



Webinaire

La qualité des données :
un facteur clé d'efficience
des dispositifs de surveillance



15h - 17h





Pourquoi un webinar sur la qualité des données issues de la surveillance ?

- Continuité des travaux du **groupe de travail inter-plateformes** sur la qualité des données de surveillance
- Composé d'**expert·es** ACTA-les instituts techniques agricoles, Anses, CNIEL, INRAE et OQUALIM
- **Mutualiser** les expériences et les bonnes pratiques pour **garantir, améliorer et suivre** la qualité des données de surveillance



Epidémiosurveillance en Santé Animale
Epidémiosurveillance en Santé Végétale
Surveillance de la Chaîne Alimentaire



Objectif : l'amélioration de l'**efficience** de la surveillance

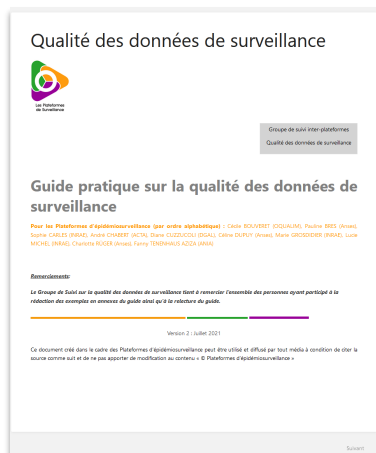


Qualité des données : enjeu fort pour la surveillance

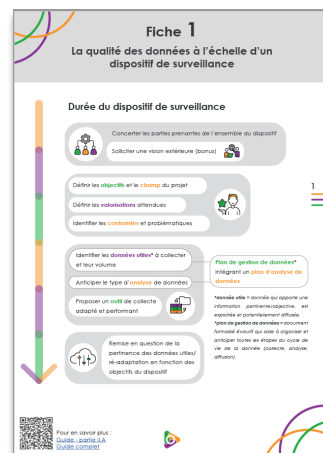


Les livrables du groupe Qualité des données inter-plateformes

- Un guide pratique (avec exemples des différents secteurs)
- Des fiches synthétiques



Accéder au guide complet



Accéder aux fiches synthétiques

- Fruit d'un travail collaboratif

Membres du groupe

- Cécile BOUVERET
- Pauline BRES
- Sophie CARLES
- André CHABERT
- Diane CUZZUCOLI
- Céline DUPUY
- Marie GROSIDIER
- Lucie MICHEL
- Charlotte RUGER
- Fanny TENENHAUS-AZIZA

○ + contributeur.rices aux exemples





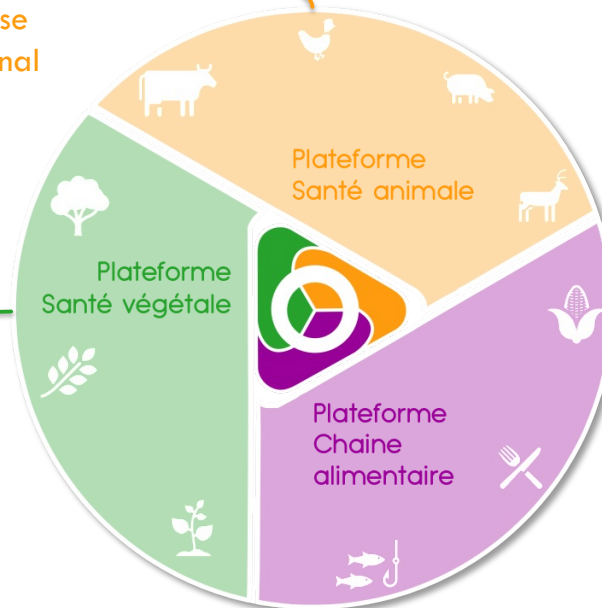
Un webinaire structuré autour de trois cas concrets



Surveillance de la tuberculose
en abattoir : Stéphanie Darnal



Surveillance du nématode
du pin : Marie Grosdidier



Surveillance du Cadmium :
Margot Bärenstrauch





CONTEXTE DE TRAVAIL DES TROIS PLATEFORMES





Pourquoi surveiller ?



Surveillance : “Collecte systématique et continue, compilation et analyse des données et diffusion dans des délais appropriés des informations à ceux qui ont besoin de savoir de manière à mettre en œuvre des actions” (J Last, 2001)

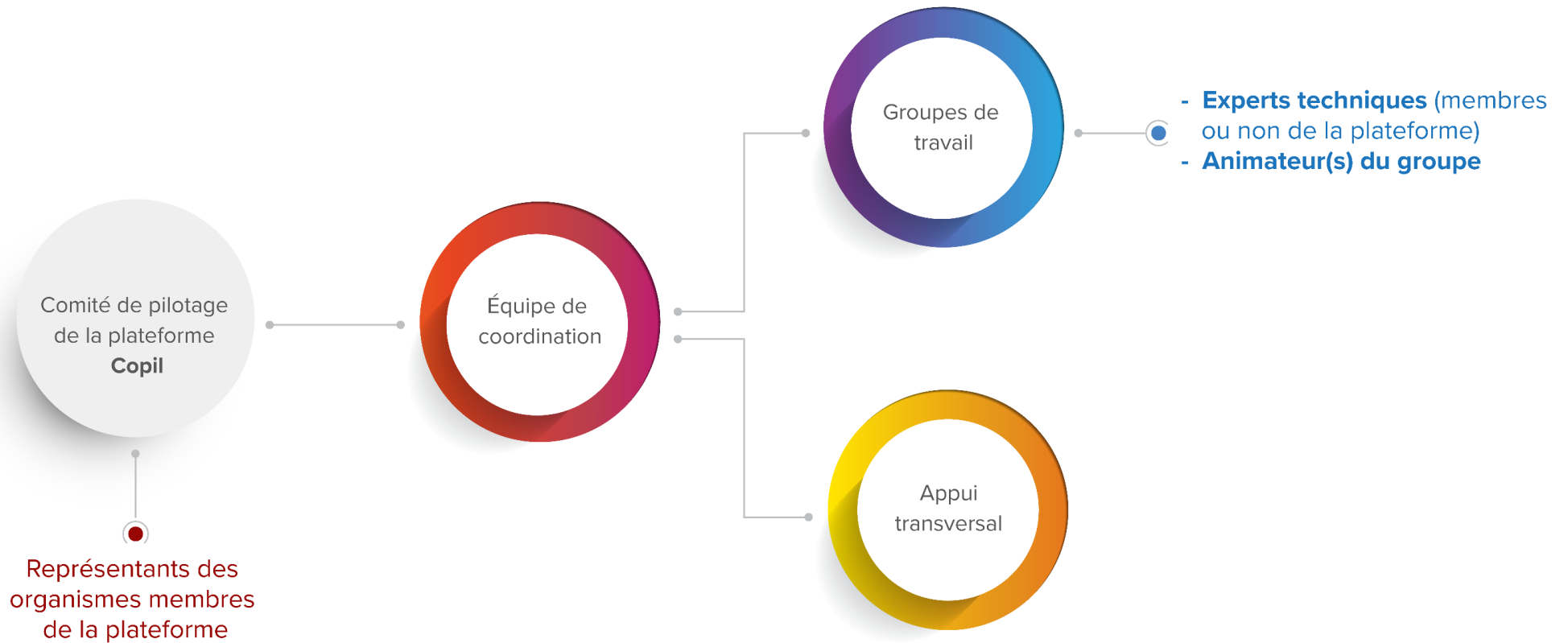


Détails sur le livre

Surveiller pour agir



Trois Plateformes, une typologie d'organisation



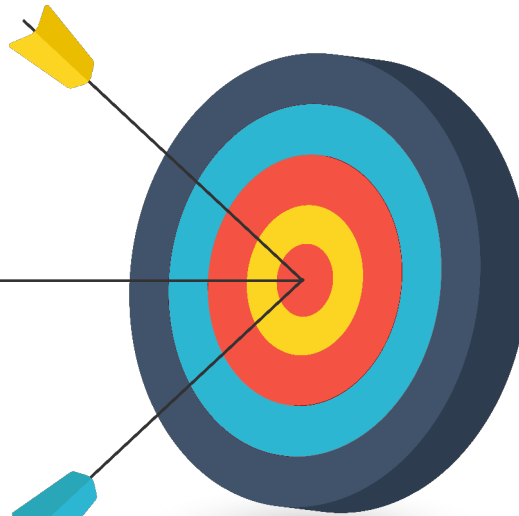


Objectifs et missions des Plateformes

**Développer, adapter
et promouvoir** des dispositifs
de surveillance

Exercer une veille
sanitaire nationale
et internationale

Réaliser des synthèses
sur la situation épidémiologique
des dangers sanitaires



**Améliorer l'efficacité
de la surveillance**
Surveiller pour agir



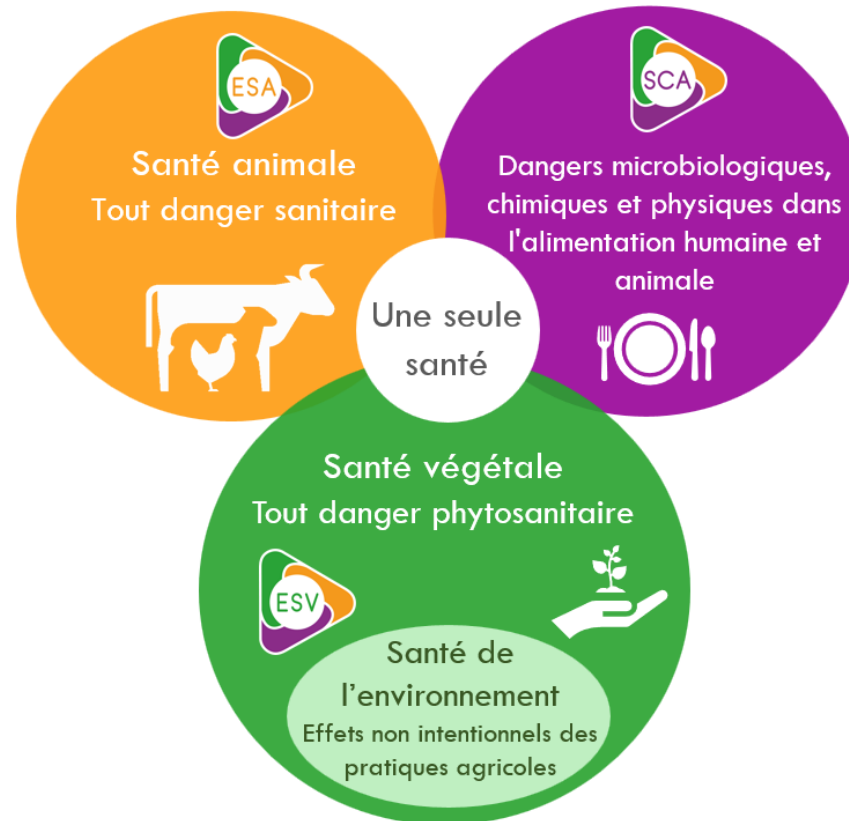
Appui méthodologique et
opérationnel aux dispositifs de
surveillance (privés/publics)



Non décisionnaires
pour les dispositifs de
surveillance



Domaine des plateformes



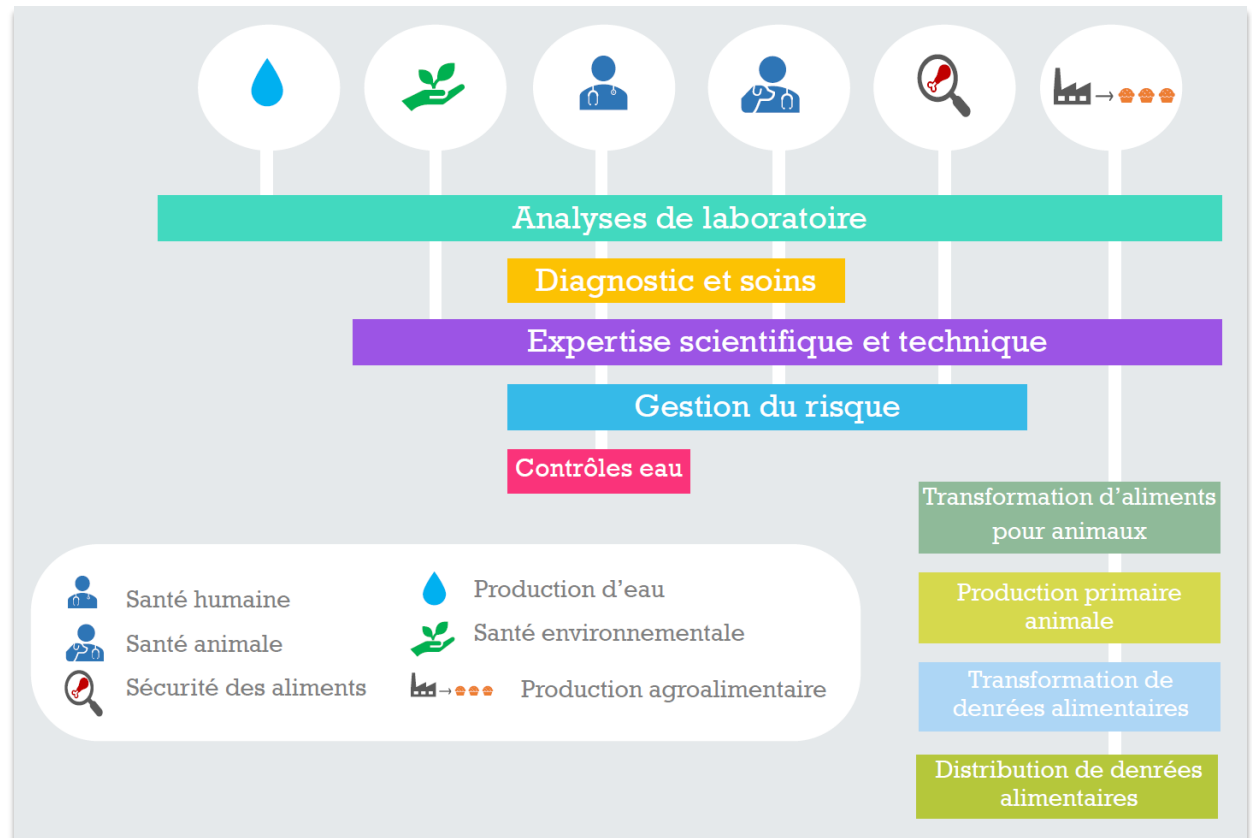


Une surveillance multisectorielle et multidisciplinaire

Des bénéfices directs de la surveillance pour les différents acteurs

☐ La qualité des données :

