



GBADs

**et Une seule santé :
ressources de
données**

D^{re} Theresa Bernardo
Chaire IDEXX, Collège
vétérinaire de l'Ontario

Pr^e Deb Stacey
École d'informatique de
l'Université de Guelph



Financement mondial de la santé en 2019

Total : 41 milliards USD

Donateur	Montant USD	% du total	Par habitant	% du PIB
Canada	1,1 milliard	3	29 (38 CAD)	0,06
États-Unis	12 milliards	30	37	0,06
Fondation Gates	3,9 milliards	10		

Bill Gates: GBD 'best chance of saving lives'

Publication Authors



Kelly Bienhoff
Research Manager



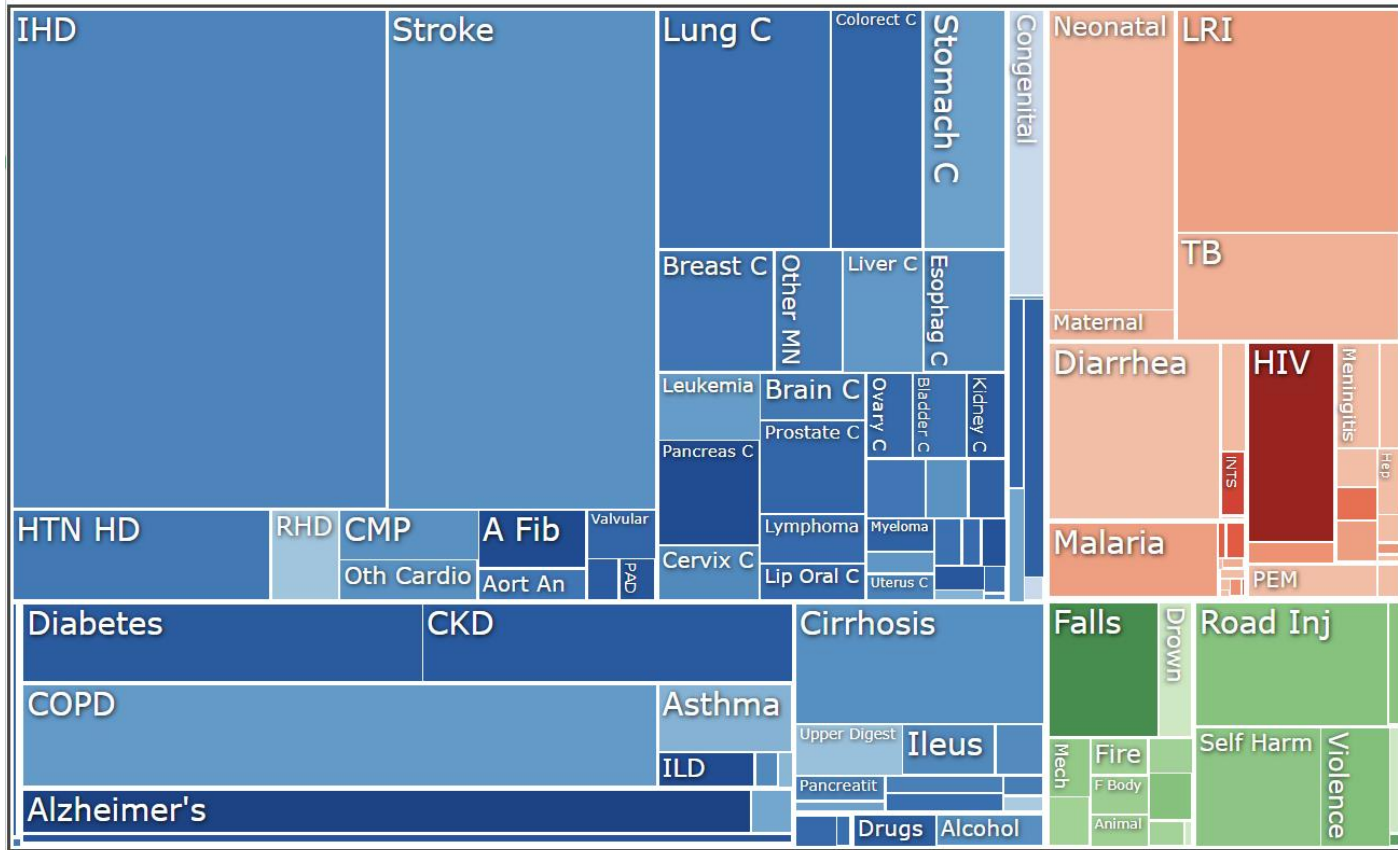
Publication date:

October 18, 2017



De quoi est-ce que les gens meurent?

