

Série de webinaires

IA EN PHYTOPROTECTION

APPROCHES PRÉDICTIVES ET OUTIL D'AIDE À LA DÉCISION EN GRANDES CULTURES

Bienvenue



Gestion des données géospatiales pour une infrastructure numérique intelligente en phytoprotection

Vers une aide à la décision intelligente et souveraine

Mickaël Germain

Université de Sherbrooke

Département de Géomatique appliquée

Contributeurs: *Gatien Romuald Kenfack Nguemo, Samuel Foucher, Ramata Magagi*

Webinaire innovations en IA et phytoprotection - 12 février 2026



Mickaël Germain, Professeur

Département de géomatique appliquée

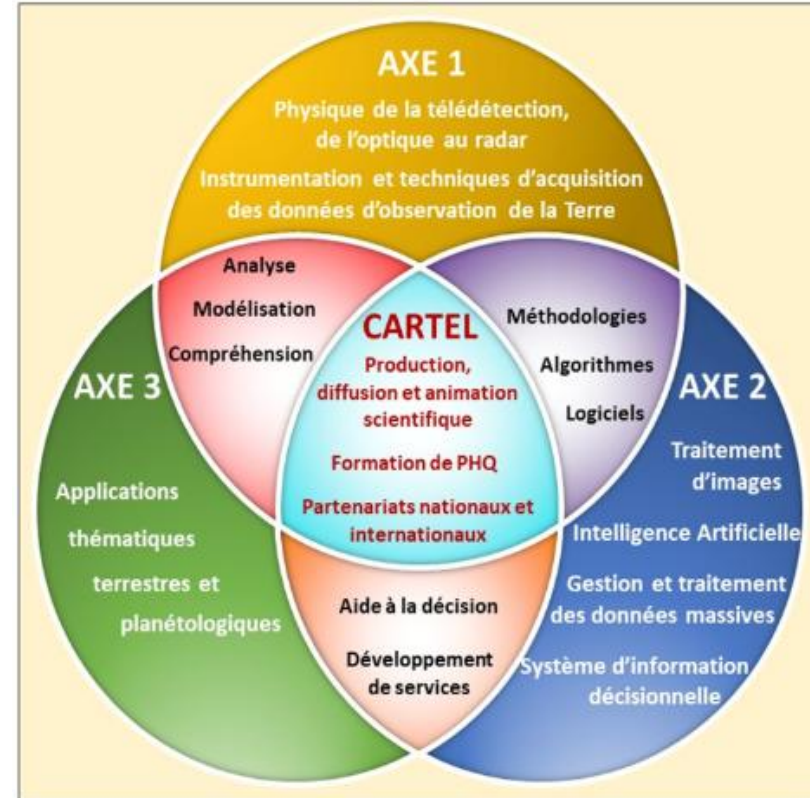


- **Expertises:**
 - SIG (web), systèmes d'aide à la décision, gestion des données géospatiales, informatique en périphérie (*edge computing*), fusion de données, intelligence artificielle, internet des objets (*IoT*)
- **Intérêts de recherche:**
 - Développement d'outils d'aide à la décision pour diverses applications:
 - Biodiversité; ■ Agriculture de précision; ■ Géologie; ■ Environnement urbain
 - Fusion de données multicapteurs et multisources
 - Mise en place d'infrastructures en *Edge Computing*
 - Apprentissage profond pour des applications de classification et de segmentation
 - Interfaces web collaboratives pour l'annotation des données
 - Systèmes de gestion des données géospatiales (bases/entrepôts/lacs de données)