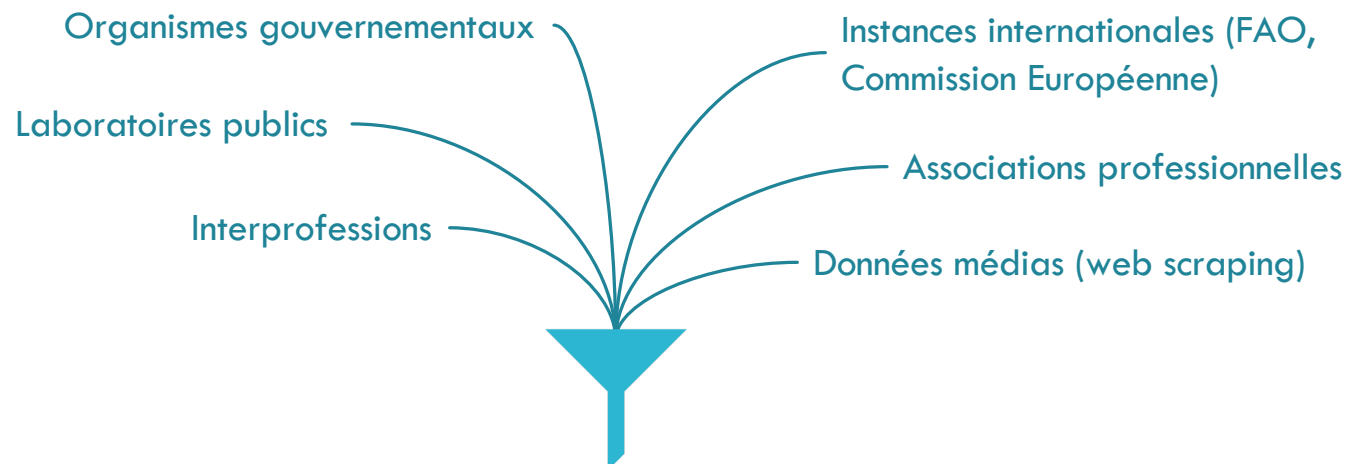
un enjeu pour tous (ex. système de surveillance des salmonelles)





D'où viennent les données des plateformes ?

- En règle générale, les plateformes **ne collectent pas** les données directement
- Fournisseurs de données **gestionnaires des dispositifs de surveillance** faisant appel à un appui des plateformes





Un des enjeux de la qualité des données de surveillance pour les plateformes

- Leur valorisation pour un appui aux politiques publiques

Participe à une prise de décision éclairée par les autorités compétentes nationales :

- Proposition **d'évolution des modalités** de surveillance en fonction de la situation sanitaire (**Sylvatub, Tuberculose bovine**)
- **Mise en place** de dispositifs de surveillance (***Xylella fastidiosa***)
- **Construction d'outils** et de **recommandations** pour améliorer la surveillance (**Cadmium**)

Permet aux autorités compétentes de suivre des indicateurs et transmettre des données aux instances internationales (**la tuberculose bovine (EFSA)**, **la SORE : Surveillance des Organismes Réglementés et Émergents (Commission européenne)**, **GT Abstract: remontée des résultats d'autocontrôles Salmonella sur carcasses bovins, ovins, caprins et volailles vers la DGAL et l'EFSA**)



Quelles plus-values des plateformes pour la qualité des données ?

- Appui pour le suivi de la qualité des données
 - Développement d'outil dédié (Qualiplan)



- Appui au développement ou évolution d'un dispositif
 - Développement d'outil de collecte (OMAA)



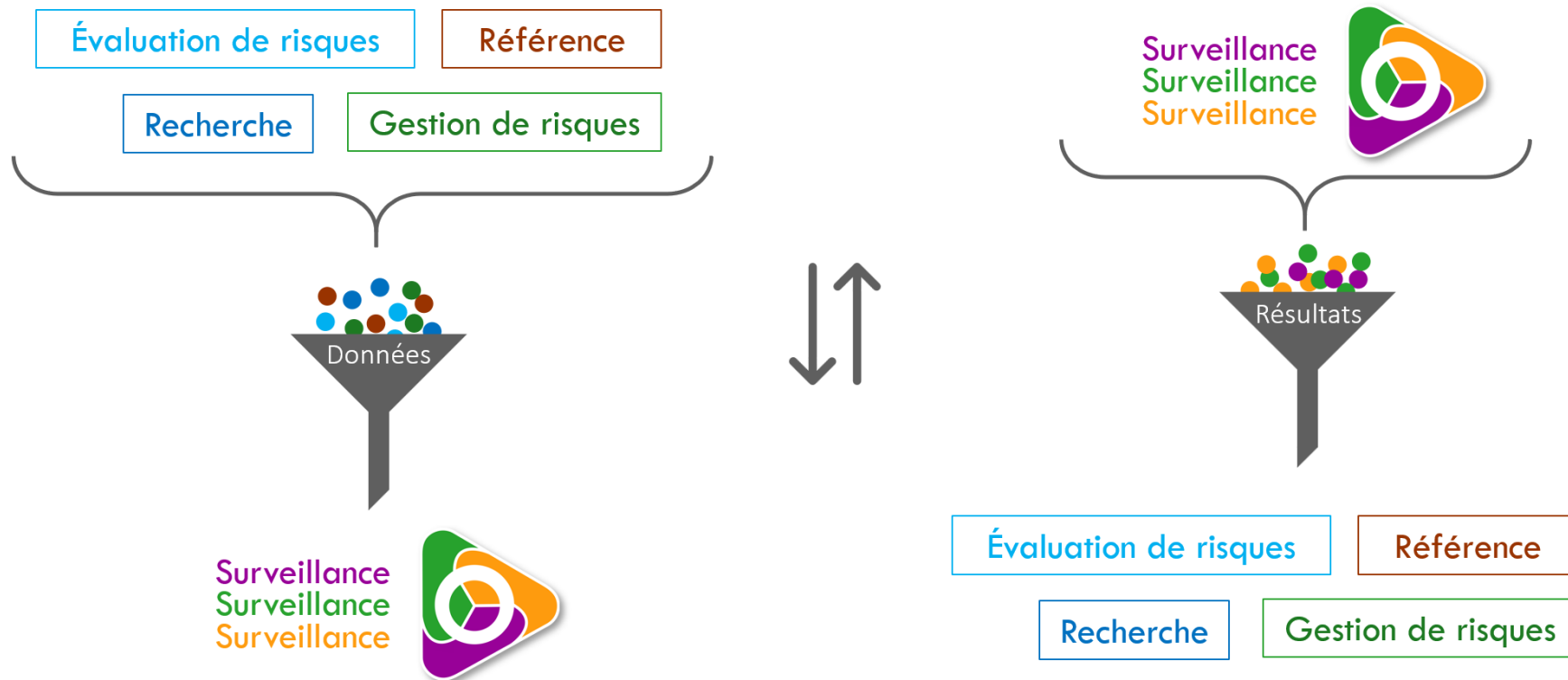
Appui pour l'amélioration en continu de la qualité des données

- + Plus-value expertise multidisciplinaire
- + Plus-value équipes dédiées aux Plateformes avec expertise métier relative aux traitements des données de surveillance





