaluation** de la performance et **sélection** de modèles;
- Qualification des cohortes d'élèves et explications par élèves.

Prédiction pour différents niveaux scolaires



Les modèles utilisent **différemment** le même **bassin de données**.

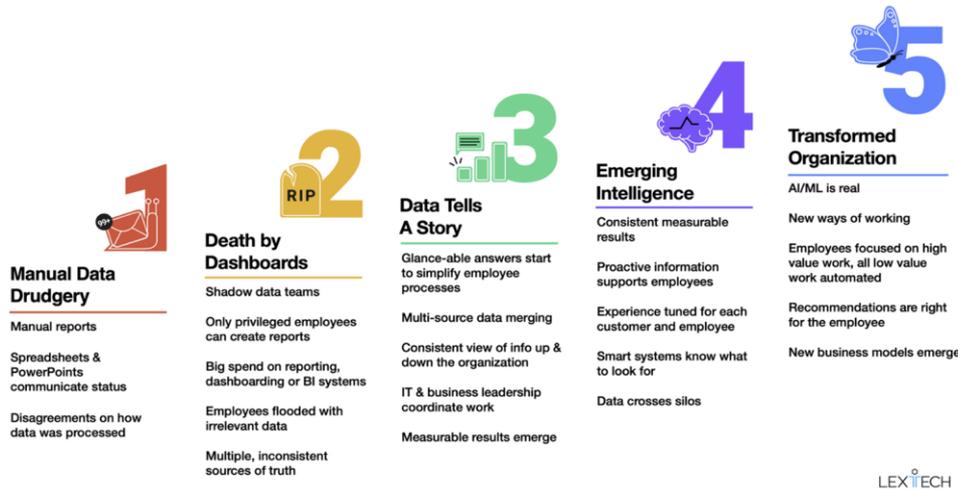


Centre
de services scolaire
du Val-des-Cerfs

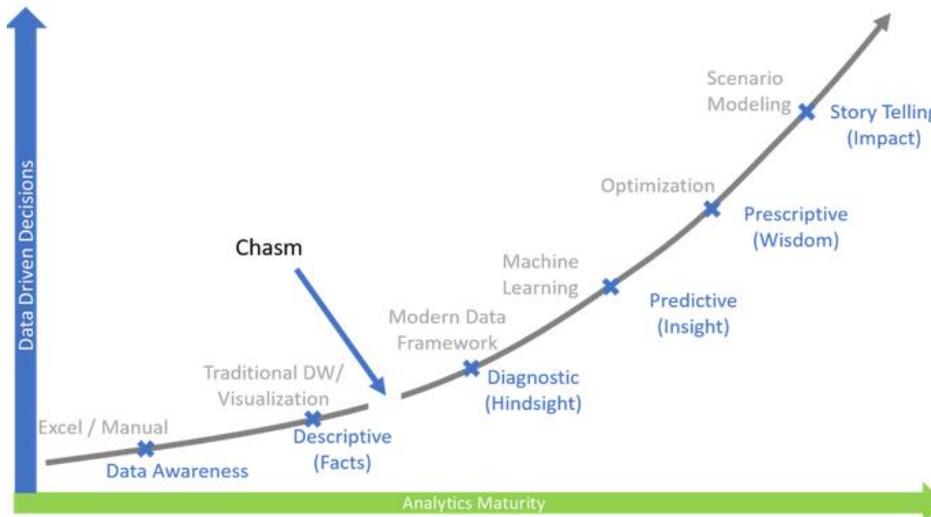
Québec

Les pièges

La mort par tableau de bord

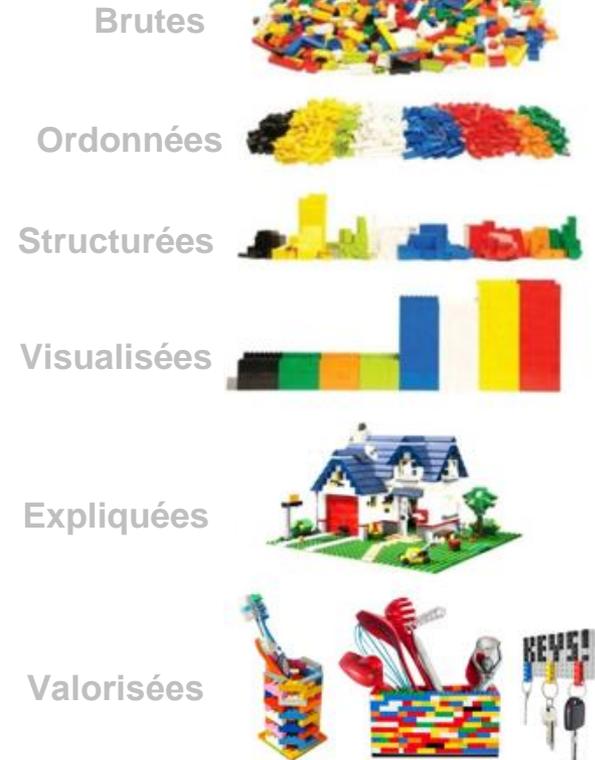


Source : [Alex Bratton](#)



Source : [The digital talk](#)

Les données :



Source : [Reddit: The Lego Data Story](#), adapted from original image by Monica Rosales Ascenio (Hot Butter Studio) 2012



**Centre
de services scolaire
du Val-des-Cerfs**

Québec 