Département de géomatique appliquée

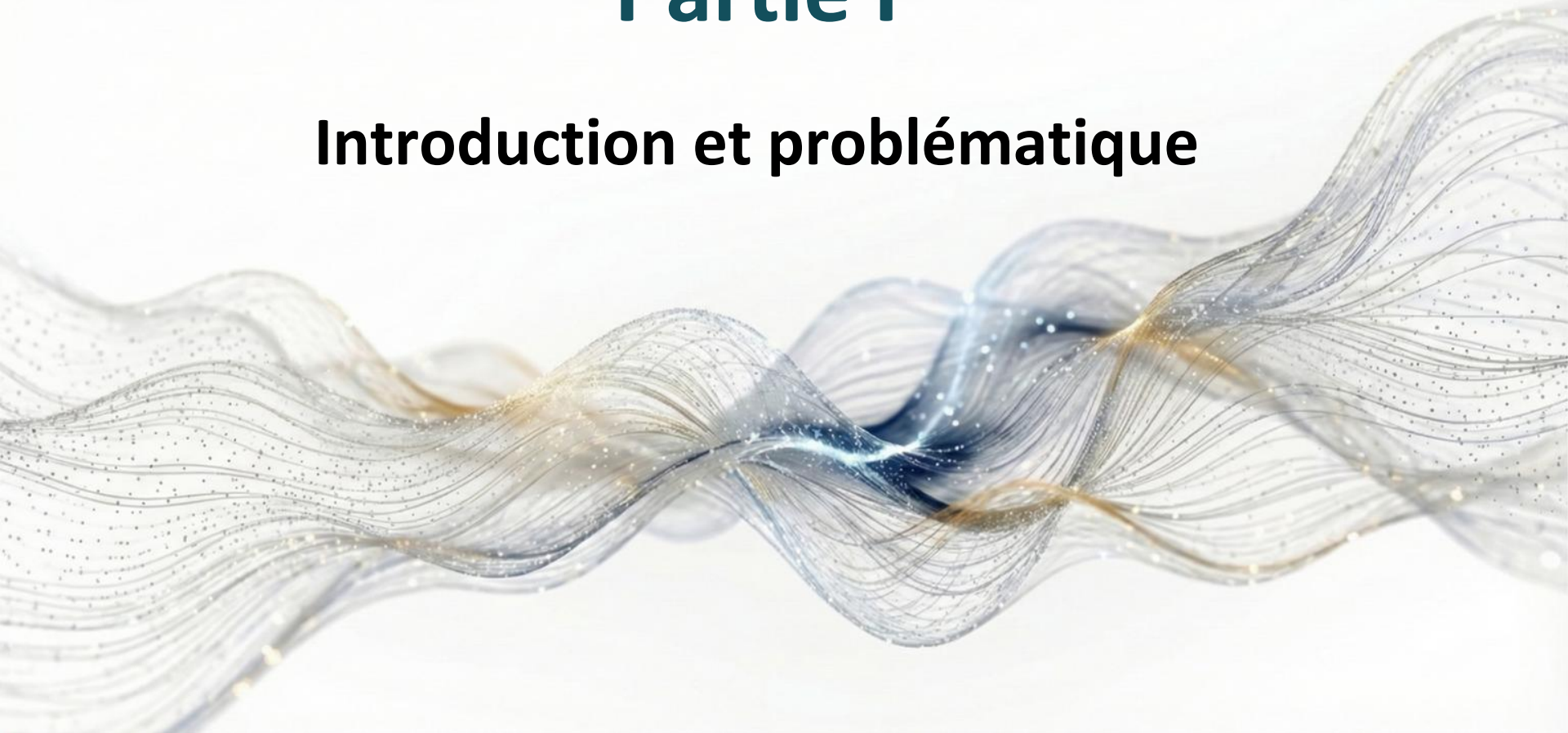


- 12 professeurs
- 100+ étudiants aux 3 cycles
- <https://www.usherbrooke.ca/geomatique/>
- Trois axes:
 - *Physique de la télédétection*
 - *Traitement de la donnée*
 - *Applications thématiques*
- Thématiques:
 - Agriculture de précision
 - Milieux nordiques
 - Foresterie
 - Biodiversité
 - Cartographie



Partie I

Introduction et problématique

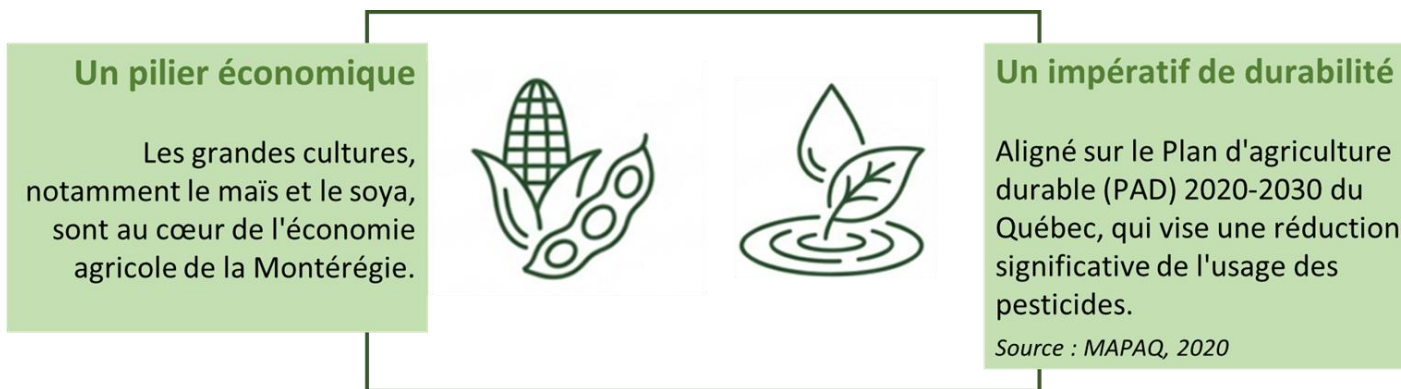


Pourquoi une infrastructure intelligente pour la gestion des grandes cultures au Québec ?

L'agriculture du Québec est confrontée à une double pression :

1. Augmenter la **productivité** pour répondre aux besoins du marché
2. Adopter des **pratiques durables** pour préserver la santé de nos sols et de nos cours d'eau.

Trouver un équilibre est essentiel pour la résilience et la rentabilité à long terme du secteur.



Ennemis visés en Montérégie

- La **Sclérotiniose du soya** est la maladie la plus dommageable au Québec.
- Son apparition dépend de trois facteurs clés, dont l'humidité et la température.
- Le **Chénopode blanc (CHEAL)** est une mauvaise herbe problématique au Québec.
- Il est très compétitif et développe une résistance aux herbicides.
- La gestion de ces deux ennemis est prioritaire en Montérégie.



Jusqu'à **20%** de pertes de rendement.

