

# **Techniques d'IA, Données et Connaissances d'Experts**

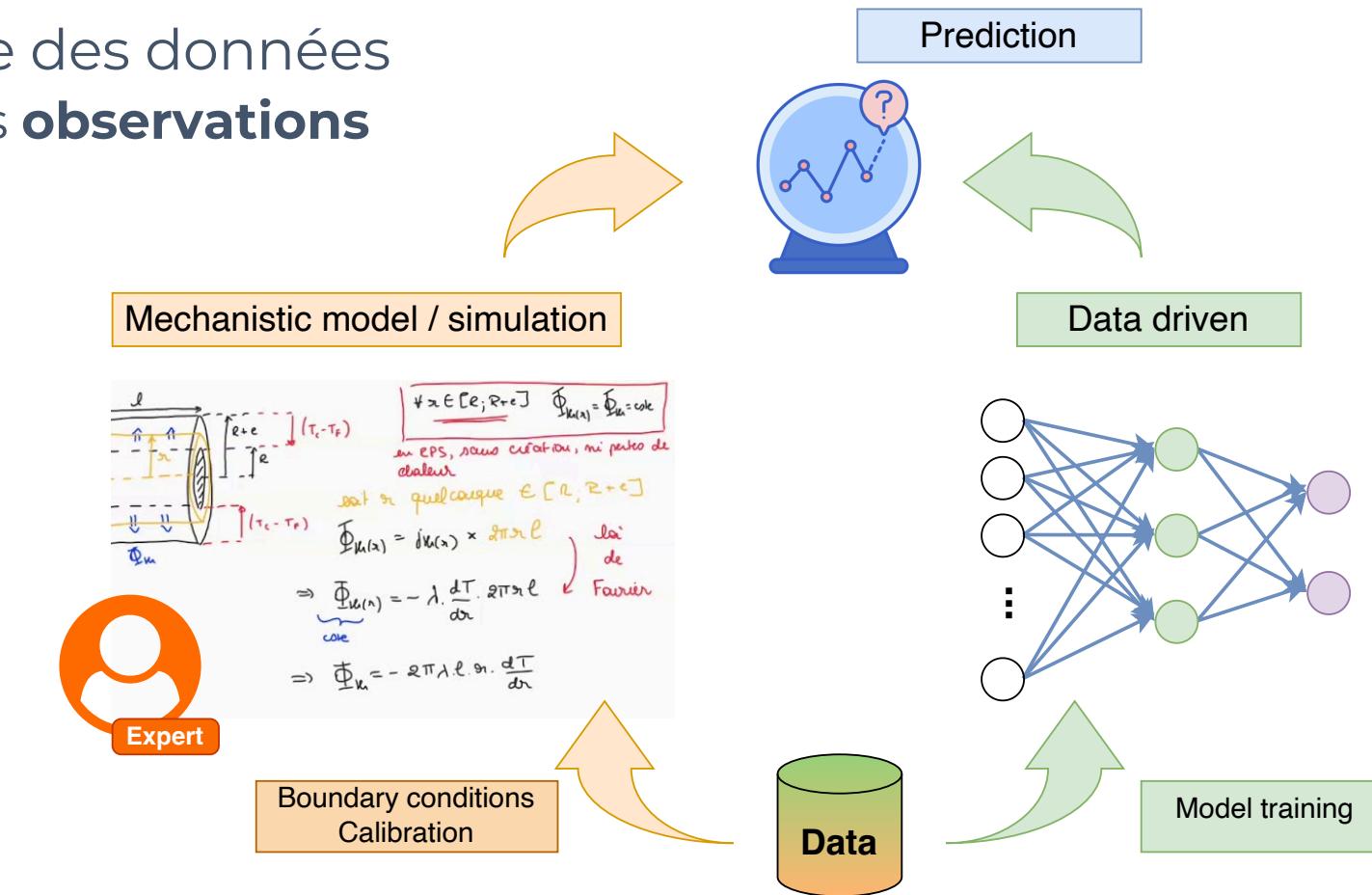
## **dans le projet AI4Diet**

**Vincent Guigue**

Professeur d'informatique

# Des données à la modélisation : différentes architectures

- Modélisation mécaniste
  - Traduire les **connaissances des experts** en équations
- Intelligence Artificielle / Science des données
  - Extraire les connaissances des **observations**

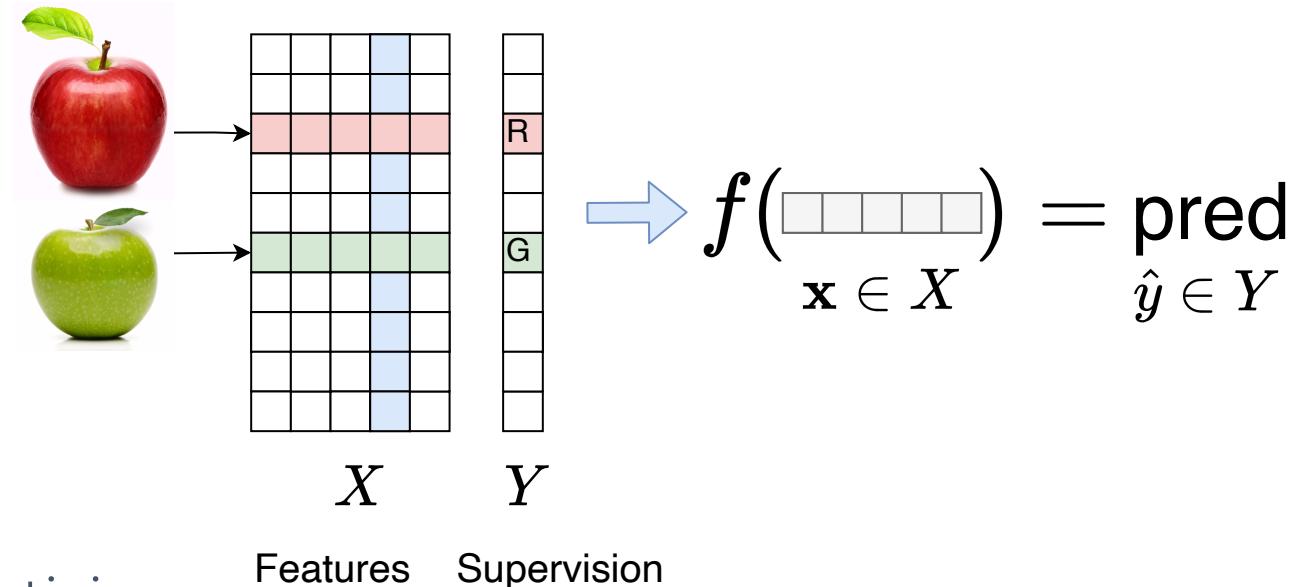


## □ Données

- Entrées (=observations)
- Sorties (=prédictions)

## □ Problématiques usuelles

- Supervision (données de sortie)
- Volume de données disponibles
- Périmètre des données
- Périmètre de validité du modèle
- Incompréhensions entre les informaticiens et le terrain