Un enjeu de la qualité des données au-delà de la surveillance





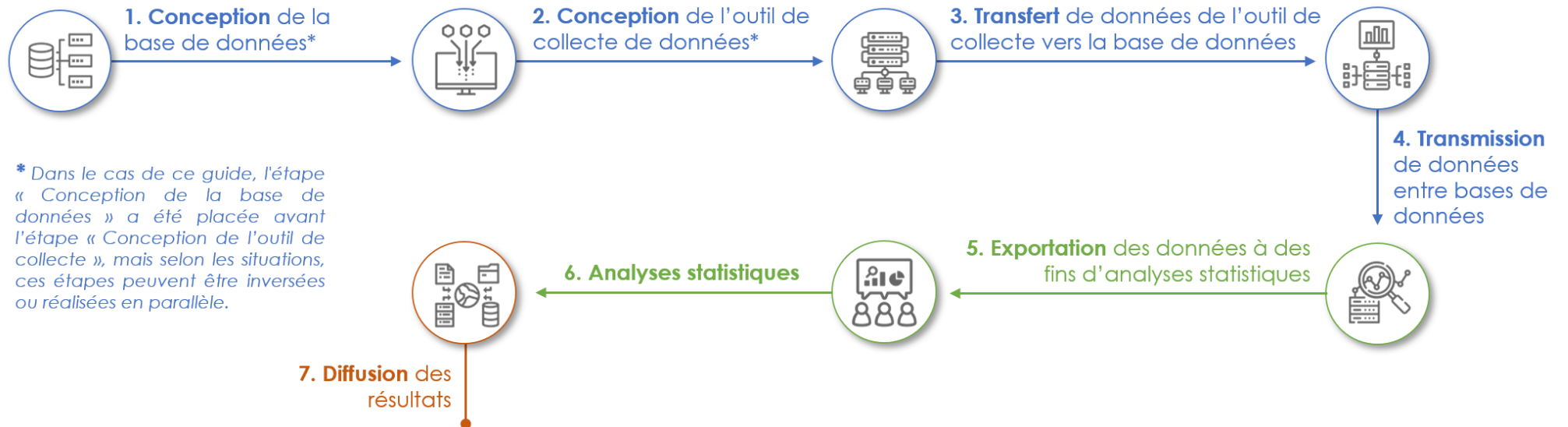
PRÉSENTATION DES CAS CONCRETS





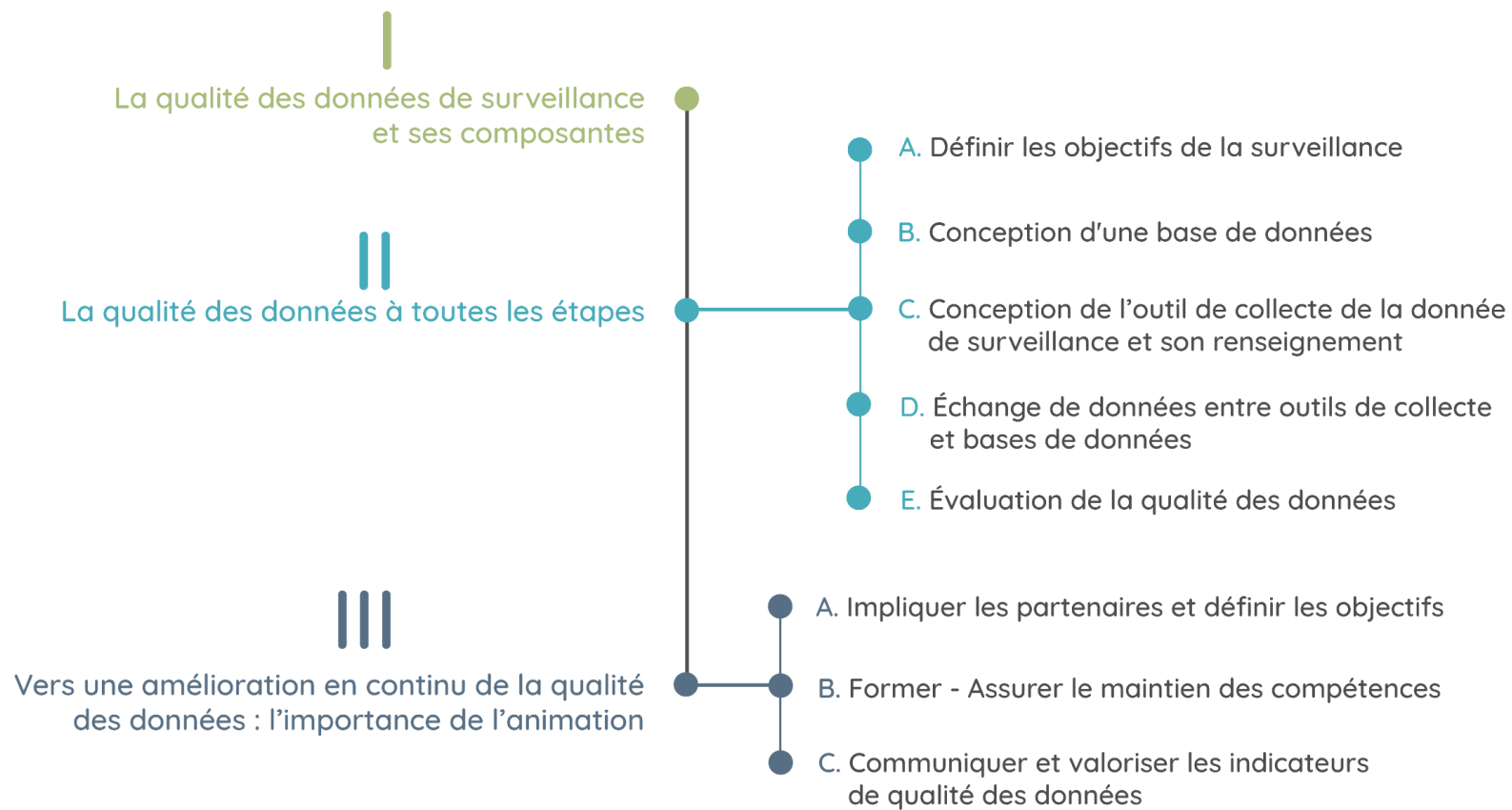
Cycle de vie des données

Collecte → Analyse → Diffusion



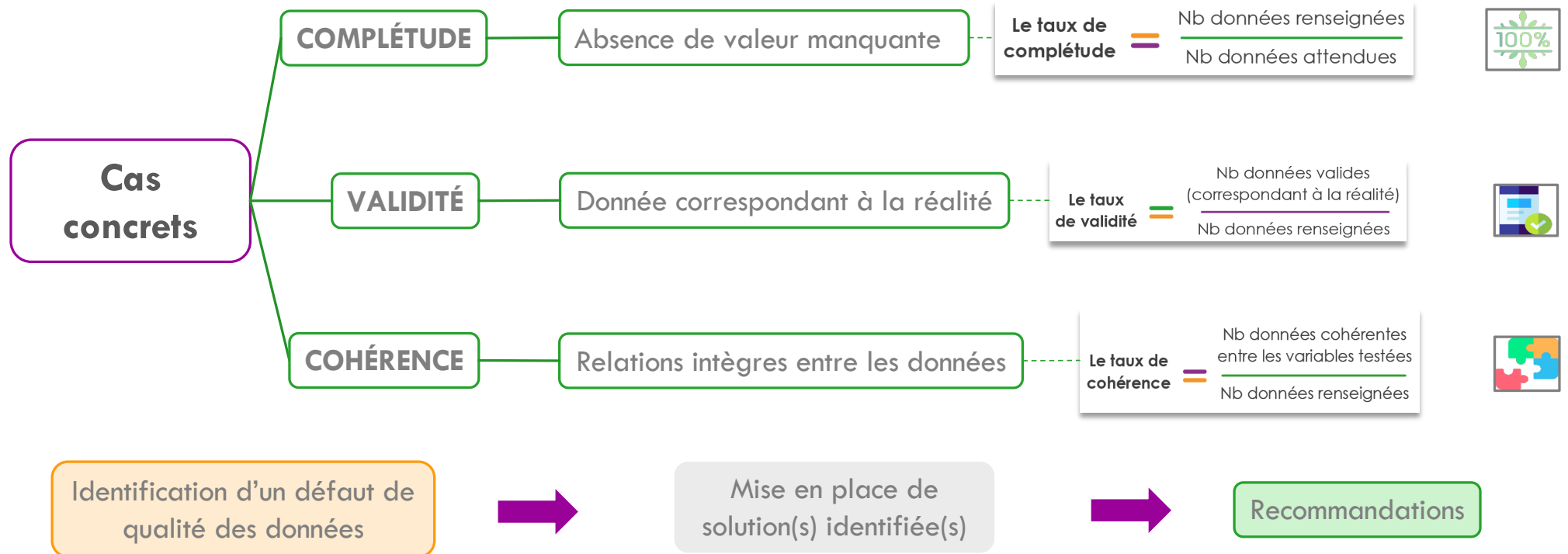


Programme du guide qualité des données



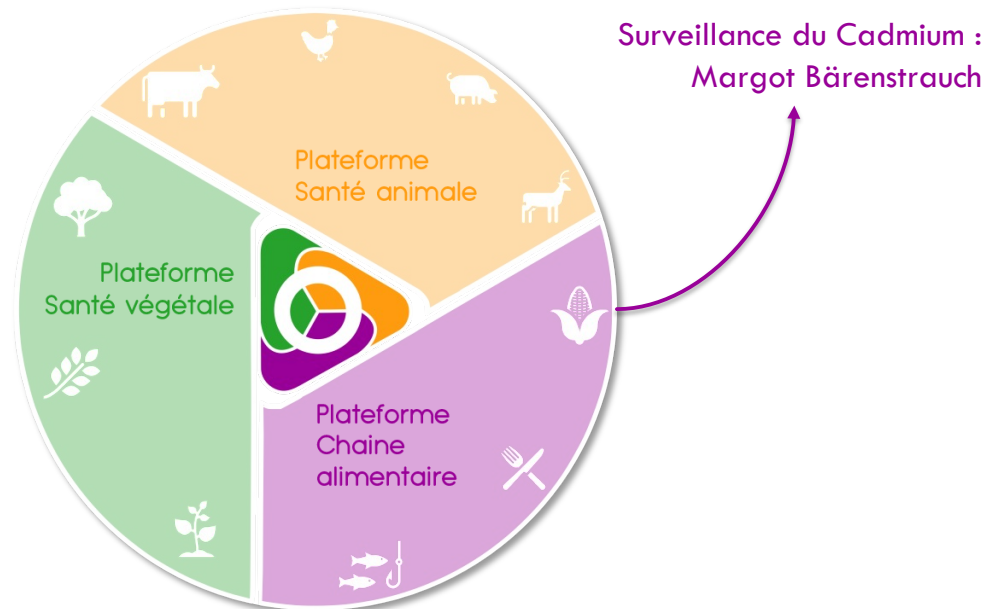


Présentation des trois cas concrets





Surveillance de la chaine alimentaire





Margot Bärenstrauch



- Ingénieure de recherche au LABERCA (Laboratoires des résidus et contaminants dans l'alimentation) à Nantes
- Membre opérationnel de la plateforme en lien avec la surveillance des **dangers chimiques**
- Co-rédactrice du BuSCA : [Abonnez-vous !](#)

Plus de détails sur les travaux du groupe ici !





Site de la Plateforme SCA,
page GT Cadmium





Exemple en lien avec les données cadmium

CONTEXTE : Le cadmium

- Élément trace métallique, **ubiquitaire**
- Sources :
 - présence naturelle dans les sols
 - activités humaines (industries, engrais phosphatés...)
- Responsable de pathologies osseuses, rénales
- **Réglementé** :
 - dans l'alimentation humaine (CE n°2023/915) 
 - dans l'alimentation animale (CE n°2002/32) 
- **Préoccupation sanitaire forte en France** :
 - Dépassements de la dose hebdomadaire tolérable pour une partie de la population française (Anses)
 - Augmentation des concentrations en Cd pour plusieurs catégories d'aliments entre EAT1 (2004) et EAT2 (2011)



La surveillance peut-elle être améliorée ?

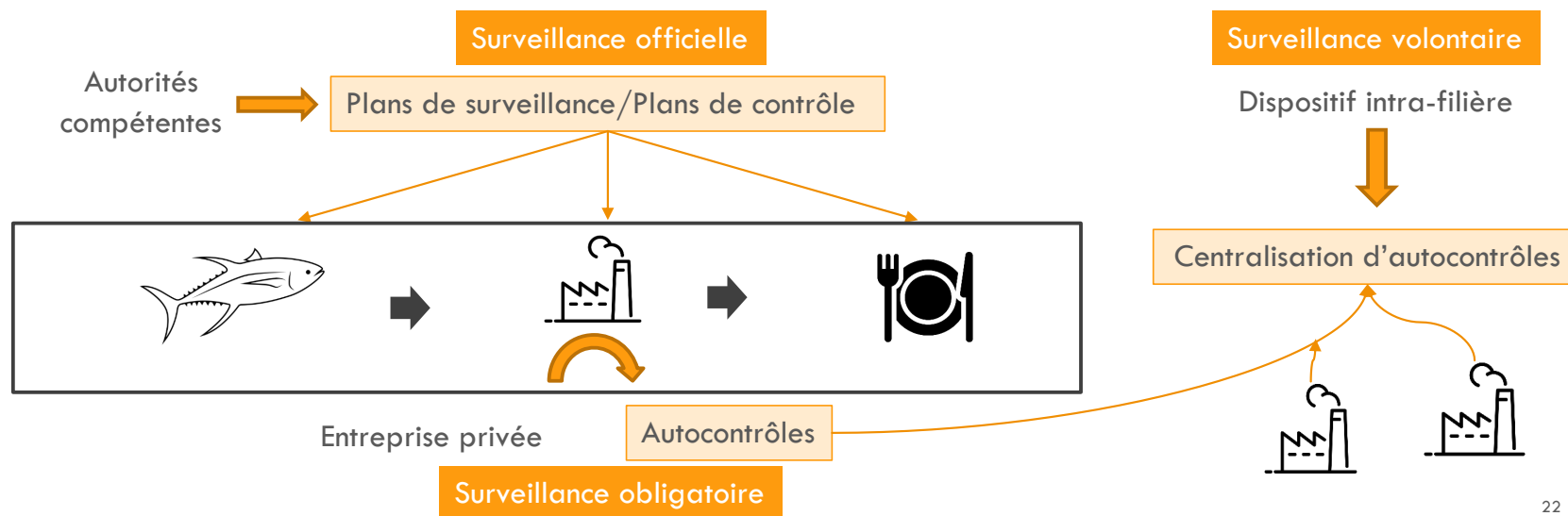
GT Cadmium (2020-2022)



CONTEXTE : La surveillance du cadmium

- **Multiplicité d'acteurs**, de l'amont à l'aval
- Différents cadres d'application : **surveillance officielle – obligatoire - volontaire**
- **19 dispositifs répertoriés** (non exhaustif)
- **Objectifs divers** : vérifier la conformité des aliments, mieux cibler les autocontrôles, participer à la fixation de seuils réglementaires réalistes, suivre des tendances géographiques/temporelles etc..

Exemples :