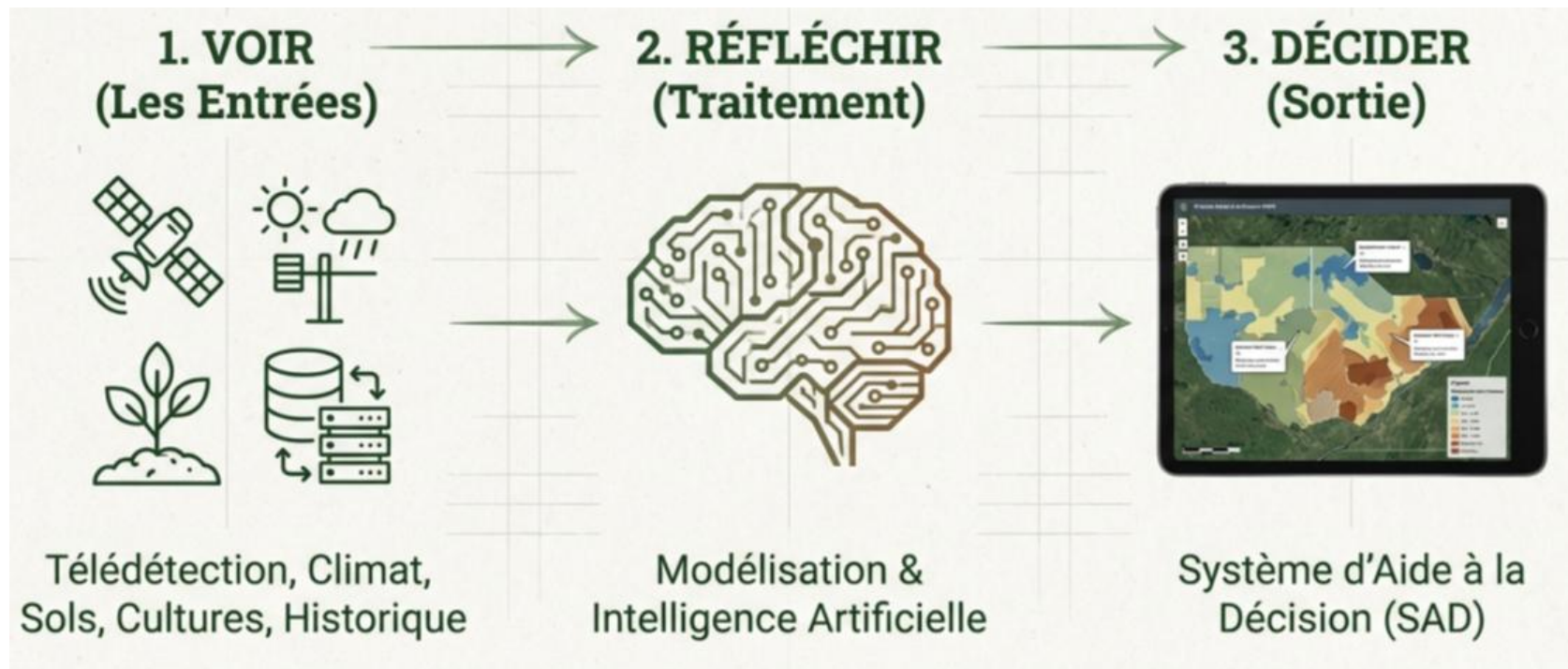
Source : CÉROM



Jusqu'à **25%** de pertes de rendement en soya et **13%** en maïs.

Source : CÉROM

Proposition d'une infrastructure numérique intelligente (INI) - un écosystème pour l'aide à la décision



Les questions centrales de ce webinaire

1



Comment stocker et organiser des données géospatiales hétérogènes ?

2



Comment préserver la confidentialité tout en alimentant des modèles d'IA ?

3



Quel rôle jouent les métadonnées dans l'interopérabilité ?

4



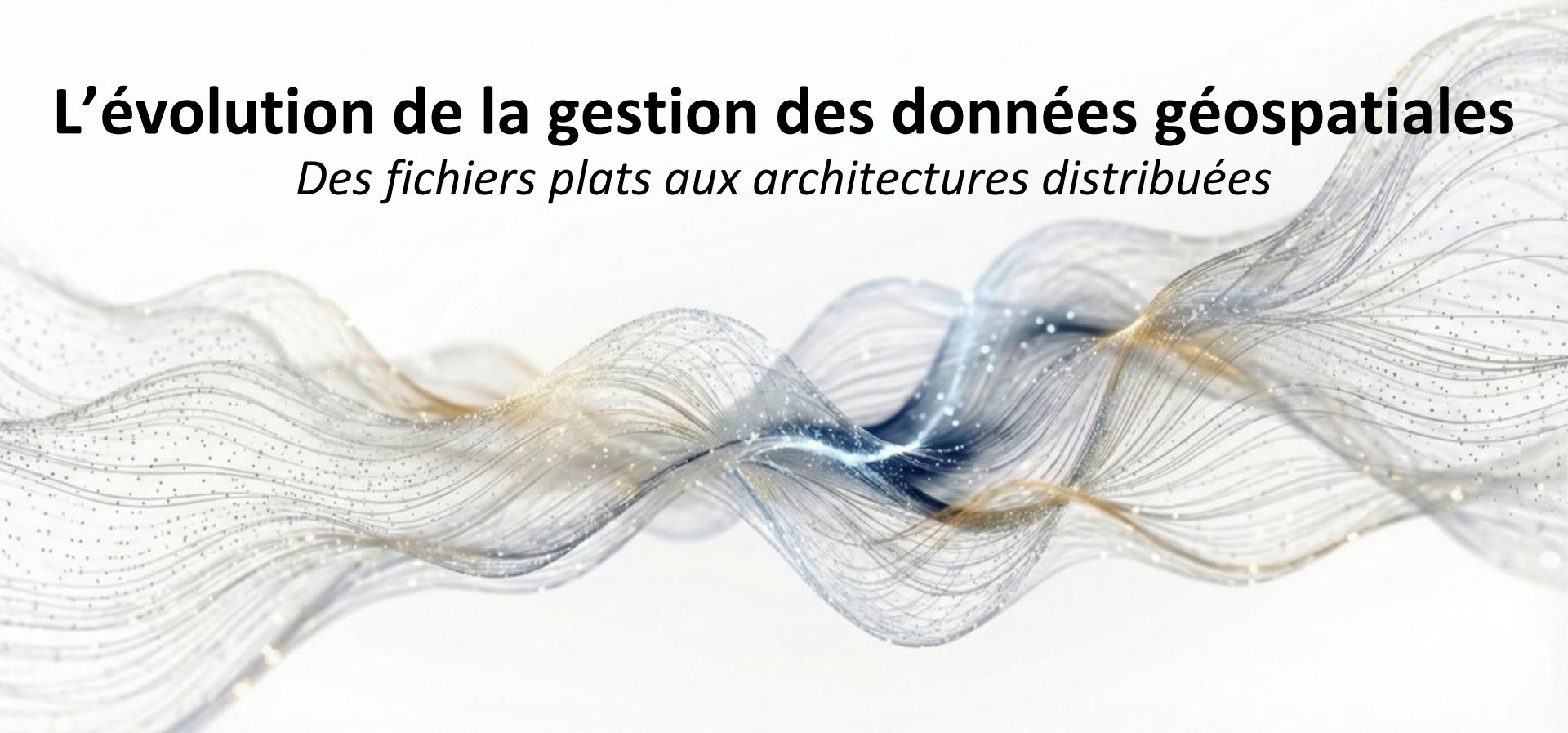
Pourquoi une approche fédérée est-elle la solution pour l'agriculture ?