Besoin d'expertise limité

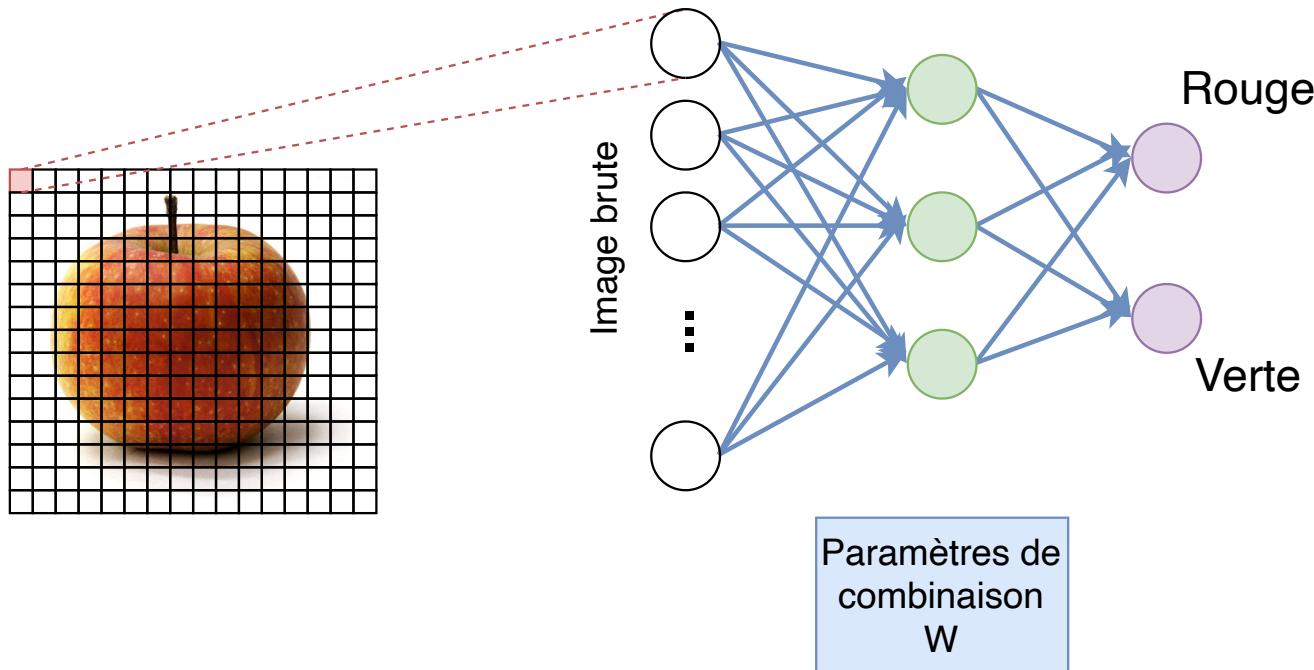
Librairies matures et faciles d'accès

Collecte des données de plus en plus facile

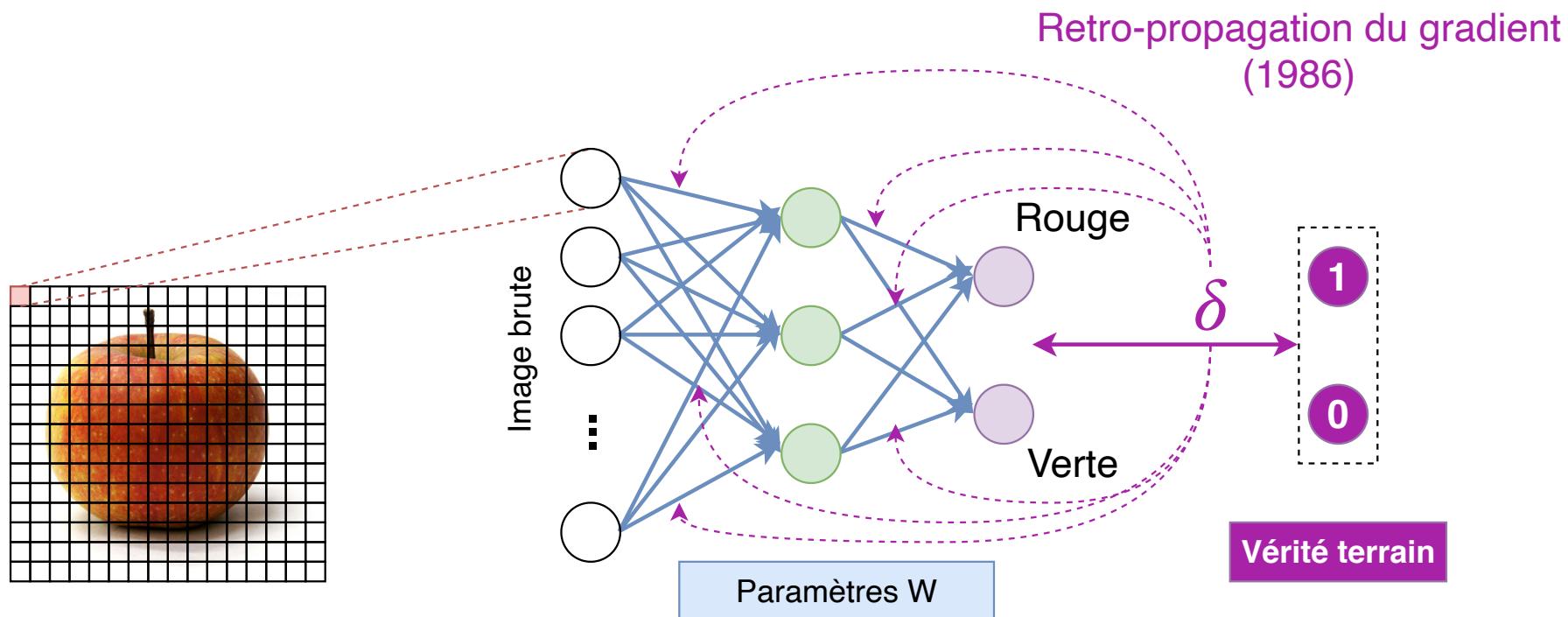
Traitement de données hétérogènes (texte, nombres, ...)

Performances non garanties  
Explicabilité / boîtes noires

- Opérateur mathématique hyper-complexe
  - Gestion de données brutes riches (texte, image, ...)
- Système de correction (progressive) des erreurs