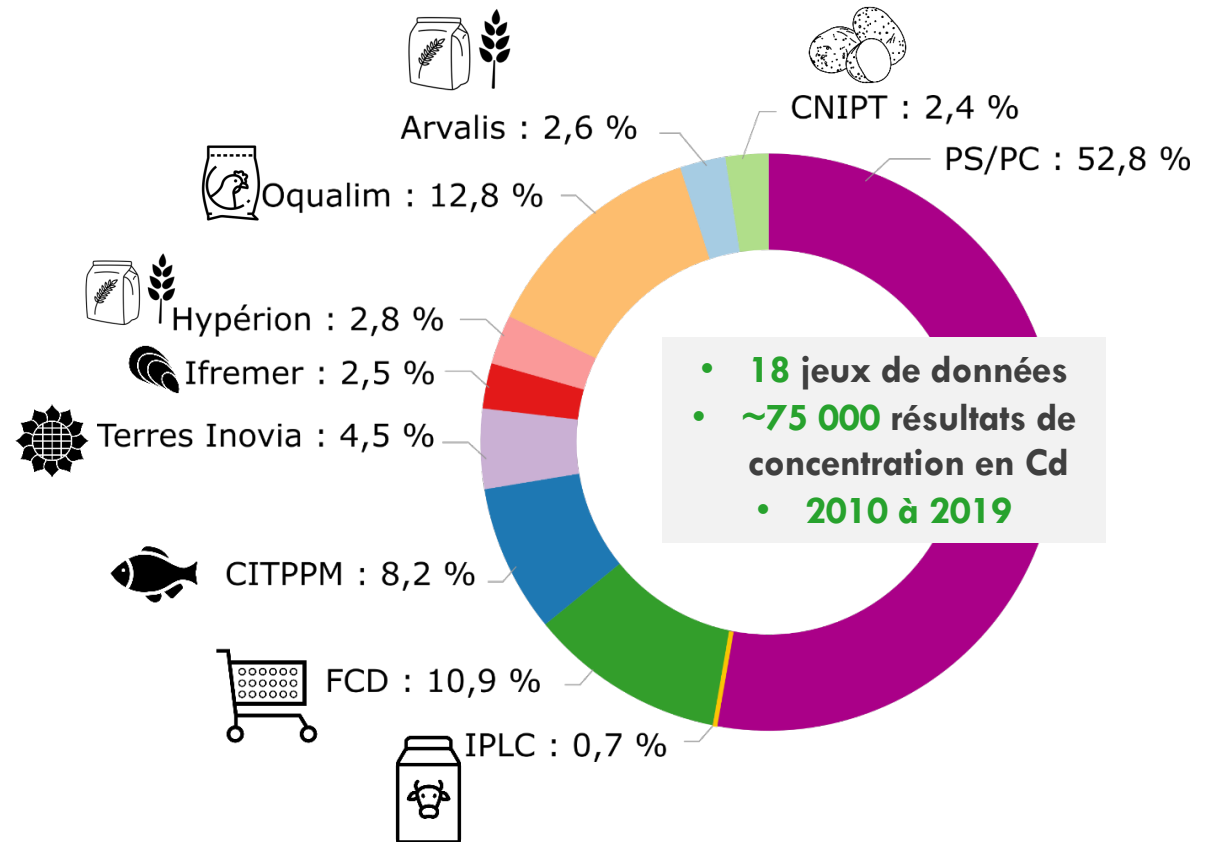


CONTEXTE : Objectifs du GT Cadmium

> Évaluer la surveillance du cadmium en France

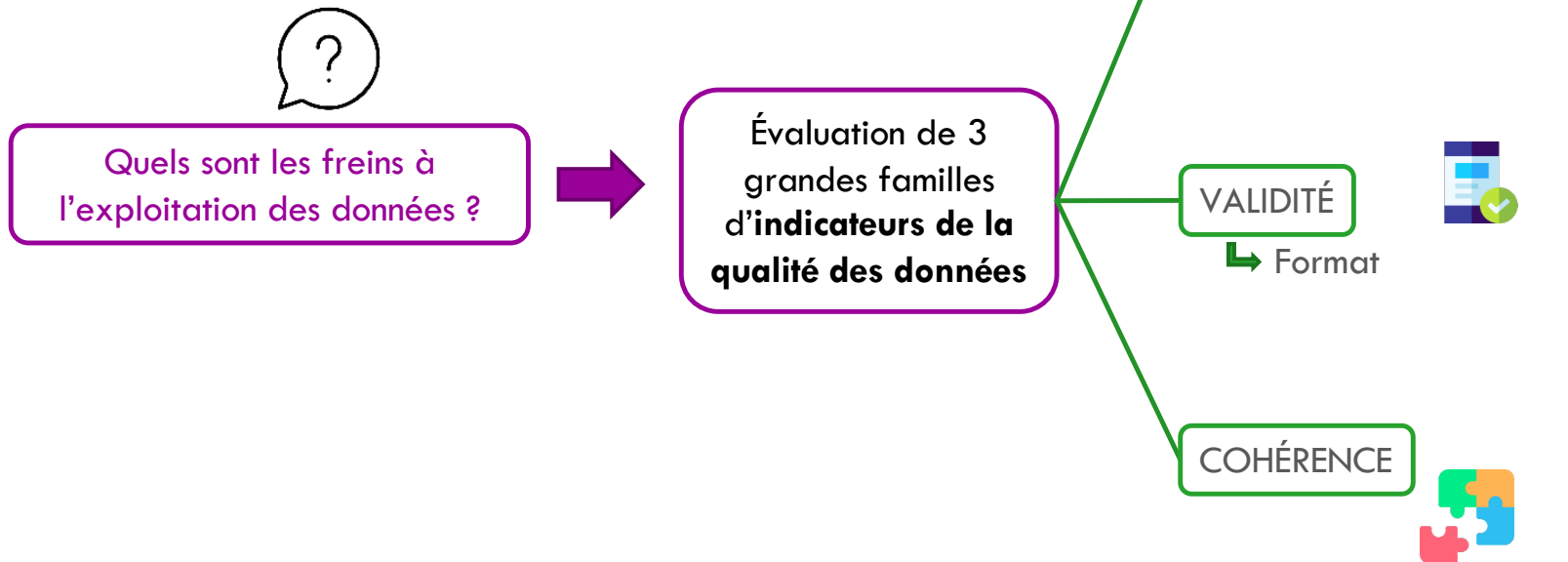
- ↳ **Mutualiser des données** de surveillance privées, publiques et issues de différentes filières (concentrations en cadmium)
- ↳ **Évaluer la qualité** des données
- ↳ **Analyser les niveaux de contamination** entre 2010-2019 (Tendances ? Matrices à cibler ?)

> Proposer des recommandations





Méthodologie





Données utiles : méthode

COMPLÉTUDE



- Travail collaboratif pour définir une liste de données indispensables ou utiles pour la surveillance du cadmium

1

Identifier les objectifs d'un dispositif « idéal »

Ex: Production primaire végétale
-Comparer les concentrations en cadmium en fonction de la zone de production

2

Identifier les données utiles

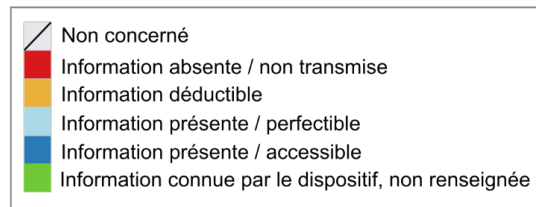
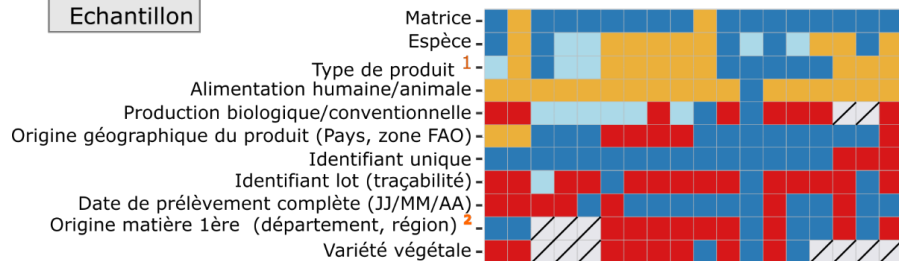
Ex: Origine géographique de production

3

Identifier les freins à la collecte : données sensibles ou difficiles à collecter

Donnée sensible

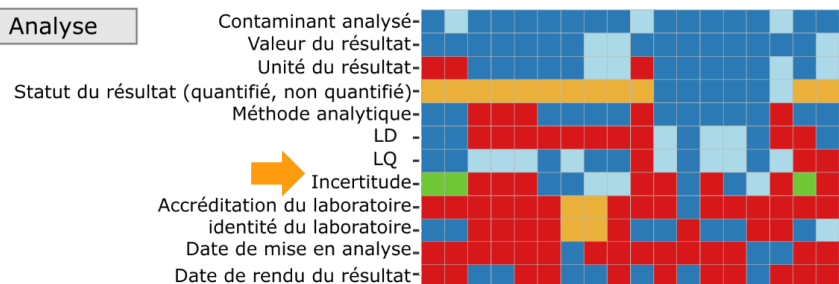
Echantillon



COMPLÉTUDE



Analyse



Données utiles : résultats

- Les informations les plus essentielles sont renseignées (contaminant, unité, LOQ, etc.)
- Certaines données rarement renseignées comme la **conformité réglementaire** et l'**incertitude de mesure**

Réglementation



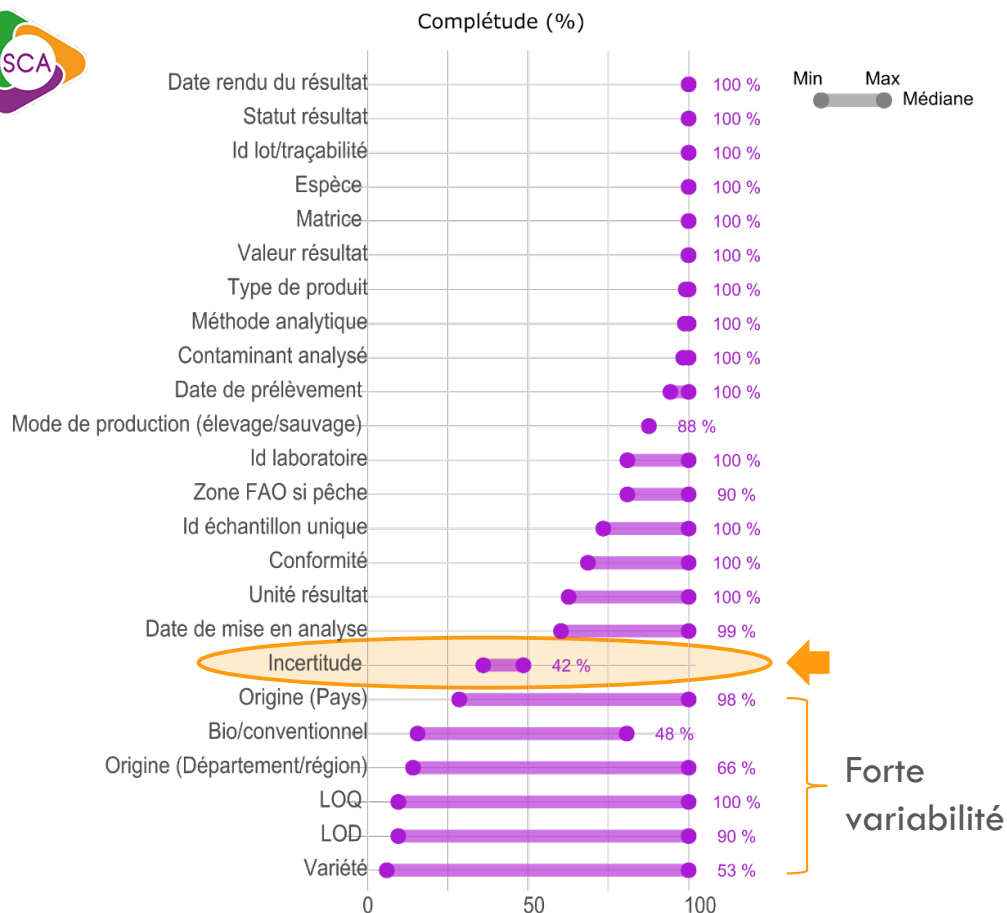
25 variables

1 colonne = 1 jeu de données

Ex: Concentration d'un échantillon de blé dur

Concentration brute	Après retranchement de l'incertitude	Teneur maximale réglementaire	Conclusion
0,20 ± 0,05	0,15 mg/kg	0,18 mg/kg	Conforme



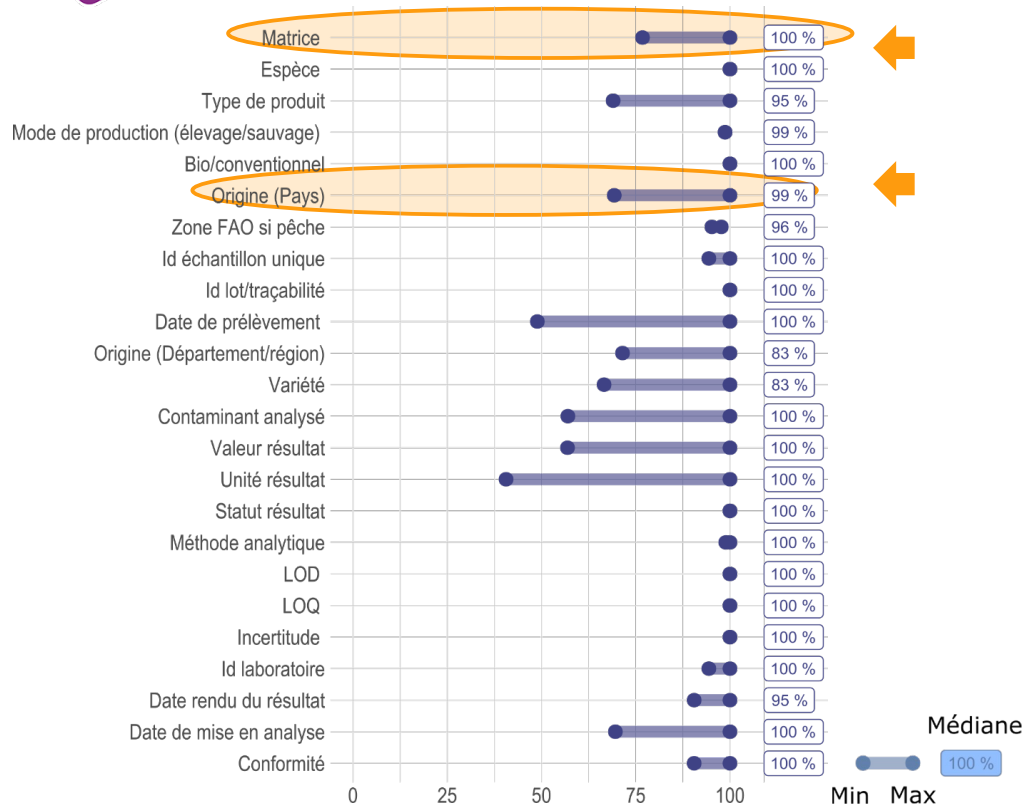


Complétude : résultats

- Complétude ~ 100 % pour les informations les plus essentielles sauf
- ...Incertitude de mesure = information la moins complète dans l'ensemble des jeux de données
- Forte variabilité en fonction des jeux de données pour certaines variables (données géographiques, performances analytiques (LD/LQ))



Validité du format (%)



Validité : résultats

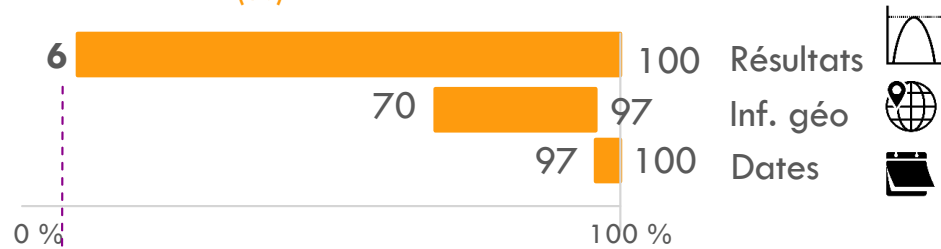
- Résultats médians très bons, quelques jeux de données à améliorer
- La plupart des corrections à envisager concernent :
 - Les dates (mélange chiffres et lettres lié à l'export vers Excel)
 - Les données numériques (mélange de points et virgules pour les résultats)
↳ +/- Rapide et facile à corriger
 - Les données textes : hétérogénéité d'écriture
↳ Très chronophage à corriger et à évaluer



Peu de jeux de données se prêtaient au calcul de ces indicateurs



Cohérences (%) minimales et maximales observées



Inversion quasi systématique entre la LD et la LQ dans un jeu de données

Cohérence : résultats

- Résultats analytiques : incohérences entre la LD et la LQ
- Informations géographiques : départements renseignés alors que le pays renseigné n'est pas la France
- Dates : cohérence excellente, rares incohérences entre date de collecte et date d'analyse



Recommandations

Complétude

- * Réfléchir aux **objectifs du dispositif** et à l'usage prévu des données
- * Appliquer le **principe de parcimonie**
 - Chaque donnée collectée doit répondre à un objectif
 - Estimer le **rapport coût de collecte/bénéfices** : mieux vaut collecter « moins et mieux »
- * Mieux flécher les informations manquantes (« NA » / « Absent » etc.)