Partie II

L'évolution de la gestion des données géospatiales

Des fichiers plats aux architectures distribuées



Quels sont les formats des données géospatiales ?

OGR Vector Formats

Format Name	Code	Creation	Georeferencing	Compiled by default
Aeronav FAA files	AeronavFAA	No	Yes	Yes
AmigoCloud API	AmigoCloud	Yes	Yes	No, needs libcurl
ESRI ArcObjects	ArcObjects	No	Yes	No, needs ESRI ArcObjects
Arc/Info Binary Coverage	AVCBin	No	Yes	Yes
Arc/Info E00 (ASCII) Coverage	AVCE00	No	Yes	Yes
Arc/Info Generate	ARCGEN	No	No	Yes
Atlas BNA	BNA	Yes	No	Yes
AutoCAD DWG	CAD	No	Yes	Yes (internal libopencad provided)
AutoCAD DWG	DWG	No	No	No, needs Open Design Alliance Teigha library
AutoCAD DXF	DXF	Yes	No	Yes
Carto	Carto	Yes	Yes	No, needs libcurl
Cloudant / CouchDB	Cloudant	Yes	Yes	No, needs libcurl
CouchDB / GeoCouch	CouchDB	Yes	Yes	No, needs libcurl
Comma Separated Value (.csv)	CSV	Yes	No	Yes
OGC CSW (Catalog Service for the Web)	CSW	No	Yes	No, needs libcurl
Czech Cadastral Exchange Data Format	VFK	No	Yes	No, needs libsqlite3
DB2 Spatial	DB2ODBC	Yes	Yes	No, needs ODBC library
DODS/OPeNDAP	DODS	No	Yes	No, needs libdap
Google Earth Engine Data API	EEDA	No	Yes	No, needs libcurl
EDIGEO	EDIGEO	No	Yes	Yes



Des centaines de formats vectoriels et matriciels

https://www.gdal.org/ogr_formats.html

Quels sont les formats des données géospatiales ?

- **La réalité : 10 formats** représentent plus de **95 % des formats géospatiaux** (exemple des jeux de données ouverts au Québec)

The screenshot shows the homepage of the 'Données Québec' portal. The header is dark blue with the 'Québec' logo and the text 'Partenariat Données Québec' and 'Nous joindre'. Below the header is a navigation bar with links: 'JEUX DE DONNÉES', 'ORGANISATIONS', 'APPLICATIONS', 'CARTE INTERACTIVE', and 'GUIDES ET INFORMATIONS'. The main content area has a large 'Données Québec' title and a subtitle: 'Le portail de données ouvertes né de la collaboration entre les villes et le gouvernement du Québec'. A search bar with the placeholder 'Rechercher des jeux de données' is present. At the bottom, a light blue bar displays statistics: '1591 jeux de données', '139 organisations', and '51 applications'.

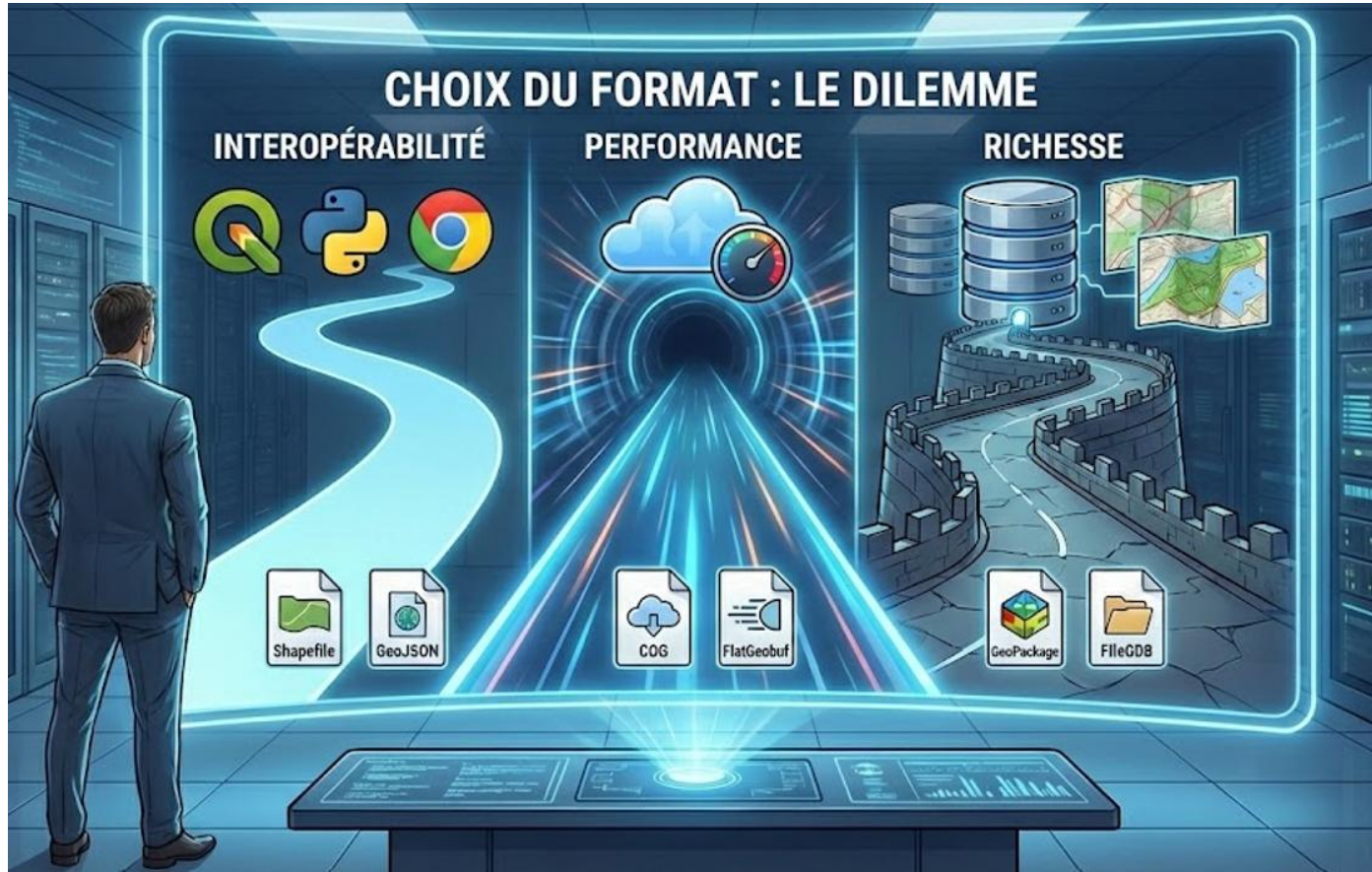
Statistique	Valeur
Jeux de données	1591
Organisations	139
Applications	51