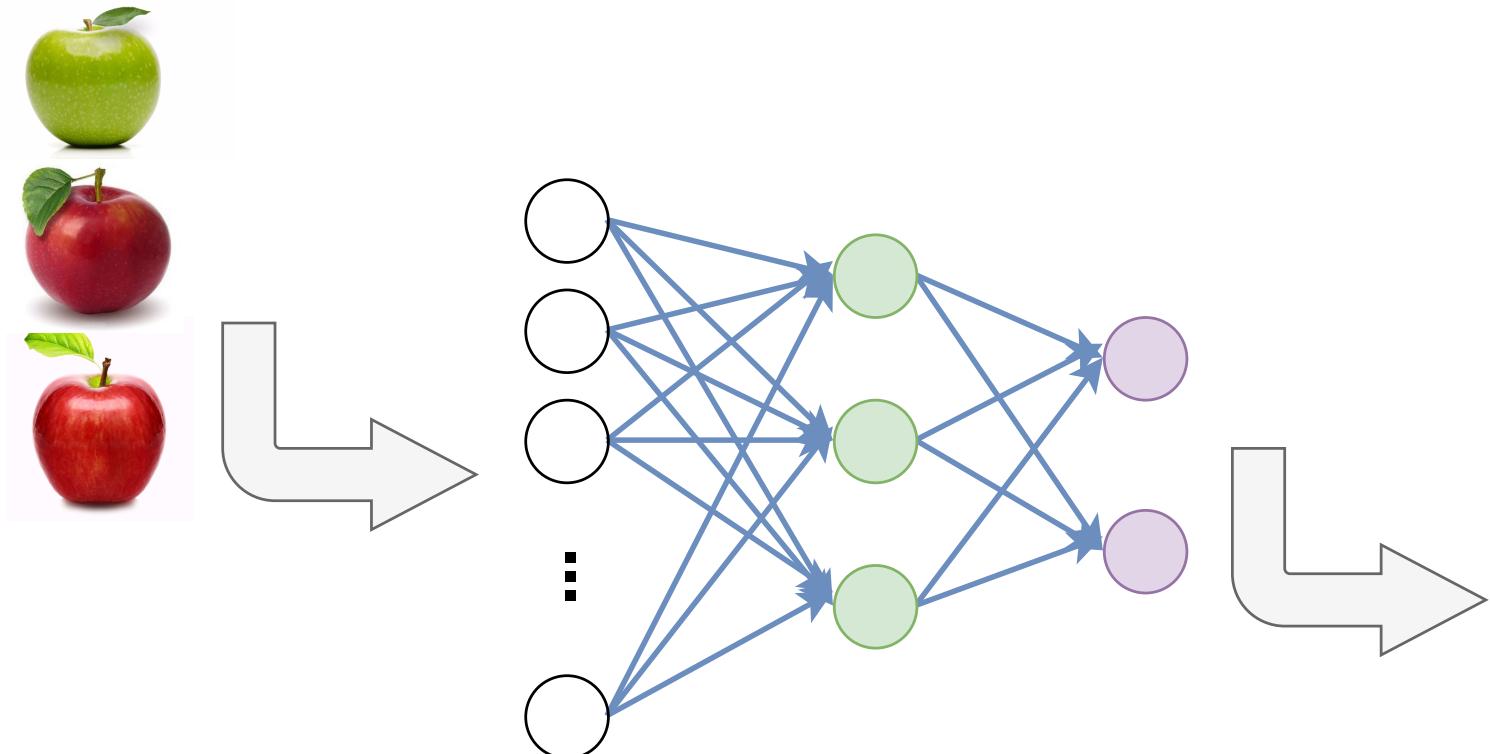


- Opérateur mathématique hyper-complexe
  - Gestion de données brutes riches (texte, image, ...)
- Système de correction (progressive) des erreurs



# Réseaux de neurones : architectures complexes...

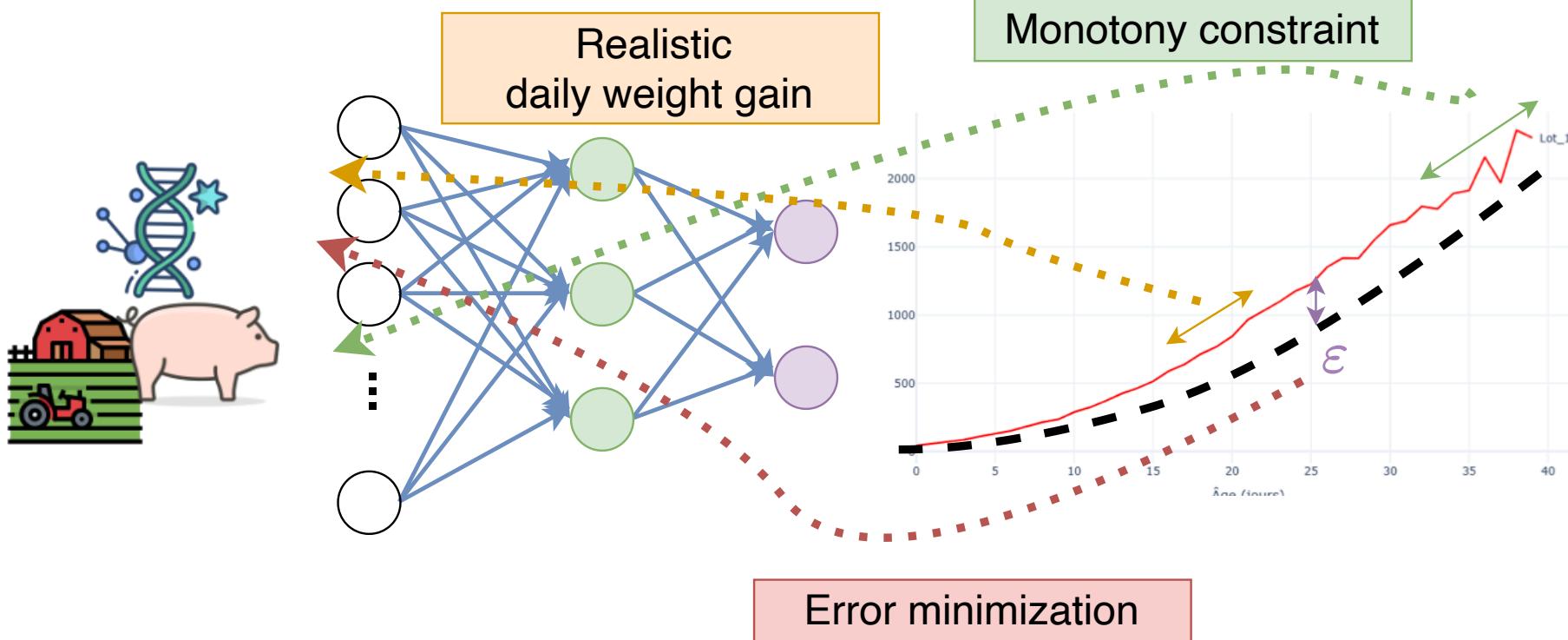
- Opérateur mathématique hyper-complexe
  - Gestion de données brutes riches (texte, image, ...)
- Système de correction (progressive) des erreurs



Verte  
Rouge  
Rouge

# Réseaux de neurones : architectures malléables

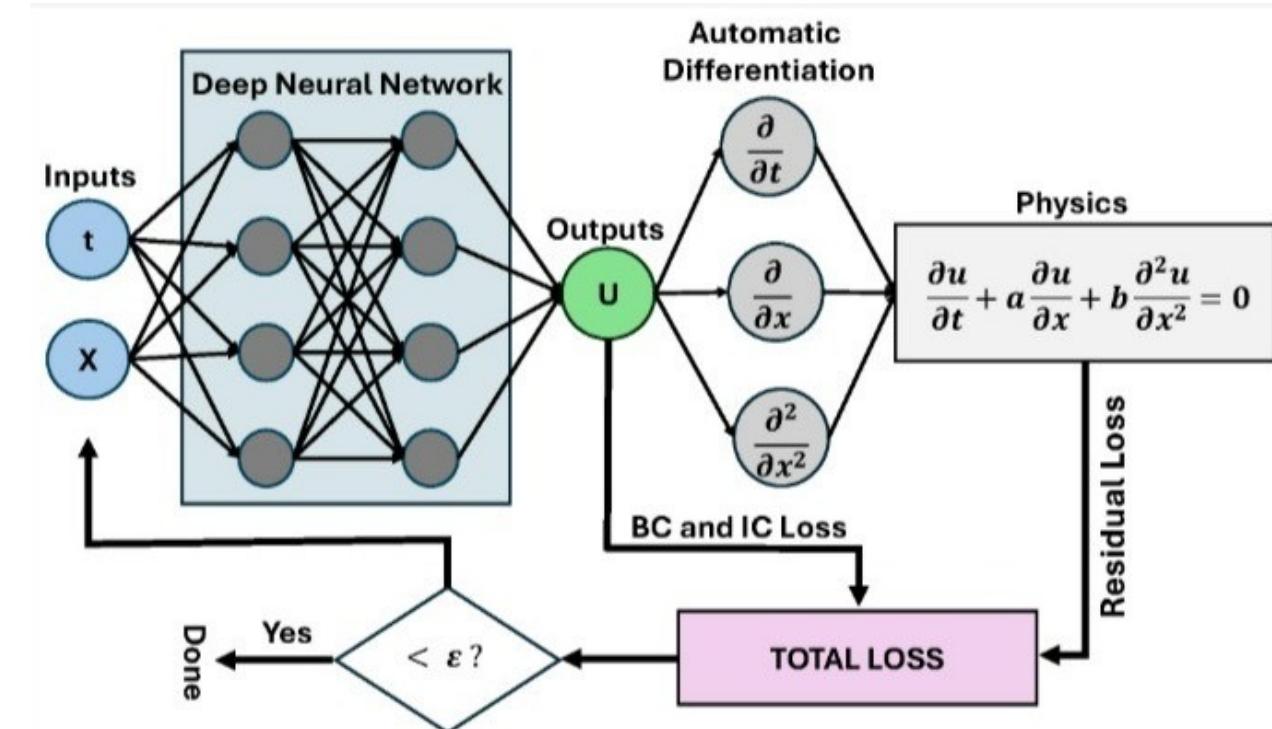
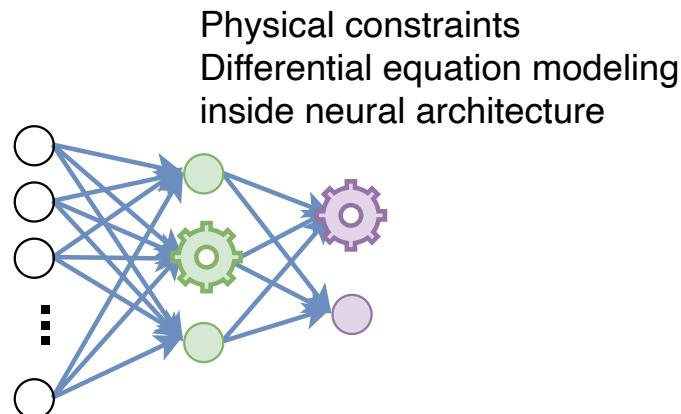
- Multiplier les contraintes = multiplier les objectifs (fonctions de coût)
  - Une autre manière d'introduire de l'expertise
- Design de l'architecture = garantie de certaines propriétés (e.g. monotonie)



