

## Script vidéo du cas « Fraude à la donnée »

Je vais vous présenter une situation de fraude potentielle à la donnée. Voici les principaux protagonistes :

Li, directrice de recherche en hydrobiologie.

Lucien, agriculteur et défenseur d'une agriculture conventionnelle.

Myriam, professeure des écoles engagée dans la protection de l'environnement.

Jean-Claude, entrepreneur local en produits chimiques agricoles.

Aujourd'hui, une bonne nouvelle vient d'arriver sur le bureau de Li. En effet, elle vient d'obtenir un financement pour un projet de recherche portant sur la cartographie de la qualité de l'eau au sein d'une vaste région agricole.

La qualité de l'eau, le traitement chimique et la géolocalisation étant des données sensibles, Li a donc contacté la Déléguée Informatique et Libertés pour la mise en conformité avec le R.G.P.D., et la Direction des Affaires Juridiques pour s'assurer du respect du Code de l'Environnement sur la façon de collecter les données et leur anonymisation.

Li envisage une démarche de sciences participatives pour une collecte massive d'échantillons. En effet, une telle méthode correspond bien aux nouvelles attentes de la société en matière de science ouverte, et d'une population sensibilisée à la question de la qualité de l'eau.

Dans cette approche, les citoyens sont impliqués dans la collecte de données, mais pas dans la construction du projet de recherche. Li récupère des centaines de candidatures pour récolter des échantillons d'eau.

Ne possédant que 100 kits d'échantillonnage de l'eau, elle et son équipe sélectionnent 100 candidats selon des critères de sexe, âge, position géographique et profession. Des agriculteurs et agricultrices, mais aussi des personnes de la société civile, dont Myriam et Jean-Claude, sont sélectionnés.

Myriam connaît bien la population de sa région et s'étonne de la présence de Jean-Claude parmi les participants. Elle s'interroge sur l'impartialité de ce dernier dans la collecte de données, craignant qu'il puisse biaiser les échantillons en faveur de l'industrie chimique.

Réalisant que le laboratoire de Li est situé à des centaines de kilomètres, et n'est donc pas au courant des réputations et professions des uns et des autres, Myriam fait part de ses doutes.

Parallèlement, Li constate que les concentrations de polluants de certains échantillons sont anormalement basses. Les rumeurs circulent, et Lucien l'agriculteur, pris de remords, reconnaît que Jean-Claude a demandé aux agriculteurs de falsifier les échantillons en échange d'une ristourne sur le prix des pesticides de leurs exploitations agricoles. Il ne sait combien ont accepté les offres de Jean-Claude.

Li se demande alors ce qu'elle peut faire :

- Fermer les yeux et analyser toutes les données pour publier et faire un retour aux financeurs ?
- Recommencer la campagne d'échantillonnage en revoyant les méthodes de prélèvement ?
- Supprimer les mesures anormalement basses ?

Après avoir pris connaissance des éléments clés de cette histoire, voyons à présent comment ce cas de fraude potentielle à la donnée résonne en vous. Pour guider votre réflexion, je vous propose de répondre au questionnaire posé sous la vidéo.