<https://www.donneesquebec.ca/>

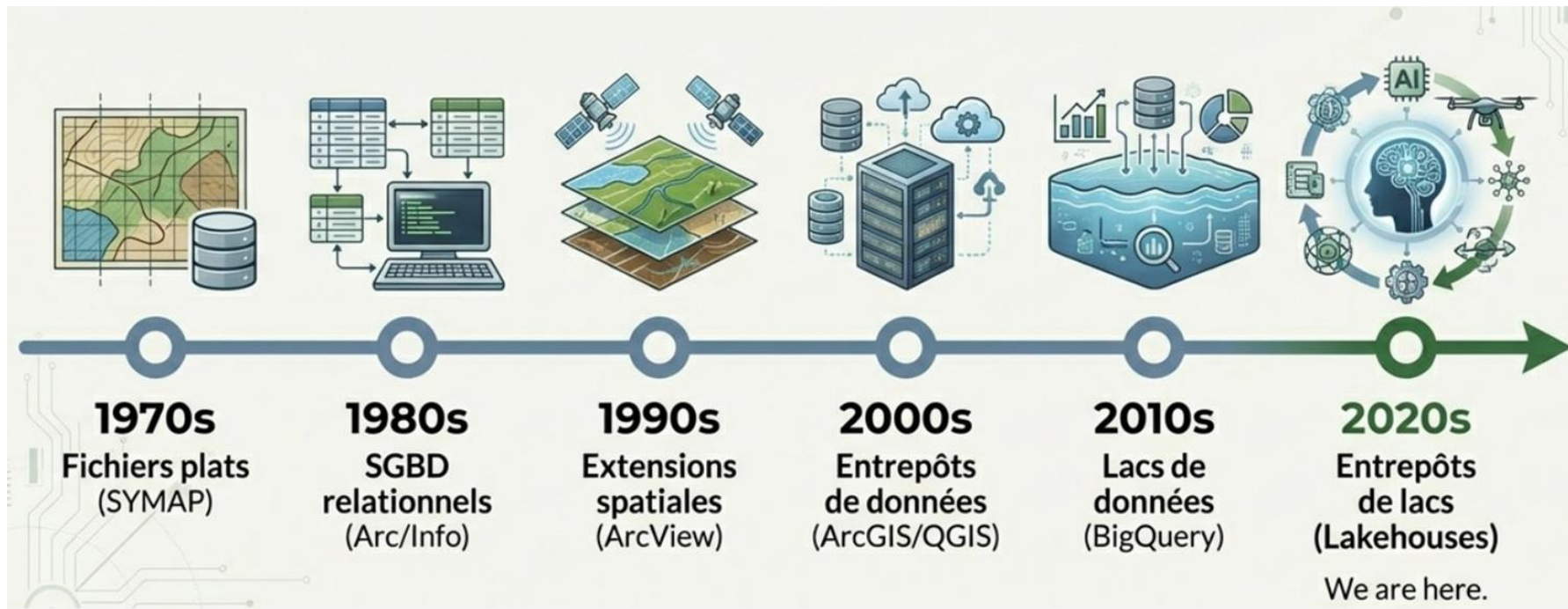
Formats

CSV (1142)
GeoJSON (629)
SHP (604)
PDF (426)
HTML (291)
KML (269)
XLSX (190)
GPKG (178)
WMS (174)
JSON (171)
FGDB (144)
REST (113)
ZIP (105)
WEB (87)
XML (74)
XLS (51)
WFS (49)
GeoTIFF (36)
KMZ (33)
GTFS (19)
ODS (18)
SQLite (17)

Quels sont les formats des données géospatiales ?



Chronologie de l'évolution des accès aux données



Problème #1 de l'INI – les limites d'un modèle centralisé



Scalabilité

Le volume de données explose exponentiellement.



Latence

Délai trop long entre acquisition et disponibilité.



Coût

Infrastructures traditionnelles onéreuses.



Interopérabilité

Formats souvent incompatibles entre eux.



Confidentialité

Données agricoles critiques et sensibles.