Validité

- * Dates : définir un format unique (ex: JJ/MM/AA)
- * Valeurs numériques : définir le séparateur des dizaines (« . » ou « , »)
- * Textes :
 - Pluriel ou singulier
 - Majuscule ou minuscule
 - Accents ou non
 - Standardiser les données en créant un référentiel contrôlé (ex: liste déroulante)



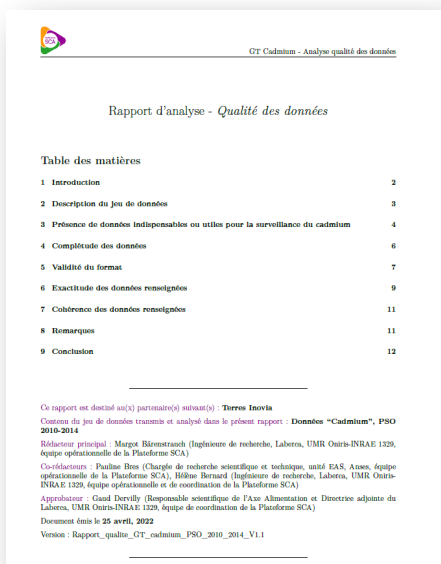
Cohérence

- * Mettre en place des tests simples pour les valeurs numériques et les dates (infériorité/égalité/supériorité)
- * Fixer des valeurs numériques minimales et maximales acceptées
- * Établir des listes contrôlées de modalités acceptées

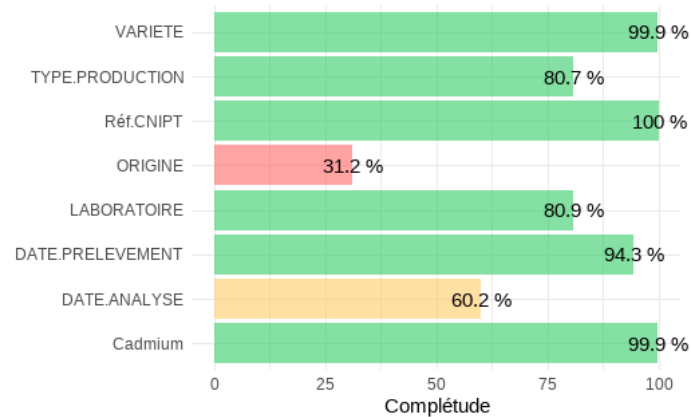


Recommandations

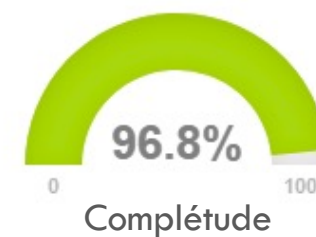
- Rapport individuel « QDD » transmis à chaque dispositif volontaire



Calculs d'indicateurs QDD par variables



Calculs d'indicateurs globaux



Recommandations pour améliorer la qualité des données





Recommandations

- Recommandations auprès des dispositifs privés/publics notamment sous forme de fiches synthétiques



A retrouver dans le rapport du GT « Cadmium »



Fiche Mémo QUALITÉ DES DONNÉES (QDD)

Optimiser un jeu de données pour le rendre interopérable et facile à exploiter, pour mieux répondre aux objectifs d'un dispositif de surveillance

Noms de colonnes : privilégier des titres concis, sans accent et sans espace

ID	Matrice	Type	Origine	Date_pre	Date_env	Analyte	Methode	Resultat	Unités	Res_type	Incertitude	LQ	Conf	Com
04_2015_01	Chou-fleur	Matière première	Belgique	02/01/2015	03/01/2015	Co	R258	1.15	mg/kg	Q	0,012	0,001	0,005	> seuil
04_2015_02	Chou-fleur	Matière première	Belgique	03/02/2015	04/02/2015	Co	R247	2.78	mg/kg	Q	0,027	0,002	0,004	> seuil
04_2015_03	Carotte	Surgele	Allemagne	03/02/2015	03/02/2015	Co	R247	0.002	mg/kg	LQ	Non applicable	0,001	0,001	< seuil
04_2015_04	Carotte	Surgele	France	08/02/2015	NULL	Co	R1578	0.005	mg/kg	LQ	Non applicable	0,0001	0,005	< seuil
04_2018_018	Carotte	Conservée	NULL	27/02/2018	28/02/2018	Co	R1578	0.005	mg/kg	LQ	Non applicable	0,0001	0,005	< seuil

Exemple de jeu de données

Fiche Mémo QUALITÉ DES DONNÉES (suite)

R1 Respecter le principe de clé d'un
Chaque résultat d'analyse doit posséder un identifiant **UNIQUE**.

- Recourir à l'implémentation unique (incrémenter un code avec au moins un chiffre et une lettre) ou créer un id en concaténant plusieurs variables
- Bloquer la saisie de doublons lors de collecte
- Nettoyer les doublons

R2 Respecter la structure standard données
1 ligne = 1 résultat d'analyse
1 colonne = 1 variable = 1 information

- Eviter la fusion de lignes, colonnes
- Eviter le mélange d'informations
- Faire figurer séparément un résultat
- Faire figurer séparément le statut du (quantifié, non quantifié, non détecté)
- Eviter de mélanger des chiffres et des (Ex : < 0,03 %)

R3 Définir les variables
Le contenu de chaque colonne doit être **sans ambiguïté** pour la personne exploitant les données.

- Nommer les variables avec des noms précis et non ambigus
- Créer un **dictionnaire des variables** en indiquant **TOUTES** les informations susceptibles d'aider l'utilisateur à interpréter les données :
 - le contenu de chaque variable
 - le format attendu
 - le caractère obligatoire ou non de l'information...

Exemple de dictionnaire des variables

Variable	Format	Description
ID	Code	Identifiant unique à chaque résultat d'analyse
Origine	Texte	Pays d'importation du produit
Date_pre	Date	Date de prélèvement
Res_incert	Numérique	Incertitude exprimée en mg/kg
Res_type	Texte	Statut du résultat (Q = quantifié, LQ = non quantifié mais détecté, LD = non détecté)

Ce fichier pourra être transmis à chaque fois que le jeu de données est partagé avec un nouvel utilisateur.

R7 Utiliser des référentiels
L'utilisation de **référentiel** est une excellente solution qui évite l'introduction d'erreurs lors de la saisie d'informations longues et complexes.

- Pour les méthodes analytiques par exemple, un référentiel peut être créé en **attribuant un code à chaque méthode**.
- Seul ce code sera saisi à l'aide d'un menu déroulant dans le jeu de données et un référentiel permettra de faire la correspondance.
- Ce référentiel doit être construit en respectant les principes QDD.

Exemple de référentiel

Code	Nom_méthode	Norme
R258	ICP-MS	EN 17053 :2018
F0247	graphite	EN 15550 :2017

Aller plus loin...

Pensez à consulter le **GUIDE PRATIQUE** sur la qualité des données de surveillance, élaboré par les plateformes d'épidémiosurveillance.
<https://wiki.ans.inrae.fr/books/guide-pratique-sur-la-qualite-des-donnees-de-surveillance>

R8 Mettre en place des étapes de nettoyage
Des tests logiques simples peuvent être mis en place afin d'identifier des erreurs de saisie qui pourront être corrigées et tracées (voir R9).

Par exemple :

- Tester les relations d'infériorité/supériorité des valeurs numériques (Ex: vérifier que LQ > LD)
- Tester les relations d'antériorité/postériorité des dates

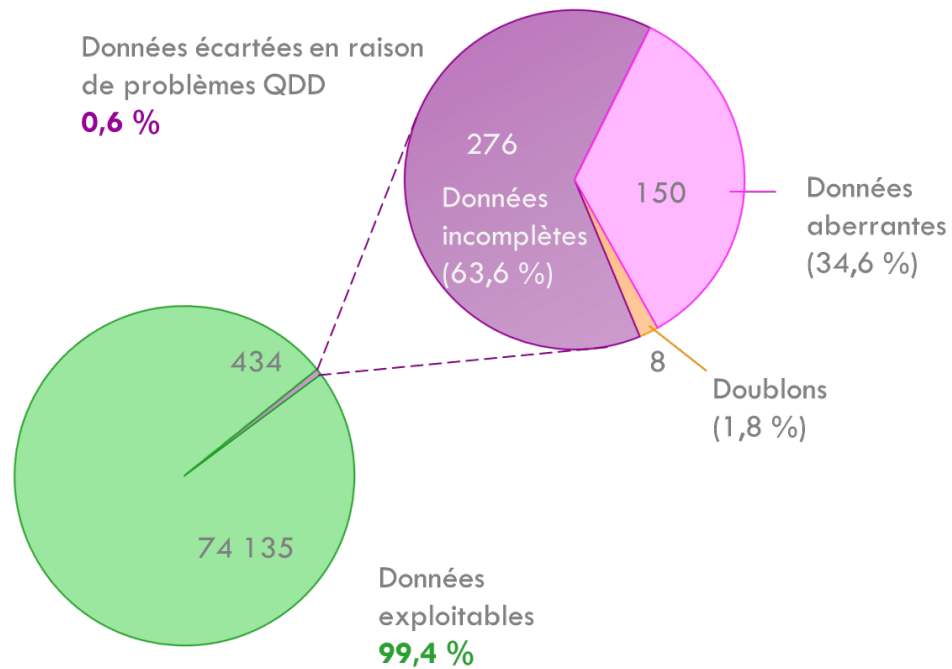
R9 Tracer les erreurs
Les informations non fiables doivent pouvoir être identifiées facilement.

- Lorsque des erreurs sont identifiées, elles peuvent être corrigées et l'ancienne valeur peut-être notée en commentaire avec une explication.
- Si des doutes sont émis sur un renseignement, ils peuvent être indiqués en commentaire également.





Conclusion



Peu de données éliminées...



Temps consacré au nettoyage et à l'exploitation des données :

3 personnes / **220** journées complètes

... mais correction/remaniement de l'ensemble des jeux de données avant exploitation



Des recommandations adaptées de celles du **guide QDD** □ travail du groupe QDD facilement **transposable** à n'importe quel dispositif de surveillance



Calcul d'indicateurs QDD simples (complétude, cohérence) = étape qui peut être aisément mise en place à l'échelle d'un dispositif pour **auto-évaluer la qualité des données de surveillance**

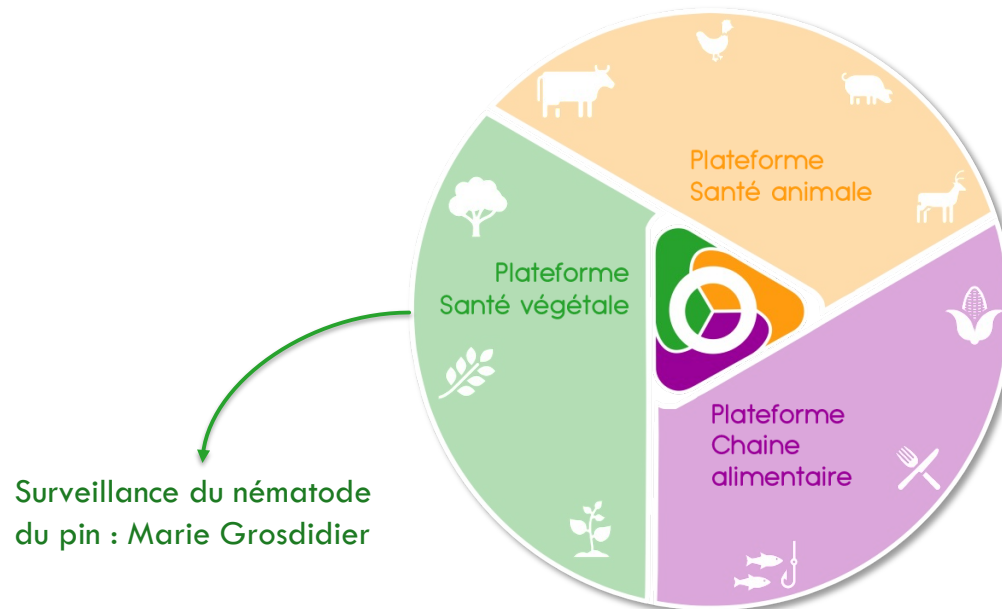


Place aux questions





Surveillance en santé végétale





Marie Grosdidier



- Ingénieure de recherche INRAE (unité BioSP, équipe OPE) à Avignon
- Epidémiologiste dans l'équipe opérationnelle de la Plateforme ESV, co-animatrice des GT SNP et VSI, membre du GS QDD
- Co-animatrice du GT SNP, viens en appui au GT SNP pour l'analyse des données de surveillance, analyse des zones à risque d'entrée, d'établissement et d'expression de la maladie en France, amélioration du dispositif de surveillance

Plus de détails sur les travaux du groupe ici !