# Hybridation des approches : data + mécaniste

- Pourquoi hybrider?
  - Gagner de la vitesse
  - Universalité
  - Exploitation des masses de données disponibles

- Physics Informed Neural Networks
  - Combiner EDP / réseaux de neurones
  - Beaucoup de théorie



# Hybridation des approches : data + mécaniste (2)

## □ Pourquoi hybrider?

- Gagner de la vitesse
- Universalité
- Exploitation des masses de données disponibles
- Limiter les entrées nécessaires

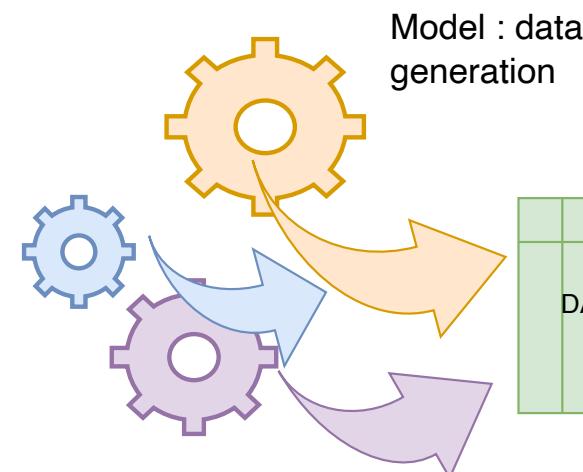
## □ Génération de données

- + Combinaison avec observations
- Calibration

## □ Unification de modèles

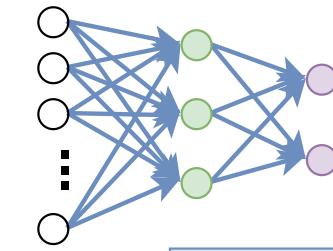
### Mechanistic model / simulation

Slow / costly  
Accurate  
Interpretable



### Data driven

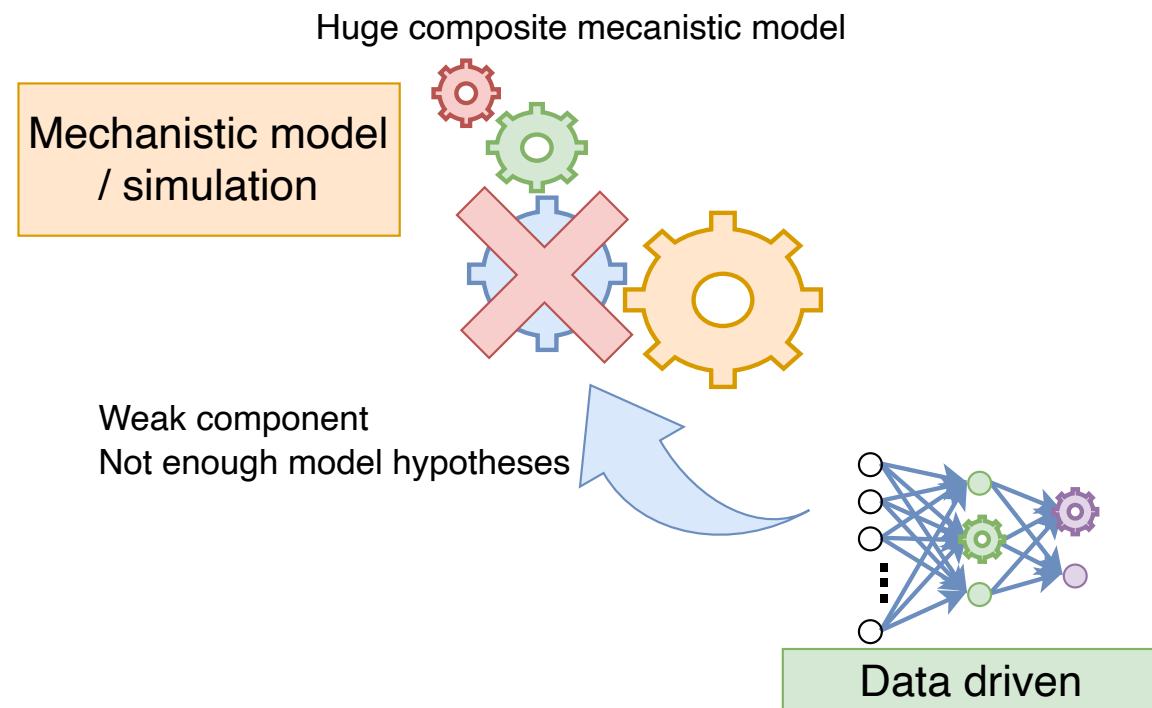
Fast  
Approximation



### Surrogate model

# Hybridation des approches : data + mécaniste (3)

- Pourquoi hybrider?
  - Gagner de la vitesse
  - Universalité
  - Exploitation des masses de données disponibles
- Système complexe
  - Remplacer le(s) maillon(s) faible(s)