Problème #2 de l'INI - l'interopérabilité des données



Systèmes de coordonnées

UTM, Lambert, WGS84... projections différentes



Résolutions spatiales

10 m (Sentinel-2) vs 30 m (Landsat) vs 1 km (MODIS)



Résolutions temporelles

Horaire (capteurs) vs 5 jours (satellite) vs mensuel



Vocabulaire et unités

Terminologie non standardisée entre organismes

Problème #3 de l'INI – la confidentialité des données



☐ Les craintes du partage des données

- La peur du **jugement** des pratiques agricoles
- Le risque de **perte d'avantage** concurrentiel en révélant ses “secrets”
- La **méfiance** légitime envers le contrôle institutionnel des informations

☐ Les obligations de la Loi 25

- La Loi 25 exige un **consentement manifeste, libre et éclairé**
- Elle impose des **obligations strictes de transparence** envers les utilisateurs
- L'**évaluation** des facteurs relatifs à la vie privée (EFVP) est maintenant obligatoire
- Ces mesures encadrent la **gestion des données numériques** au Québec

Problème #3 de l'INI – la confidentialité des données

❑ Des solutions classiques et leurs limites

Anonymisation

Suppression des identifiants



Limite : Ré-identification possible avec données spatiales !

Agrégation

Données à l'échelle régionale



Limite : Perte de la granularité intra-parcellaire

Accords de partage

Contrats bilatéraux



Limite : Lourdeur administrative, non scalable

Solution de l'INI : l'approche fédérée

« Apporter le modèle aux données, pas les données au modèle »

Principes clés

- Entraînement/analyse local sur chaque source
- Seuls les résultats sont partagés
- Agrégation centrale des modèles

Avantages

- Confidentialité préservée
- Scalabilité
- Robustesse aux données hétérogènes

**Clé de voûte de l'approche fédérée :
les métadonnées**

Partie III

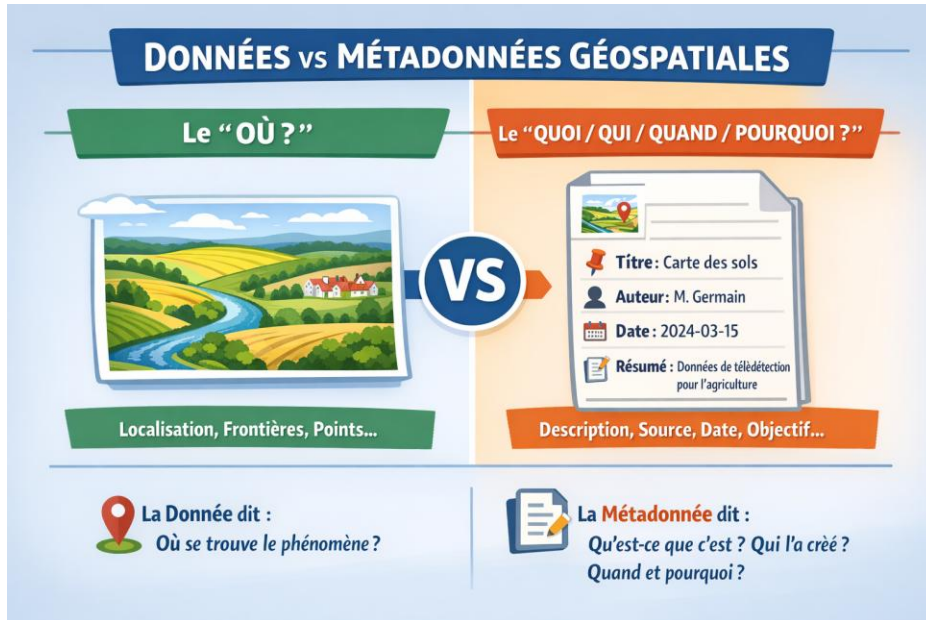
Les métadonnées : clé de voûte de l'approche fédérée

Documenter pour mieux intégrer



Qu'est-ce qu'une métadonnée ?

« Donnée sur la donnée » – Information décrivant le contenu, la structure et le contexte



- **Informations descriptives** : titre, date, auteur, résumé, etc.
- **Informations structurelles** : format, résolutions spatiale et temporelle, projection, description des champs, etc.
- **Information administrative** : droits, licences, provenances, etc.

Les normes de métadonnées en géomatique

Standard	Domaine	Usage
ISO 19115	Géographique	Métadonnées géospatiales
ISO 19139	Implémentation	Encodage XML
Dublin Core	Générique	Métadonnées documentaires
STAC	Raster / EO	Catalogue spatio-temporel
DCAT	Catalogue	Description de catalogues

Focus : STAC (SpatioTemporal Asset Catalog)

Norme moderne adoptée par : Microsoft Planetary Computer, Google Earth Engine, AWS Open Data, Copernicus.

**Pour le système INI : adoption de STAC et ISO 19115
pour le catalogue fédéré**

Le rôle des métadonnées dans une base fédérée



Découvrabilité

Trouver les données sans les centraliser

Avantage INI : catalogue unifié sans partage des données



Interopérabilité

Comprendre les données sans conversion

Avantage INI : création d'une ontologie agronomique



Traçabilité

Connaître l'origine (≠ confidentialité) et la qualité

Avantage INI : données de qualité



Automatisation

Permettre aux agents IA d'interroger le catalogue

Avantage INI : intégration IA

Partie IV

L'architecture fédérée : la solution pour l'Agriculture 5.0

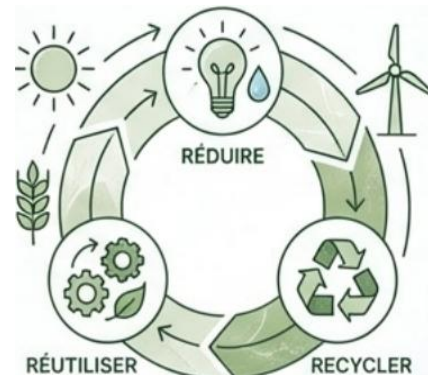
Base de données fédérée + Apprentissage fédéré