Décès dans le monde en 2019



Non transmissibles

Infectieuses et alimentaires

Traumatismes

GBADs

PUBLIC RELEASE: 27-MAR-2018

Founders of Global Burden of Disease study receive award for research excellence

INSTITUTE FOR HEALTH METRICS AND EVALUATION



SHARE



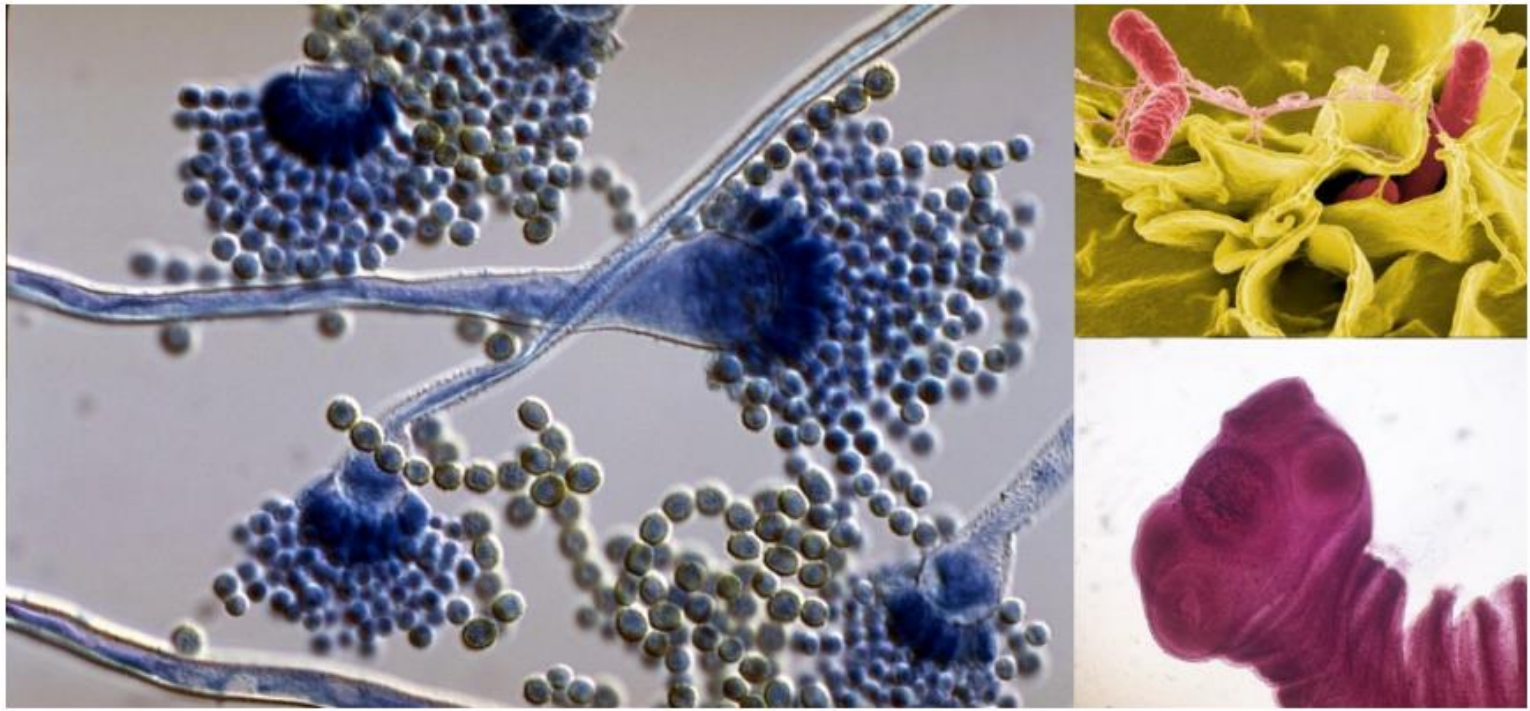
PRINT



E-MAIL

TORONTO - The co-founders of the groundbreaking Global Burden of Disease study (GBD), Professors Christopher Murray and Alan Lopez, have been selected for an international award honoring the "world's top scientists who have made outstanding achievements in global health research."