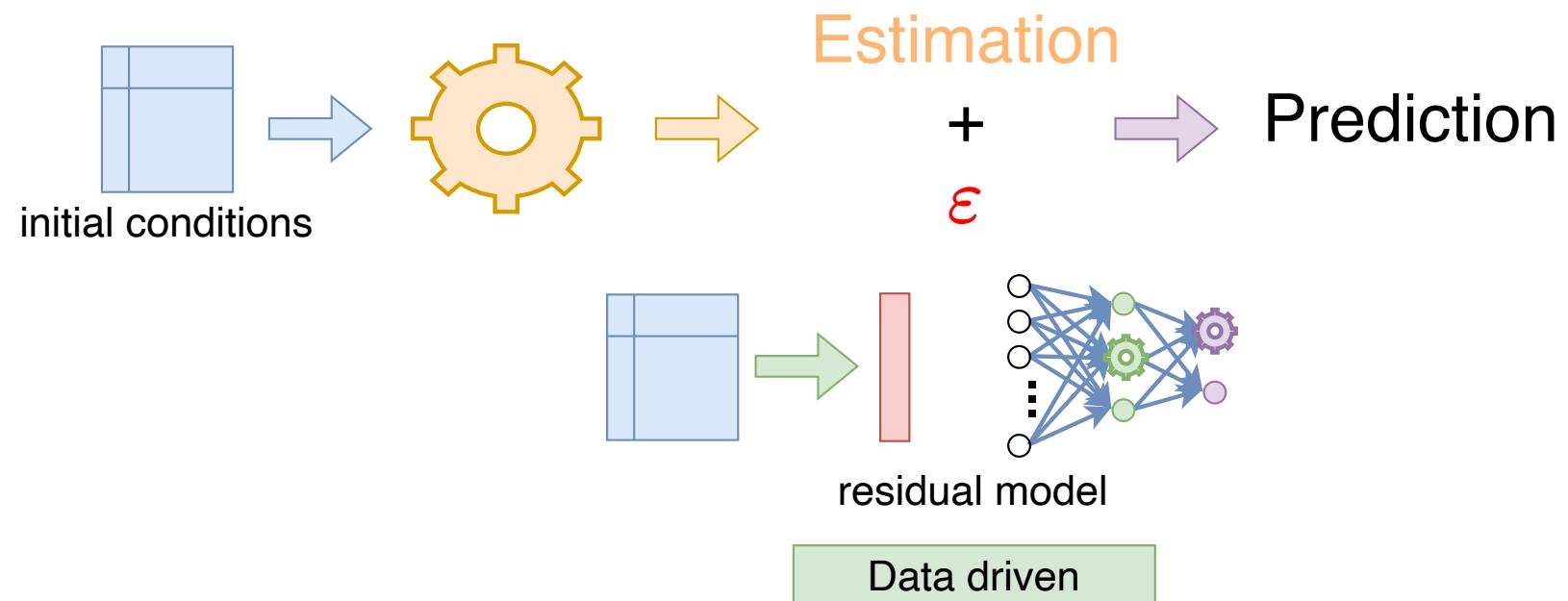
# Hybridation des approches : data + mécaniste (4)

## ❑ Pourquoi hybrider?

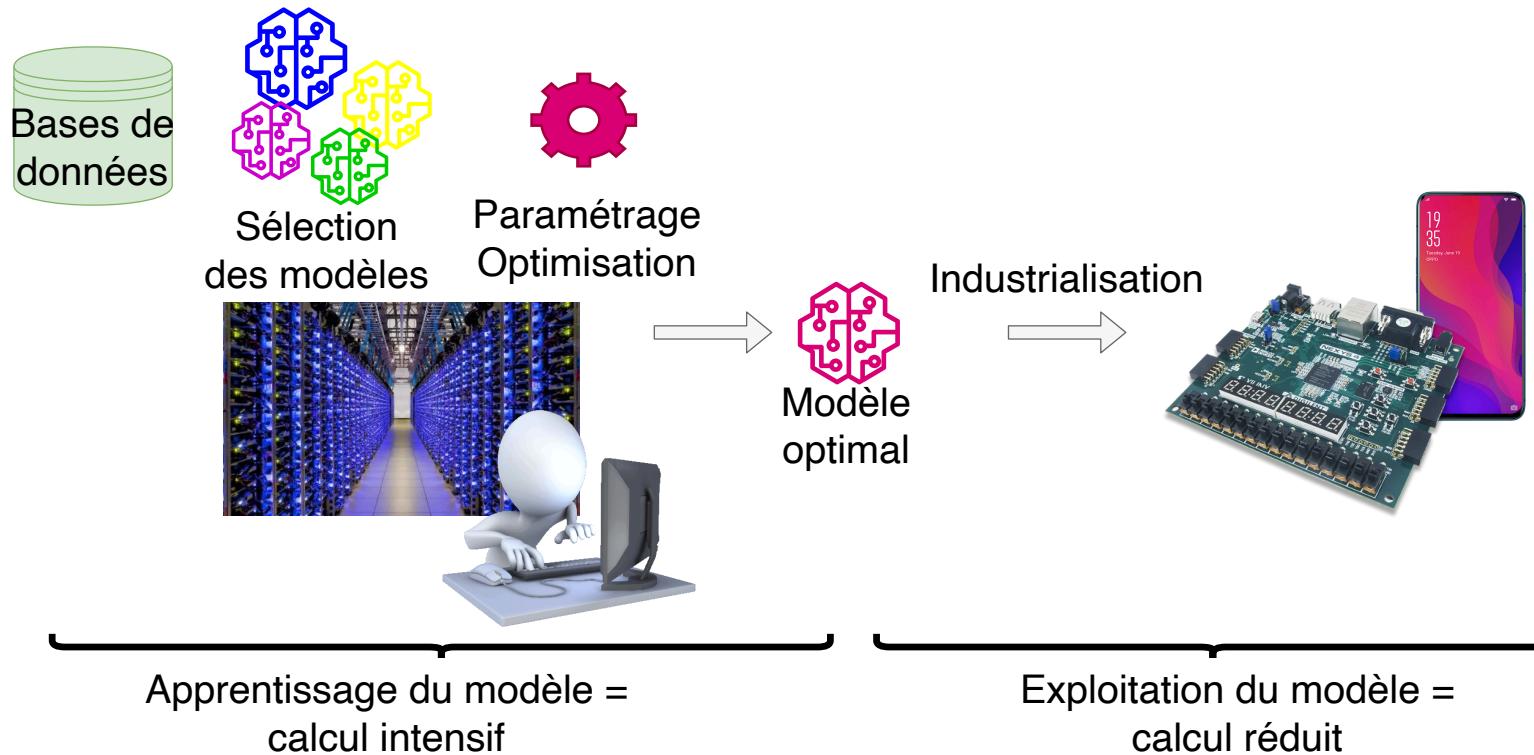
- Gagner de la vitesse
- Universalité
- Exploitation des masses de données disponibles