Site de la Plateforme ESV,
page GT Surveillance
Nématode du pin



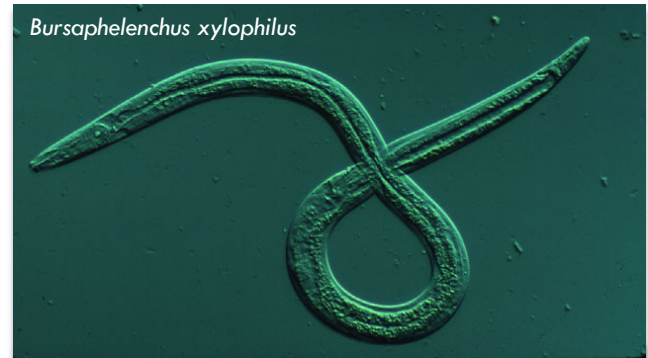
Exemple en lien avec le nématode du pin

CONTEXTE

- **Surveillance du nématode du pin :**
- Nouvelle réglementation européenne: à partir du 14 décembre 2019 (règlement 2016/2031/UE) + 200 ON à suivre
 - 20 organismes de quarantaine prioritaire (OQP) (Reg UE 2019/1702)
 - ± 180 organismes de quarantaine de l'Union (OQ, OQZP)

Bursaphelenchus xylophilus (OQP) :

- ver microscopique parasite de conifères (Pins surtout pin maritime, pin sylvestre et pin noir),
- vecteurs du genre *Monochamus*,
- symptômes ...dépérissement de l'arbre,
- présent au Portugal et en Espagne,
- propagation par la circulation de bois de conifères contaminés,
- conséquences économiques et environnementales importantes.





CONTEXTE : objectifs de la surveillance

- **Surveillance du nématode du pin en France**
 - S'assurer que le territoire en est exempt
 - Détecter le plus précocement possible un foyer
 - Connaitre la population des vecteurs

CONTEXTE : Comment on surveille

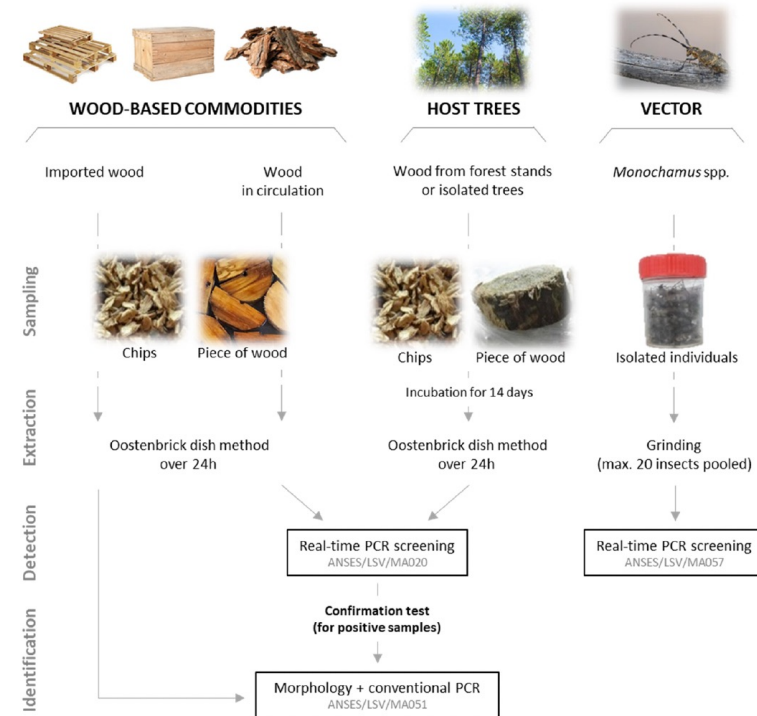


Fig. 1 Routine analytical process for *B. xylophilus* detection within the metropolitan France monitoring framework (in grey: the reference for the analysis method)

Mariette et al., 2023



CONTEXTE : Effort de surveillance en France de 2000 à 2019

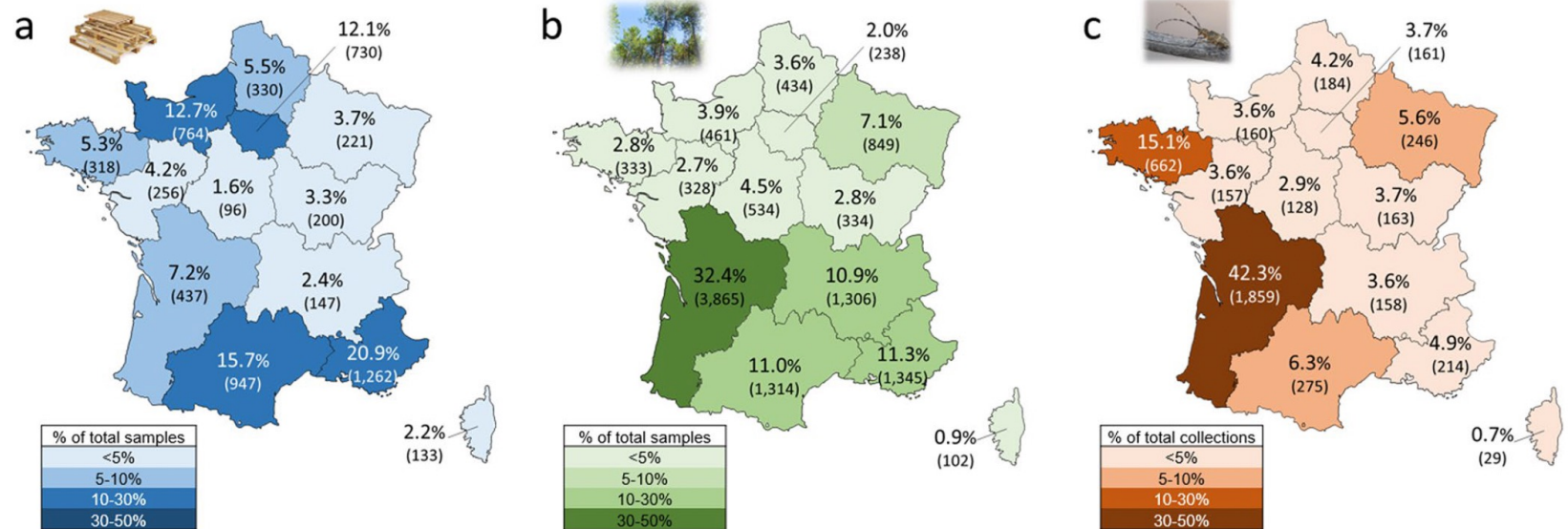
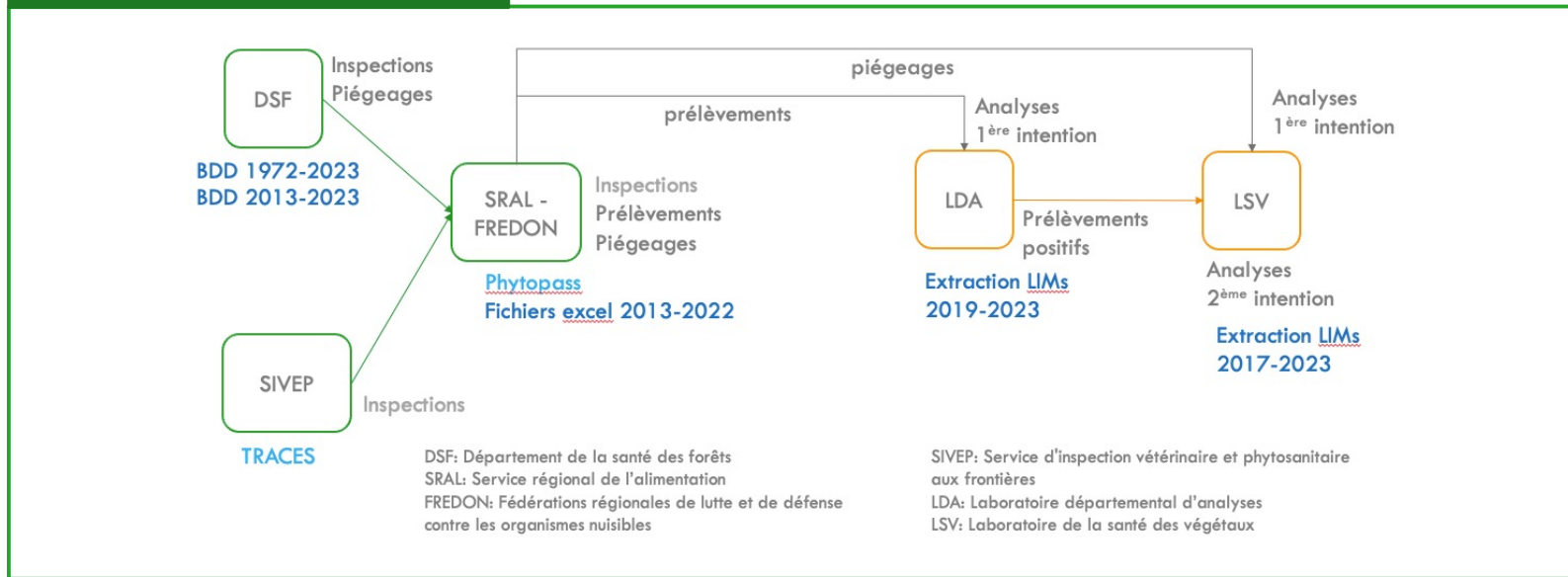


Fig. 3 Sampling effort in the framework of French monitoring of the PWN from 2000 to 2019 for **a** wood-based commodities, **b** samples of wood collected from standing trees and **c** collections of insects from *Monochamus* spp. traps. For each map, we have indicated the proportion of samples for each region (out of the total sampling), while the number of samples is indicated in brackets

Mariette et al., 2023



CONTEXTE : Flux des données



Centralisation des données
Plateforme ESV

Qualité des données



CONTEXTE : Importance de la qualité des données

• Qualité des données de surveillance

- Pour quels besoins ?
 - Améliorer le dispositif de surveillance / objectifs de la surveillance
- Quel est l'impact potentiel d'une mauvaise qualité des données ?
 - Chiffres de surveillance pour la Commission européenne erronés
 - Conséquence directe sur la réglementation et sur la filière bois (traitements ...)
 - Indirectement, passer à côté d'une détection précoce ou d'un foyer

METHODE de la Plateforme : pour des données de qualité



Données brutes

Centralise les données : BDD
Plateforme ESV

Nettoyage des données :
scripts R

Données propres



- Package en cours de construction mais déjà utilisé en partie



Complétude : méthode

- Package
 - Fonction complétude
 - Précise ce qu'est une donnée manquante (NA, "")

COMPLÉTUDE



Complétude : résultats

variable	pourcentage_complete	NbLigne_complete
num_echantillon	100.00	2353
num_echantillon_bis	26.35	620
num_echantillon_lnr	100.00	2353
commune_prelevement	81.26	1912
provenance_echantillon	94.43	2222
date_prelevement	35.32	831
expediteur_echantillon	100.00	2353
date_reception_echantillon	100.00	2353
type_echantillon	81.47	1917
modalite_surveillance	100.00	2353
matrice_echantillon	95.54	2248
methode_analyse	100.00	2353
nombre_insectes_pieges	59.03	1389
resultat	99.75	2347
date_envoi_resultat_sral	99.66	2345
LDA	14.66	345

- Vision globale des données manquantes
- Définir une stratégie pour obtenir les données nécessaires (ne peut pas se permettre d'avoir une valeur manquante), ex : le résultat des analyses



Validité : méthode - rapport

- Package
 - Fonction `qdd_dates`, `qdd_modalités`

Validité : résultats - rapport

----- Rapport de qualité -----

QDD sur la variable : `date_prelevement`
Format souhaité : `%Y-%m-%d`
Intervalle de date : 2017-01-01 au 2023-10-11

Date min : 2017-04-13
Date max : 2023-08-24

Nb NA : 1522 / 2353 (64.68 %)
Nb ligne au bon format : 831 / 2353 (35.32 %)
Nb ligne au bon format et dans l'intervalle : 831 / 2353 (35.32 %)

Merci de vous référer à la fonction `nett_format_dates()` pour nettoyer vos données

QDD sur la variable : `resultat`
Modalités à respecter : Positif, Négatif, Indéterminé, Non renseigné

Nb modalités correspondant au vecteur 'modalites' : 3
Pourcentage : 0.06%

Liste des modalités non renseignées dans le vecteur 'modalites' :
>> DETECTE, Non détecté

- Définir une stratégie pour nettoyer les données ou obtenir directement les données dans le bon format



Validité : méthode - nettoyage

- Package
 - Fonction nett_dates



Validité : résultats - nettoyage

```
-----  
Avant nettoyage :  
- nombre de valeur manquante : 2  
-----  
Après nettoyage :  
- nombre de date non nettoyées (hors valeurs manquantes) : 0 (0%)  
- nombre de date nettoyées : 2015 (99.9%)  
-----  
date_reception_echantillon date_reception_echantillon_net  
43614 2019-05-31  
2021-12-10 2021-12-10  
12/10/2022 2022-10-12
```

resultat	val_net
Non détecté	Négatif
Test négatif	Négatif
Test positif	Positif
Détecté	Positif
<NA>	<NA>

Non détecté + Test négatif valeur non renseignée dans le référentiel
Détecté + Non détecté valeur non renseignée dans le référentiel

-
- Définir une stratégie pour les données qui n'ont pas pu être nettoyées



Conclusion

BioSP Shiny Apps

CGU

Surveillance Nématode du pin

Dernière mise à jour le 2022-05-10

Sélectionnez une date de début

2007-01-01

Sélectionnez une date de fin

2022-05-10

Informations générales

Analyses Labo

SRAL/FREDON

DSF/FREDON

Carte Inspections DSF - Prélèvements SRAL/FREDON

Informations générales

Bienvenue !