L'Agriculture 5.0

Caractéristique	Agriculture 4.0	Agriculture 5.0
Focus principal	Connectivité et données (IoT)	Intelligence et Autonomie (IA/Robots)
Rôle de l'humain	Analyseur de données	Superviseur stratégique
Objectif	Productivité et efficacité	Durabilité et résilience sociale

- Le 5.0 met l'accent sur l'**intelligence artificielle (IA)**, les **agents autonomes** et, surtout, la **collaboration entre l'humain et la machine**
- La grande différence avec l'industrie 4.0. **L'agriculture 5.0 réintègre l'agriculteur au centre du processus**
- **Intégrer des modèles de production circulaires**

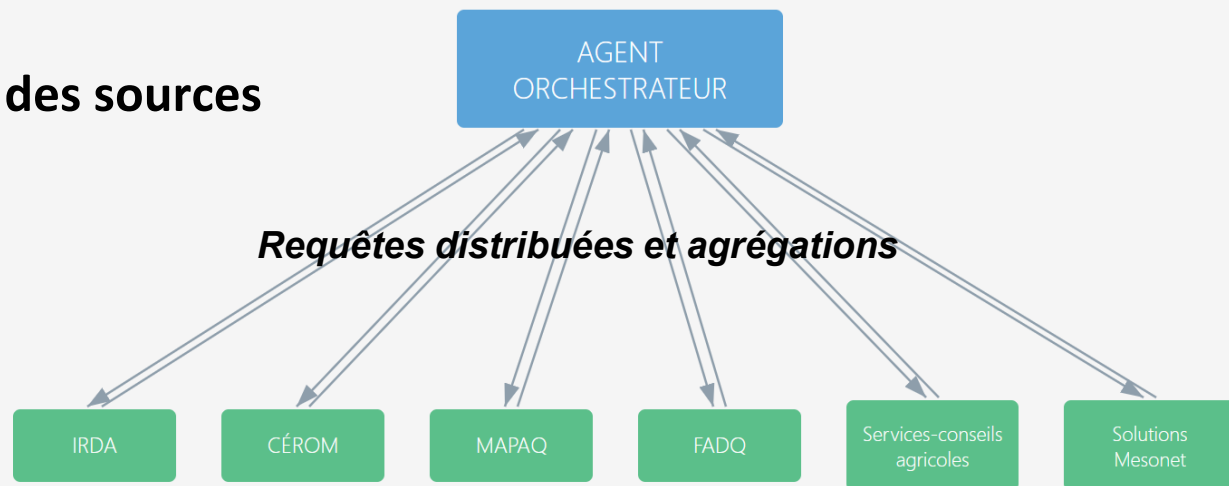


L'Agriculture 5.0 – rôle de la base de données fédérée

Système INI intégrant plusieurs bases autonomes qui conservent leur indépendance

Principes clés

- **Pas de centralisation** des données brutes
- **Requêtes distribuées** via un agent (administrateur) qui a accès aux **métadonnées**
- Respect de l'**autonomie des sources**

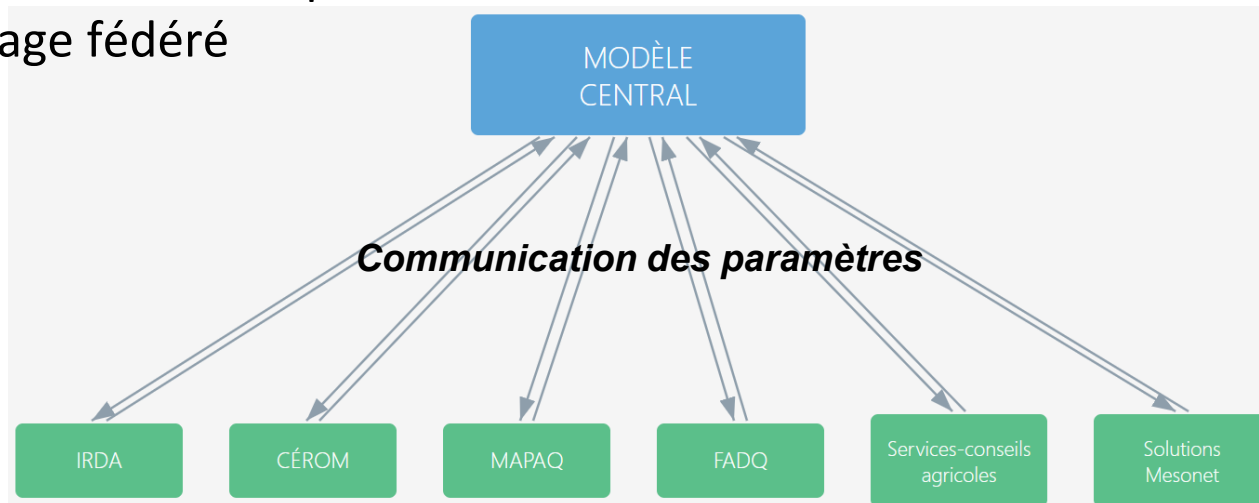


L'Agriculture 5.0 – rôle de l'apprentissage fédéré

Système INI intégrant un modèle central qui intègre plusieurs modèles locaux

Principes clés

- Seuls les « apprentissages » (modèles) sont partagés
- Le modèle global s'améliore sans compromettre la fiabilité
- Analogie de l'apprentissage fédéré en santé



L'Agriculture 5.0 – rôle de l'apprentissage fédéré

Quand la médecine montre la voie à l'agriculture

	MÉDECINE 	AGRICULTURE 
Données sensibles	 Dossiers patients	 Rendements / Pratiques
Lieu de stockage	 Hôpitaux	 Fermes
Objectif de l'IA	 Prédire les maladies	 Prédire les ravageurs

Dans les deux cas, le frein est la confidentialité.
La solution? L'apprentissage fédéré.

