Usage des données

Avant de déployer une solution technique, il est essentiel de définir précisément les données à collecter (leur nature, leur fréquence, leur objectif, le niveau de précision attendu…). Ces éléments qui émanent de la définition des besoins varient fortement selon les usages et orientent le choix des solutions techniques à tout niveau (capteurs, matériels, logiciels, interfaces, etc.). Il faut ensuite s'interroger sur la manière de capter cette donnée. Enfin, l'échantillonnage dépendra de l'utilisation finale de la donnée, notamment si elle sert à modéliser ou entraîner une IA.
L'exemple de l'outil Farm-e-val développé par CCPA illustre cette démarche. Les données à collecter au service de l'optimisation, des performances et du bien-être ont été sélectionnées pour répondre à 3 objectifs :

1. Piloter l'innovation en nutrition,
2. Accompagner le conseil aux éleveurs,
3. Comparer pour un conseil stratégique.

> Selon les cas, il pourra s'avérer nécessaire de s'assurer que les données pourront répondre à l'usage de modélisation ou d'entraînement de l'IA.

Par exemple si les données sont issues de capteurs, le parcours de choix et de validation des capteurs pourra être le suivant :

- Veille technologique, sélection du capteur,
- Tests et validation du capteur hors site et sur site pour validation des mesures et possibilité de traitement selon utilisation attendue, validation de la fiabilité de la remontée des données en situation réelle, validation de l'utilité par rapport aux objectifs fixés,
- Tests terrain avec utilisateurs pour valider l'usage, optimiser les réglages, le positionnement, la durée de vie des capteurs in situ,
- Intégration des données pour modélisation ou entraînement.

Outre les biais de mesure relatifs aux outils de collecte des données, il est également crucial d'anticiper les biais
pouvant affecter la qualité des données :

- **Biais de sélection** : s'assurer de la représentativité des données, population suffisamment large,
- **Biais de confirmation**: s'assurer que l'algorithme n'a pas été fait de manière à arriver au résultat auquel le développeur s'attend.
- **Biais d'observation** : s'assurer que la méthode de collecte n'influence pas les données,
- **Biais de rapportage** : s'assurer que les observations ne sont pas rapportées de manière disproportionnée,
- **Biais d'exclusion** : s'assurer que certaines caractéristiques ne sont pas sous-estimées.

Révision #1
Créé 1 juin 2026 14:40:03 par Céline Ravel
Mis à jour 1 juin 2026 14:46:14 par Céline Ravel