Cette application concerne uniquement la surveillance du nématode du pin en France.

Elle vous permet de visualiser les analyses laboratoires réalisées par le LSV, LDA33, LDA31 et LABOCEA (depuis 2018) ; les prélèvements et relevés et FREDON (depuis 2013) ; les inspections réalisées par le DSF (depuis 1972) et les relevés de pièges réalisés par la FREDON et recensés dans la B...

Chaque partie (voir le menu sur la gauche) contient une sous-partie sur le matériel végétal et sur le piégeage.

Au sein de ces sous-parties, 3 onglets sont présentés (dans la majorité des cas) :

- un premier avec les données géolocalisées,
- un deuxième avec les données non géolocalisées et,
- un troisième 'Data' qui correspond à la visualisation du jeu de données.

La date de dernière mise à jour des données est notée en haut à droite (à côté du logo de la Plateforme ESV).

Voici quelques informations sur les données :

- Pour les prélèvements SRAL/FREDON, la Date de prélèvement est utilisée. Lorsqu'elle n'est pas renseignée, elle est notifiée en 'NA'.
- Pour les analyses LABO, la date de réception de l'échantillon au laboratoire est utilisée (la date de prélèvement n'étant pas renseignée par les données).
- Les analyses LABO positives correspondent aux analyses confirmées positives par le LSV (2ème intention).

BioSP Shiny Apps

CGU

Nématode du pin

Présence du nématode du pin

Présence des vecteurs

Surface hôtes

Surveillance

Autour de chez moi

Surveillance

Nombre d'inspections, prélèvements ou pièges

7
20
55
148
403
1,097
2,981
8,103

Carte

Filter

Types de surveillance

Inspections

Pièges

Prélèvements

Année

2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022

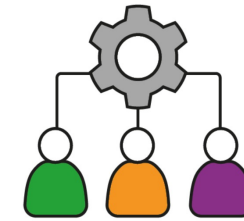
Filter

→ Sensibilisation des acteurs de la surveillance à la qualité de leurs données





Le travail sur la qualité des données rassemble tous les acteurs du programme de surveillance à toutes les étapes :



- **Sensibilisation des acteurs** (terrain, laboratoires) sur l'importance des données de qualité :
Application R shiny
- **Adaptabilité des analystes** par rapport aux données demandant le moins de contraintes possibles aux acteurs : **package R qdd**
- **Décisionnaires :**
 - Sensibilisation avec le guide QDD
 - Participent aux réflexions pour créer un outil unique de collecte des données (DGAI) limitant les erreurs possibles
 - Actions réglementaires avec les ordres de méthodes chapeau

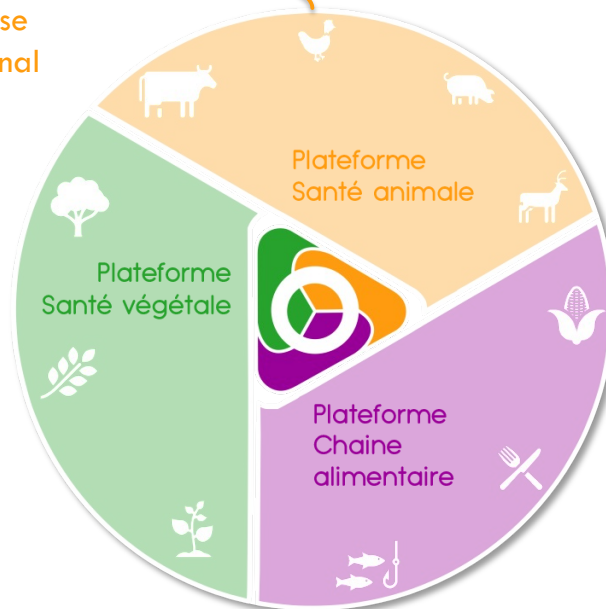
Place aux questions





Surveillance en santé animale

Surveillance de la tuberculose
en abattoir : Stéphanie Darnal





Stéphanie Darnal



- Chargée d'études (BEAD : Bureau des Etablissements d'Abattage et de découpe) à Paris
 - En charge du pilotage du Système d'Information de l'inspection en abattoir (SI2A), du suivi réglementaire relatif à la gestion des encéphalopathies spongiformes transmissibles et du suivi règlementaire de la tuberculose
 - Experte au sein de la plateforme ESA depuis 3 ans, co-anime le groupe de travail sur la surveillance de la tuberculose en abattoir
- Plus de détails sur les travaux du groupe ici !



Site de la Plateforme ESA,
page GS Tuberculose en
abattoir

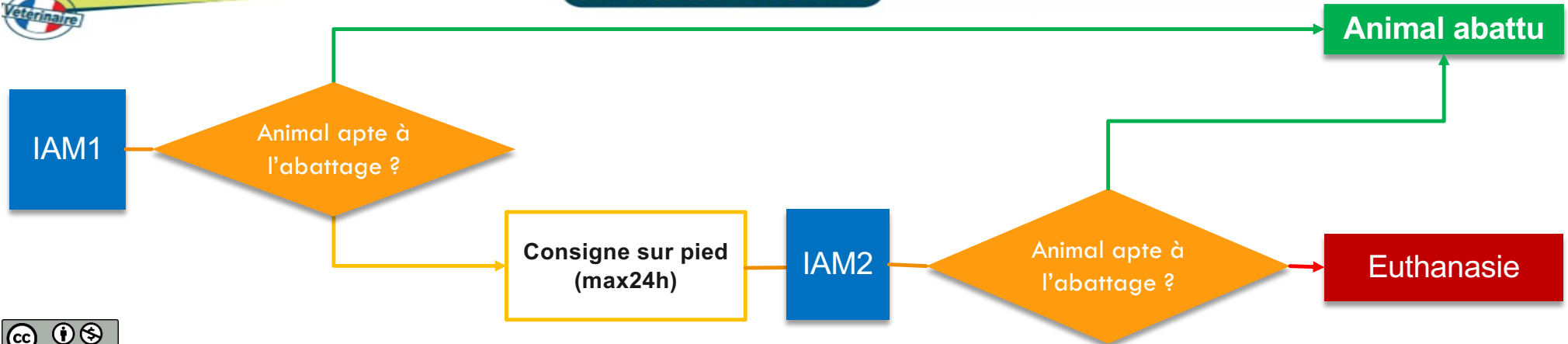
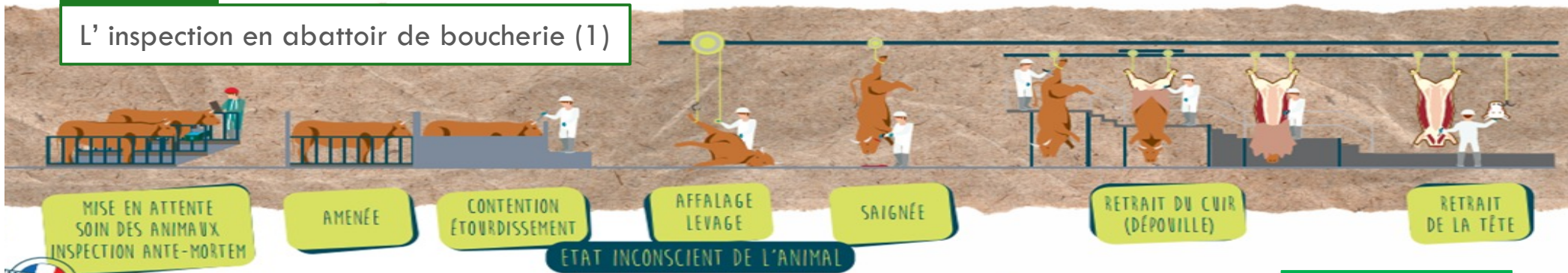




Exemple en lien avec la surveillance de la tuberculose en abattoir

CONTEXTE

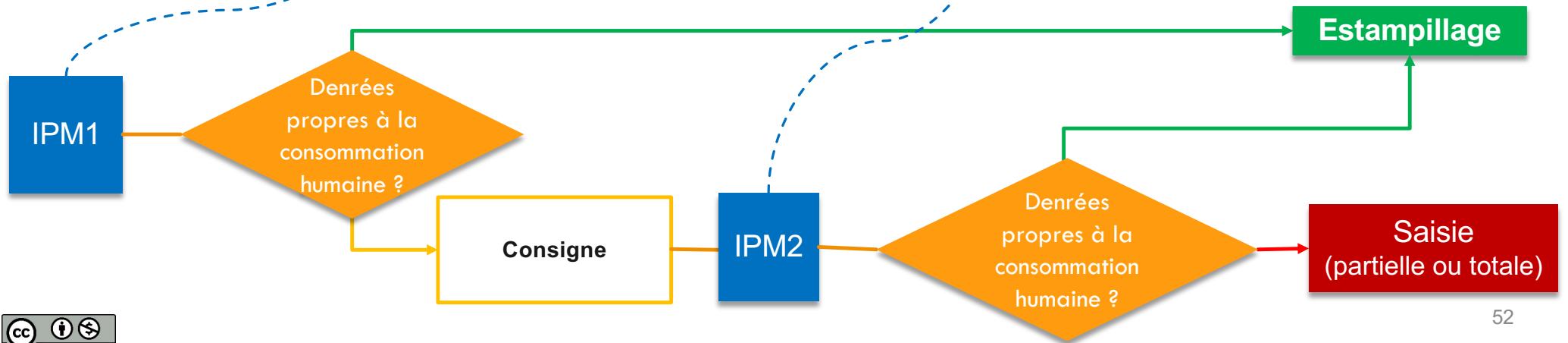
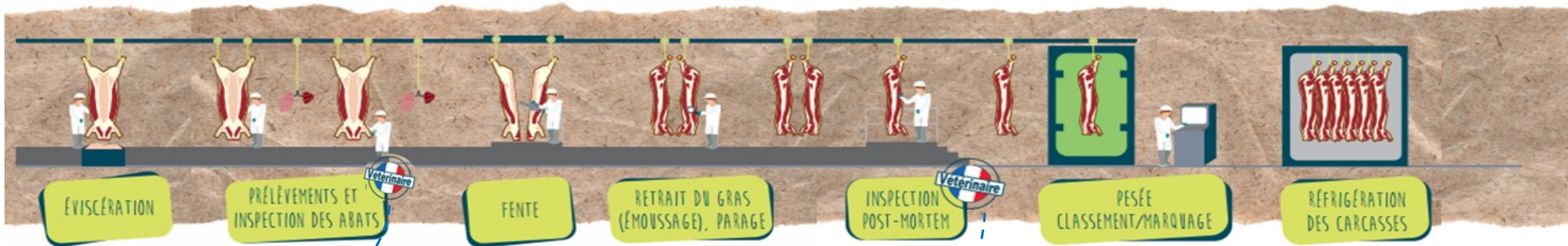
L'inspection en abattoir de boucherie (1)





CONTEXTE

L'inspection en abattoir de boucherie (2)