Les avantages de l'approche fédérée pour l'INI

« Une intelligence distribuée pour une agriculture durable »



Souveraineté des données

Chaque partenaire garde
le contrôle total



Scalabilité

Ajout facile de nouveaux
partenaires et sources



Robustesse

Pas de point unique de
défaillance



Conformité réglementaire

Respect de la Loi 25



Performance IA

Modèles généralisables grâce
à la diversité

Partie V

Collecte et accès aux données

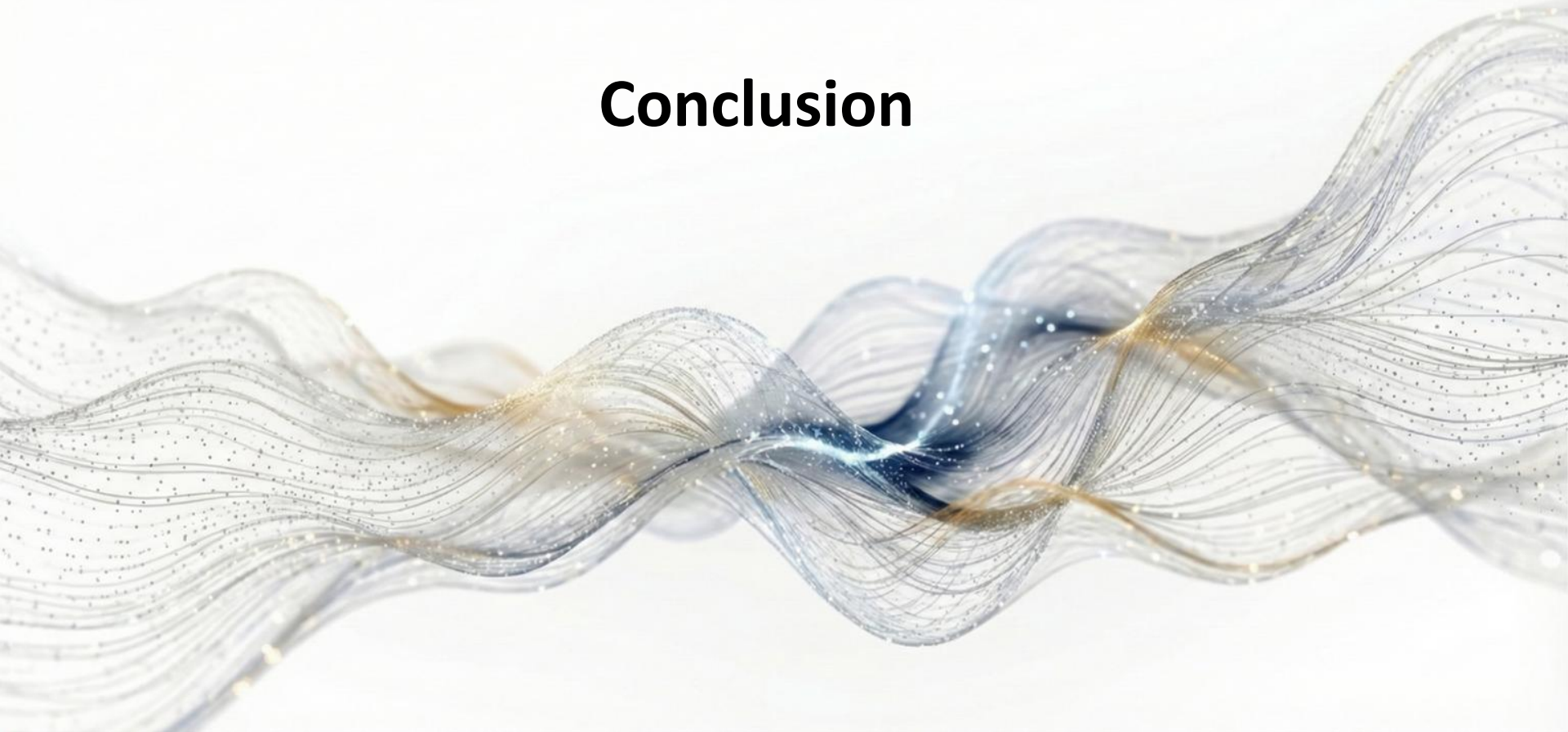


Collecte et accès aux données

- **Données historiques en Montérégie**
 - Archives de l'IRDA, du CÉROM, de la FADQ et du MAPAQ,
 - Inventaires de mauvaises herbes remontant aux années 1980
- **Données climatiques**
 - Scénarios actuels et futurs (jusqu'en 2100)
- **Données collectées durant les dépistages (2024 et 2025)**
 - Dépistage du CHEAL sur **10 champs de maïs** et des parcelles expérimentales (IRDA, CÉROM)
 - Dépistage de la sclérotiniose sur **15 champs de soya / année**
- **Réseaux de sondes installées dans les champs**
 - Mesures en continu (toutes les 15 ou 30 minutes) : humidité et température sol
- **Données de télédétection**
 - Données optiques et radar

Partie VI

Conclusion



Des résultats concrets pour l'agriculture québécoise

Le projet produira un ensemble d'outils et de connaissances directement applicables par le milieu agricole.



Une plateforme de données unifiée (INI).

Une infrastructure fédérée regroupant des informations multisources fiables pour la modélisation en phytoprotection.



Des modèles prédictifs précis.

Des modèles spatio-temporels pour la sclérotiniose et le CHEAL, améliorés par l'IA et validés pour les conditions du Québec.



Un système de recommandation intelligent (SAD).

Un outil d'aide à la décision pour optimiser l'usage des pesticides et d'autres pratiques agricoles.



Une nouvelle génération d'experts.

Formation de personnel hautement qualifié (7 Ph.D, 2 M.Sc prévus) pour relever les défis de l'agriculture durable.

Ce projet est financé par le ministère de l'Agriculture, des
Pêcheries et de l'Alimentation (MAPAQ) et le Fonds de
recherche du Québec (FRQ)



Partenaires