## ❑ Combinaison résiduelle

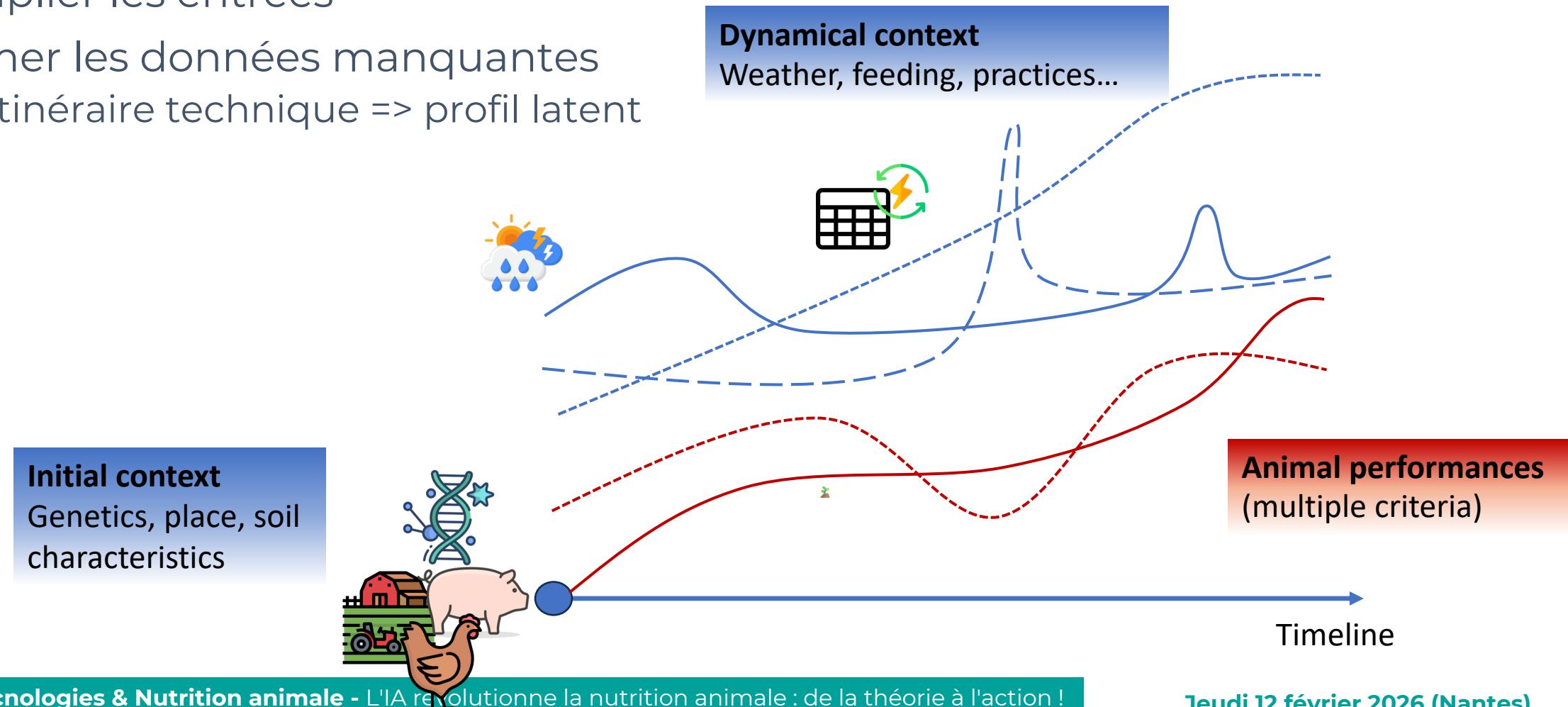
Mechanistic model / simulation



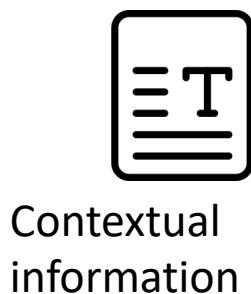
- Ne pas confondre entraînement et exploitation du modèle
- Le machine-learning est souvent plus léger que les approches mécanistes



- Prédire: Contexte (initial + dynamique) => performance animale
- Multiplier les entrées
- Estimer les données manquantes
  - Itinéraire technique => profil latent