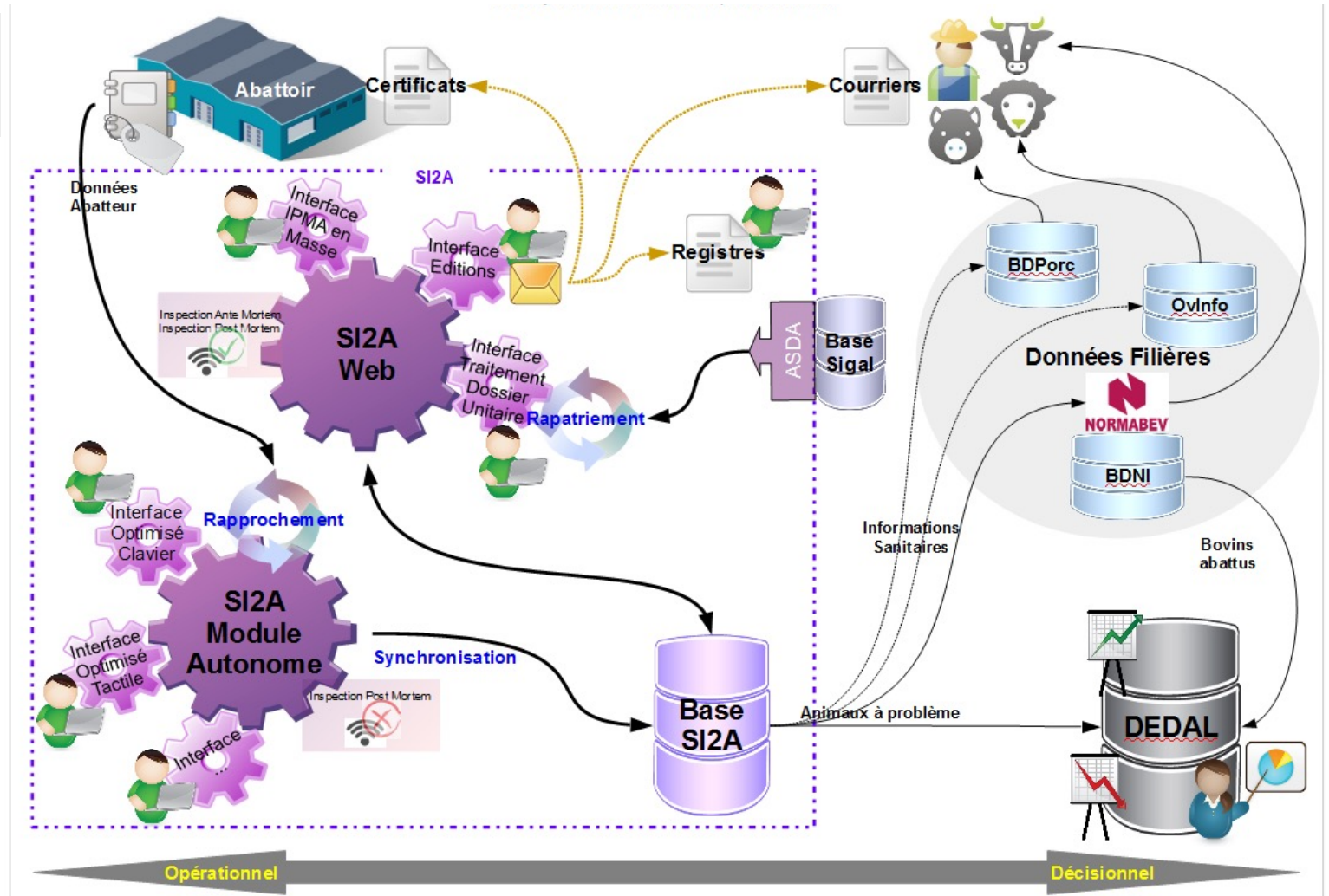
Avant SI2A...

- Diversité des outils selon les abattoirs
- Pas de :
 - Centralisation des données
 - De sauvegarde automatique des données
 - De référentiel national des motifs. *Premier test de référentiel via Nergal-Abattoir*
 - D'outil propre au ministère de l'agriculture





SI2A c'est ...

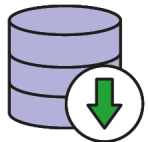




Présentation du système d'information sur l'inspection en abattoir (1)

- **Premier logiciel national** de recueil des données d'inspection en abattoir d'animaux de boucherie
- Quels besoins ?

Exigences réglementaires = générer les certificats de consigne / saisie collecter les données



Surveillance : pourquoi est-ce important de réaliser une surveillance en abattoir pour la santé publique, la santé et la protection animale?

- Santé Publique : cysticerose bovine
- Santé Animale : suivi indicateurs relatifs à la tuberculose bovine, les douves
- Protection Animale : suivi lésions traumatiques, décisions IAM



Présentation du système d'information sur l'inspection en abattoir (2)

Quel est l'impact potentiel d'une mauvaise qualité des données ?



- Certificat de saisie/consigne erroné (défaut d'information de l'éleveur limitant les mesures correctives/corrections en élevage = impact santé animale, acte administratif erroné)
- Indicateurs de suivi erronés pour la surveillance en santé publique, santé animale, protection animale
- Indicateurs pour les audits de la Commission européenne



Constat : Quantité d'enregistrements affecté à un champ « Autre motif » anormalement élevé dans un des abattoirs utilisant l'ancien outil Nergal-Abattoir

- Pourtant interface permettant renseignement a priori de tous les motifs de saisie possibles à partir d'une liste de champs fermés

Actions mises en place et solutions trouvées avec l'outil Nergal-Abattoir

Actions

- Entretien avec le personnel :
 - Selon eux un motif fréquent n'était pas présent dans la liste et par conséquent ils renseignaient ce motif dans un champ 'Autre motif '.
 - Motif de saisie en question bel et bien présent mais liste de modalités trop longue accessible en cliquant sur 'page suivante' non connue de tous

Solutions

- Respect du principe de parcimonie dans l'affichage = proposition de choix restreints de motifs en fonction de la pièce saisie permettant de limiter les affichages à l'utilisateur et optimiser l'enregistrement. Motifs les plus fréquemment utilisés également mis en premier.



Solutions trouvées

- Référentiel des pièces et motifs construit avec les agents de terrain
- Simplification dans l'affichage du motif sur l'interface « motif court » et motif officiel sur le certificat de saisie/consigne

Code	Libellé court	Libellé long
B015	Absence CVI dûment renseigné d'un ongulé domestique abattu d'urgence hors abattoir	Viandes provenant d'un ongulé domestique abattu d'urgence hors abattoir présenté sans certificat vétérinaire d'information dûment renseigné
B017	Animal malade (autre que maladie de l'OIE)	Viandes provenant d'un animal malade (autre que maladie de l'OIE)
B018	Animal atteint d'une maladie de la liste de l'OIE	Viandes provenant d'un animal atteint d'une maladie de la liste de l'OIE
B019	Préparation non conforme au paquet hygiène/PMS/process	Conditions de préparation des viandes à l'abattoir non conformes aux dispositions réglementaires en matière d'hygiène
B026	Viandes avec contaminants > TM / LMR	Viandes contenant des résidus de contaminants dépassant la TM (teneur maximale) ou la LMR (Limite Maximale de Résidus)
B027	Viandes avec métaux lourds	Mesures de retrait liées à la présence de métaux lourds dans l'environnement
B033	Bovin testé EST avec résultat non négatif	Bovin soumis à un test de dépistage des ESST avec résultat non négatif au test de dépistage des ESST
B034	Bovin testé EST prélèvement non analysable	Bovin soumis à un test de dépistage des ESST avec prélèvement non analysable
B035	Bovin non testé	Bovin non soumis à un test de dépistage des ESST alors que la réglementation en vigueur le prévoit
B036	Bovin non marqué issu du même élevage qu'un bovin positif ESB	Bovin du même élevage qu'un bovin positif au test de dépistage de l'ESB et qui aurait fait l'objet d'un marquage
B037	Bovin abattu <u>avant</u> le bovin résultat EST non négatif	Bovin abattu juste <u>avant</u> le bovin ayant un résultat non négatif (confirmé par la laboratoire de référence) au test de dépistage des ESST dans un abattoir dans lequel un procédé de décontamination entre carcasses n'est pas validé ou efficace.



Solutions trouvées

- Avec SI2A module autonome
- Contextualisation des pièces/motifs : en fonction des pièces saisies la liste des motifs est réduite / top 5 des motifs pour la pièce saisie
- Création de motifs et pièces favoris

Jambon avec pied

Numéro d'abattage : 27900004

Détail pièce	Lésion/Motif*	Détail	Destination*	Observations
entier(e)	Abcès	multiples	C2	Jambon avec pied Abcès unique
partiel(le)	Arthrite(s)	unique		partiel(le) C2
parage	Souillures			
Latéralité				
droit(e)	Infiltration			
gauche	Non conforme hygiène			
droit(e) et gauche	Autres Motifs ▶			





Constat : Quantité d'enregistrement affecté à un champ « Autre motif » anormalement élevé dans un des abattoirs utilisant l'ancien outil Nergal-Abattoir



Messages clés

- Importance de l'analyse descriptive d'un jeu de données
- Importance d'avoir un champ « Autre » pour détecter ce type de problématique sinon les utilisateurs contournent en mettant un motif « approchant » erroné et c'est impossible à détecter. (permet de faire évoluer le référentiel si besoins identifiés)





Complétude



Indicateur : Pourcentage de données manquantes

- Besoin : un certificat de saisie est attachable sans précision de certaines informations relatives à la saisie. Il ne faut pas permettre l'édition d'un certificat de saisie sans ces données.

Actions réalisées

- Anticipation pour s'assurer d'un pourcentage de données manquantes (exemple poids= 0 par défaut, identité du VO)
- Champs identifiés pour être à enregistrement obligatoire avec Pop-up d'information de l'utilisateur

ERREUR : IPM2 - Enregistrement d'une inspection avec décision: au moins une pièce du volet Observation doit être renseignée

Date IPM2 (★)	Agent du SVI	Protocole d'inspection (★)
<input type="text" value="17/05/2021"/>	<input type="text" value="- Non renseigné -"/>	<input type="text" value="Standard"/>

Volet Observations

Exemple de message alertant l'utilisateur de l'oubli d'enregistrement d'une pièce et motif de saisie.

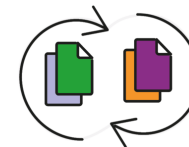


Constat : identification de données essentielles dont l'absence est non envisageable



Messages clés

- Importance d'identifier les données absolument nécessaires et rendre leur enregistrement obligatoire
- Attention à ne pas rendre obligatoire ce qui ne le nécessite pas sinon risque d'être contre-productif avec une absence totale d'enregistrement





- **Indicateur : % d'enregistrement avec date d'IPM (inspection post-mortem) > 7 jours de la date d'abattage**
 - Il n'est pas normal d'avoir un délai entre l'abattage et l'IPM aussi élevé

Solutions

- Mise en place d'un contrôle automatisé de cohérence directement lors de la saisie dans l'outil en alertant l'opérateur de saisie avec un message d'alerte (pop-up) lorsque une incohérence simple

Veillez confirmer votre action...

Il existe une ou plusieurs alertes : consultez la zone de message sous le bandeau pour les découvrir.

> Annuler

☑ Confirmer

Figure 3 : exemple de message alertant l'utilisateur car il a saisi une date pour les données d'IPM qui est à plus de 7 jours de la date d'abattage ce qui n'est pas normal. Le système n'est pas totalement bloquant car ce cas n'est pas impossible même si très rare, il alerte donc l'utilisateur en lui demandant explicitement de confirmer que c'est bien cette date qu'il souhaite renseigner.



SI2A : Un travail sur la qualité des données depuis sa création

Le travail sur la qualité des données a été réfléchi depuis le début du projet SI2A *via* :

- Co-construction de l'outil avec la méthode Agile (pour également être adapté aux contraintes de terrain des utilisateurs)
- Multidisciplinarité des acteurs impliqués dans le projet : informaticiens, agents de terrain, référents nationaux abattoirs, experts régaliens, épidémiologiste



La qualité des données est un travail en continu

- Importance de la formation
- Importance de la documentation pour les utilisateurs pour une bonne appropriation de l'outil
- Sensibilisation aux enjeux-impacts de la qualité des données pour la surveillance

Avec la qualité des données, rien n'est jamais terminé
Pour aller plus loin : chapitre dédié à l'animation de la qualité des données du guide



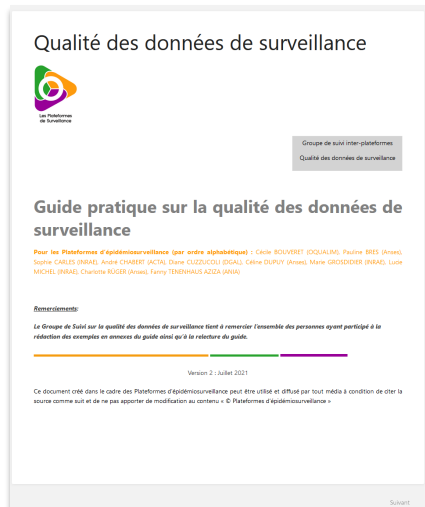
Place aux questions





Pour en savoir plus, allez voir le guide et les fiches !

Guide pratique



Accéder au guide complet






Les fiches synthétiques








Accéder aux fiches synthétiques

Fruit d'un travail collaboratif

Membres du groupe

-  Cécile BOUVERET
-  Pauline BRES
-  Sophie CARLES
-  André CHABERT
-  Diane CUZZUCOLI

-  Céline DUPUY
-  Marie GROSIDIER
-  Lucie MICHEL
-  Charlotte RUGER
-  Fanny TENENHAUS-AZIZA

 + contributeur.rices aux exemples





Pour plus d'informations sur les Plateformes

Plateforme ESA

www.plateforme-esa.fr



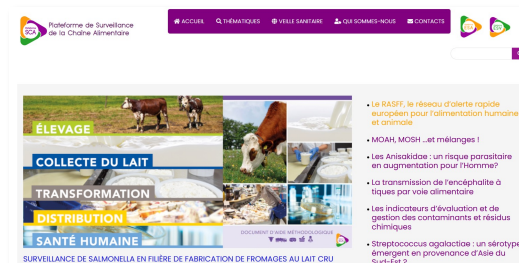
Plateforme ESV

www.plateforme-esv.fr



Plateforme SCA

www.plateforme-sca.fr





Merci pour votre attention
et votre écoute !