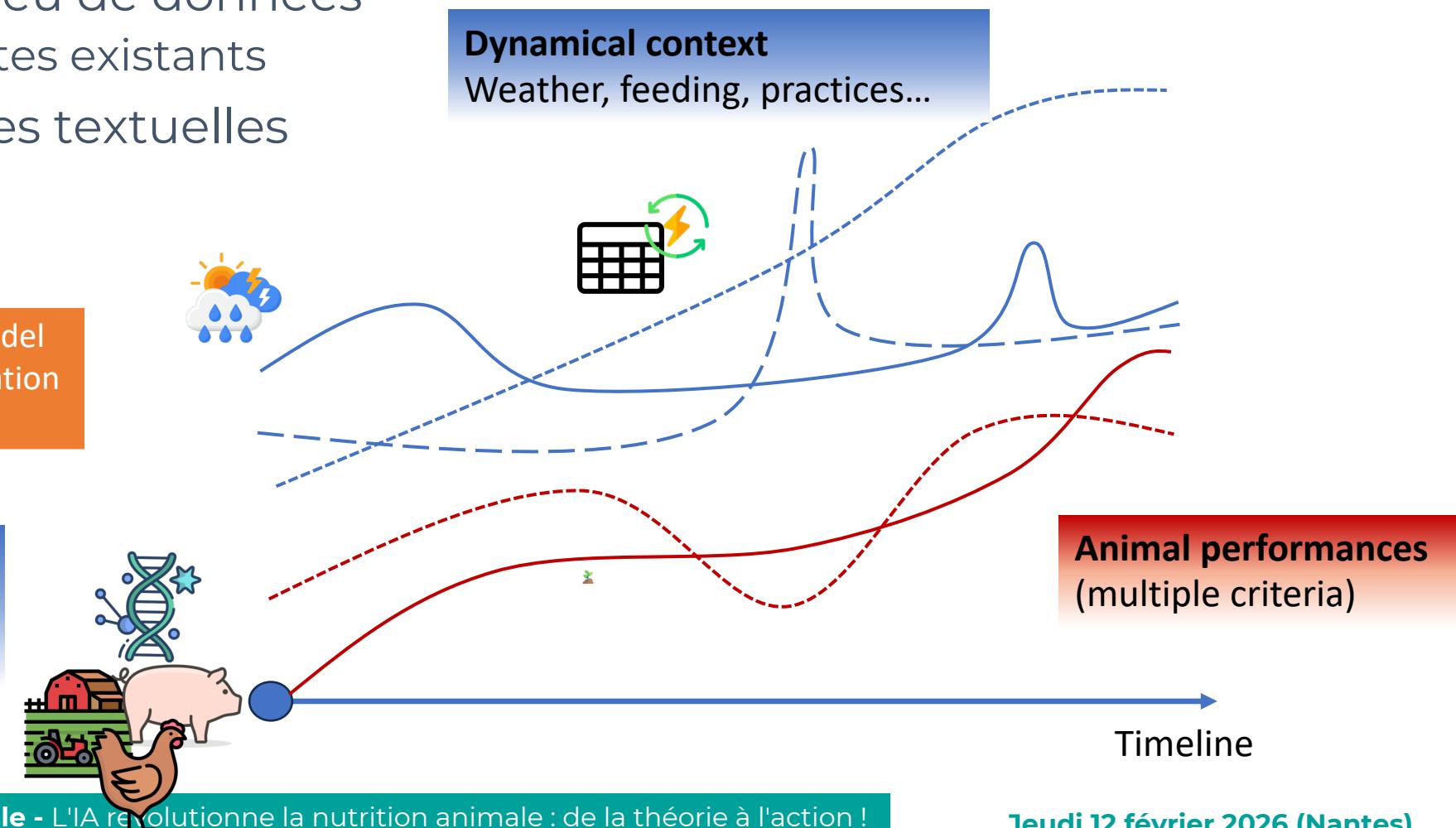


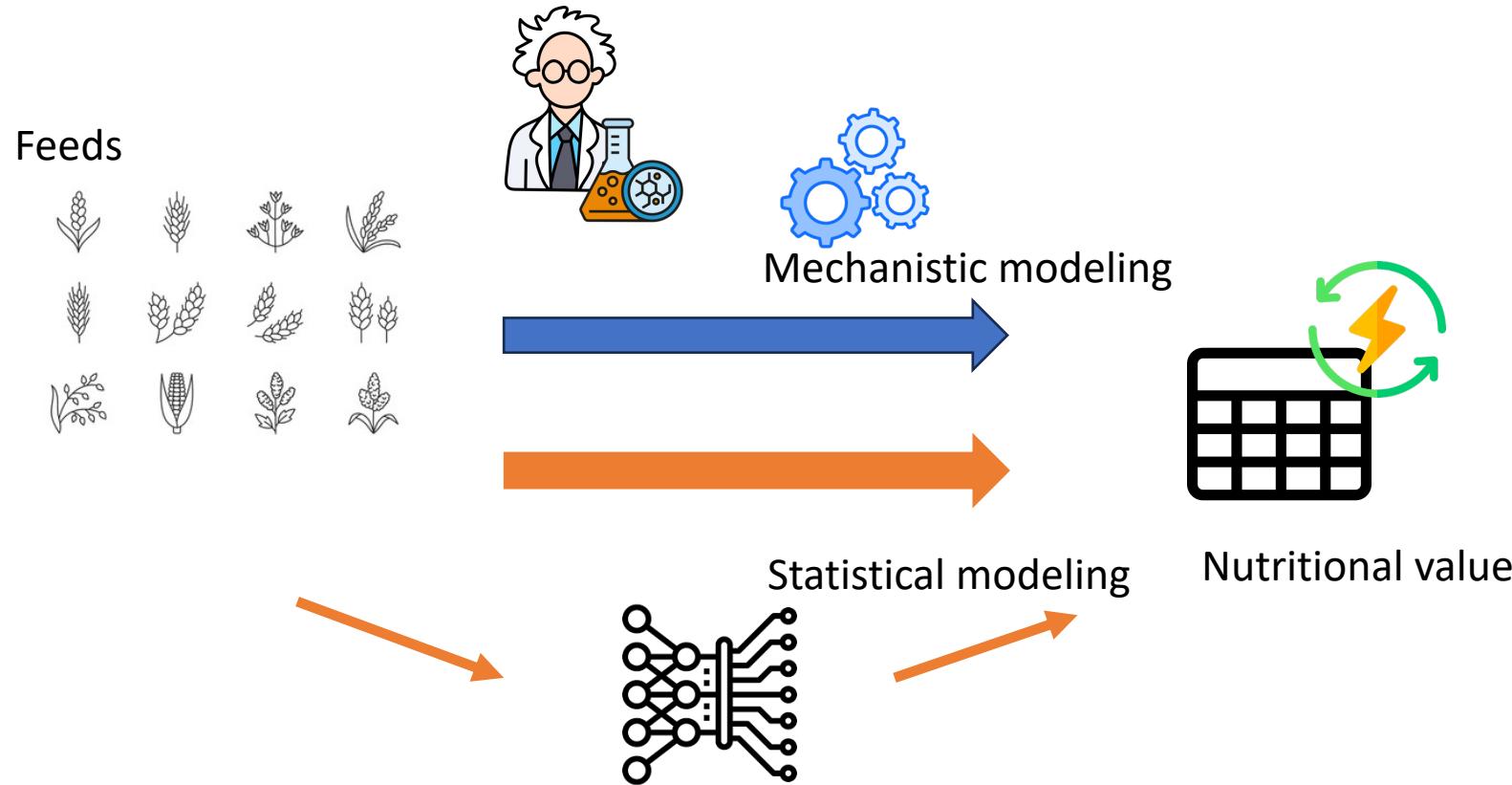
- Observer des variations pour en déduire l'impact sur les performances
- => Besoin de grand jeu de données
  - Modèles mécanistes existants
- Tirer parti de données textuelles descriptive

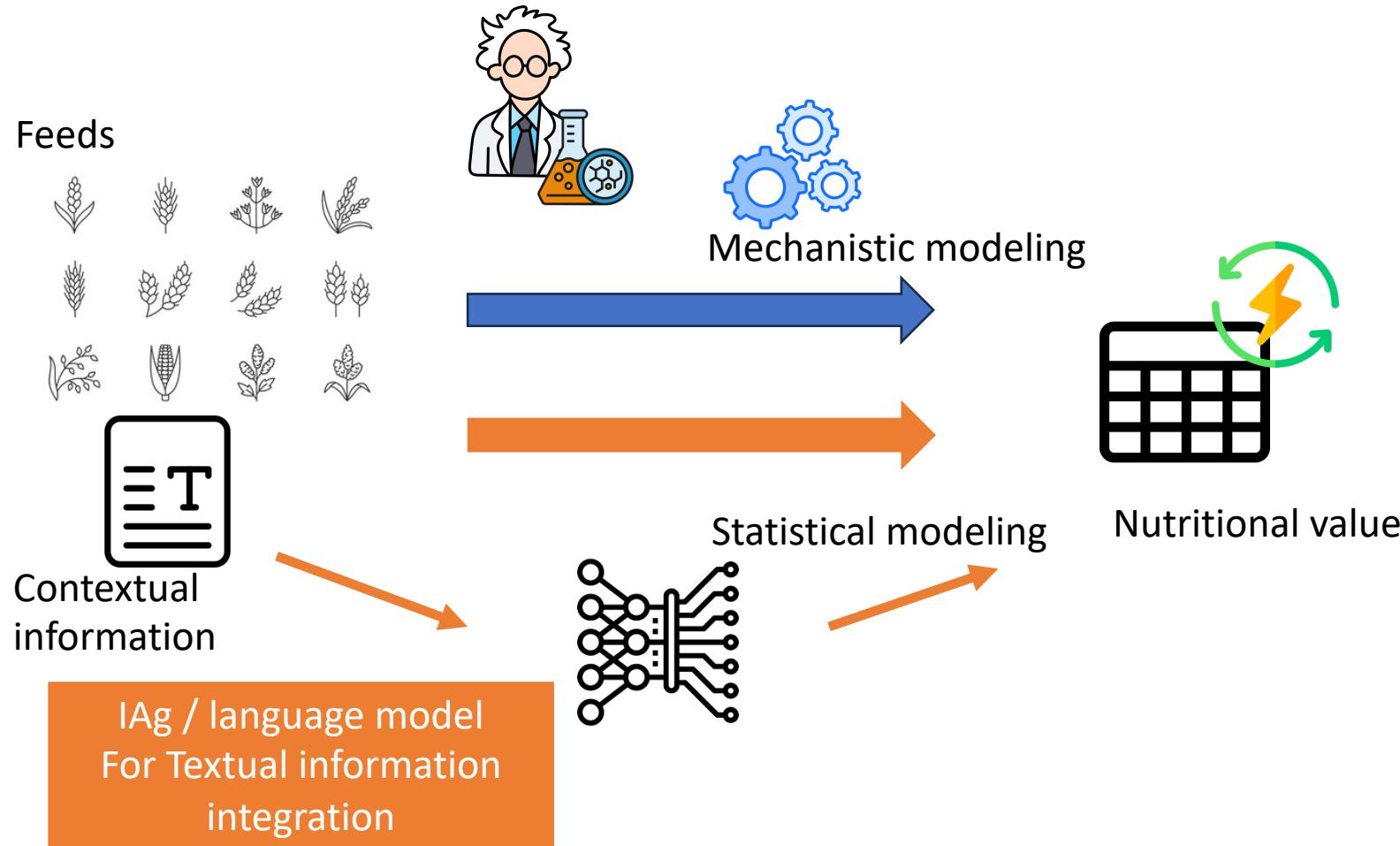


Contextual information

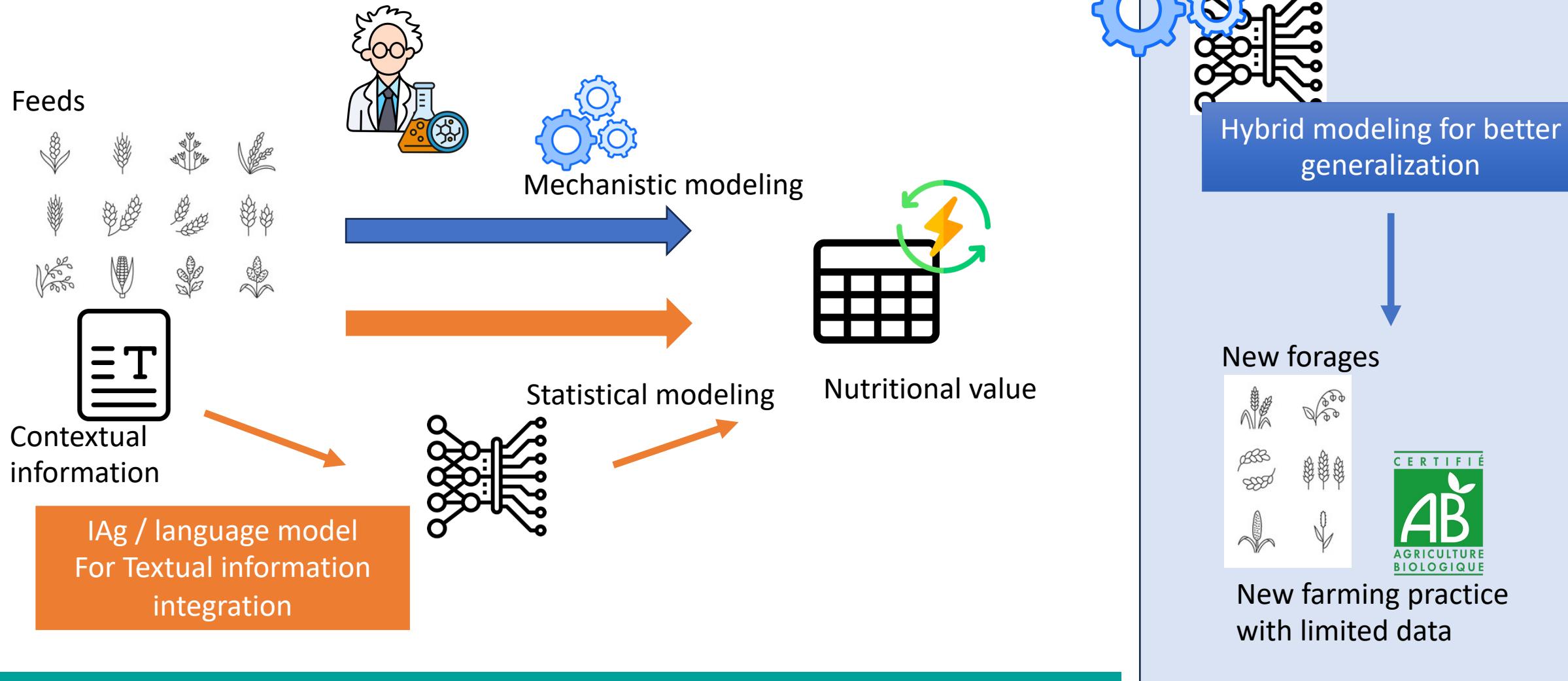
IAg / language model  
For Textual information  
integration





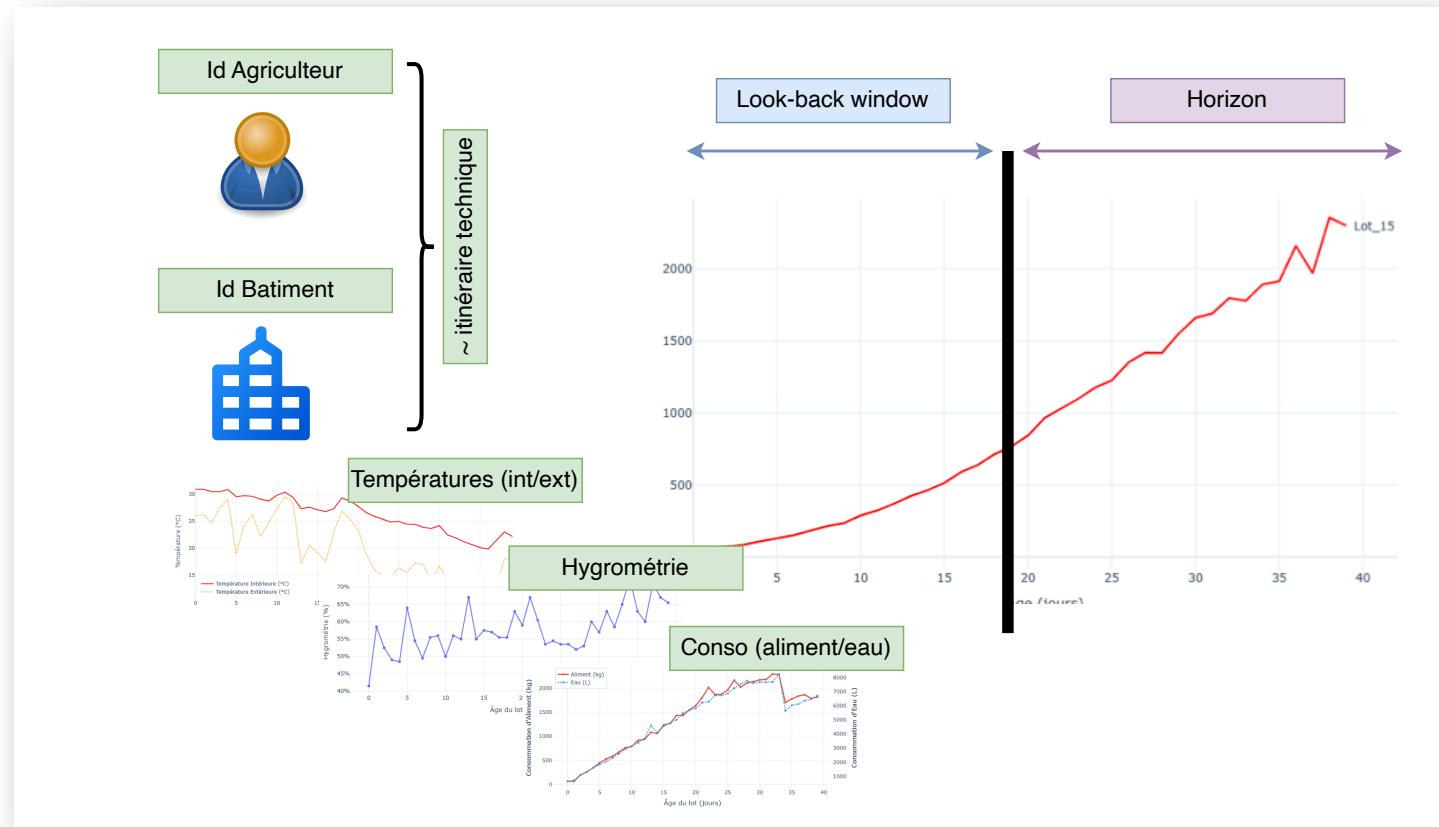


# AI4Diet : Remonter la chaîne vers l'alimentation



- Data = même modèle (+quelques variantes internes)
  - Différents jeux de données
- Mécaniste = Modèles complètement différents
  - Différents experts
- Applications :
  - Prédiction du poids des porcs, des poulets
  - Prédiction de la production de lait
  - Production de gaz dans un méthaniseur
- Mais à quelle échelle de prédition?
  - Individu / population
  - Heure, Jour, Semaine...

- Prédire = ne pas connaître le futur (contexte dynamique)
- Simuler = suivre un scénario pour le contexte futur



- Data: apprentissage sur des observations = un domaine
- Mécaniste : définition d'équations... Valables sur un domaine
- Que se passe-t-il lorsque l'on s'éloigne?
  - Nouvelle génétique
  - Réchauffement climatique
  - Aliments locaux mal caractérisés
  - Aliments exotiques

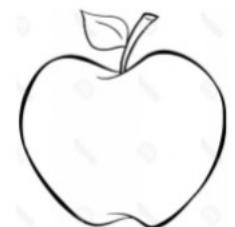
Photo ( train )



Clipart ( test )



Sketch ( test )



## => Augmenter le domaine de validité des modèles

- Combinaison de modèles
- Hypothèses sur les autres domaines
- Régularisation / Intégration de nouvelles données (e.g. textuelles)