Since its launch over a quarter of a century ago, the GBD collaboration has generated nearly 20,000 peer-reviewed publications and has received more than 700,000 citations in scientific studies and reports.



Reflections on the WHO Initiative to Estimate the Global Burden of Foodborne Diseases

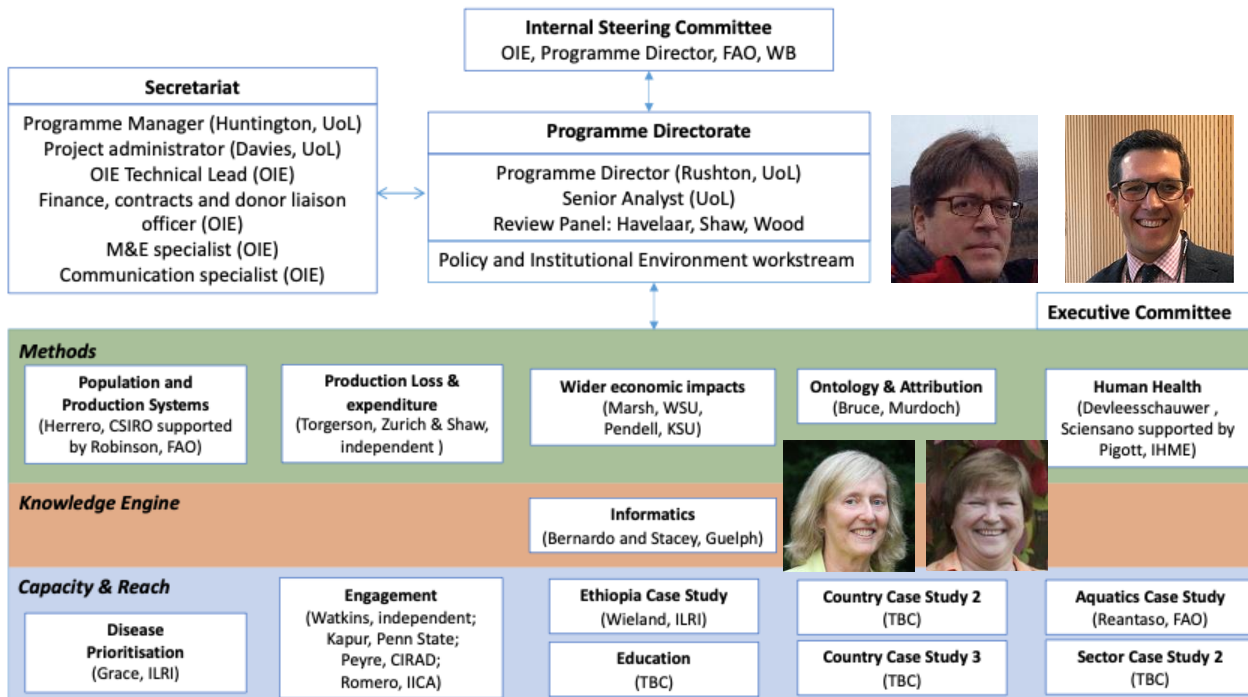
You are here: [Home](#) / [Projects](#) / [Global Burden of Crop Loss](#)

Global Burden of Crop Loss

Efforts to reach Sustainable Development Goals in food security, nutrition and livelihoods are being hindered by crop loss. Up 40% of crop yields are lost to pests and disease but the data available to prove and show trends is limited. The Global Burden of Crop Loss project will collect, validate, analyse and disseminate data on the extent and causes of crop loss, with the aim of gathering sufficient and reliable data that can act as evidence to enable prioritisation of research and policy in plant health, improving our ability to predict the impact of emerging diseases.

Global Burden of Animal Diseases

GBADS WILL MEASURE AND IMPROVE SOCIETAL OUTCOMES FROM LIVESTOCK AND HAVE A POSITIVE IMPACT ON THE SUSTAINABLE DEVELOPMENT GOALS CONTRIBUTING TO A WORLD IN WHICH THERE IS ZERO HUNGER, GOOD HEALTH AND WELL-BEING, GENDER EQUALITY, DECENT WORK AND ECONOMIC GROWTH AND RESPONSIBLE CONSUMPTION AND PRODUCTION





Groupe de travail informatique



Groupe de travail informatique

Extrants

An 1

1,5

An 2

An 3

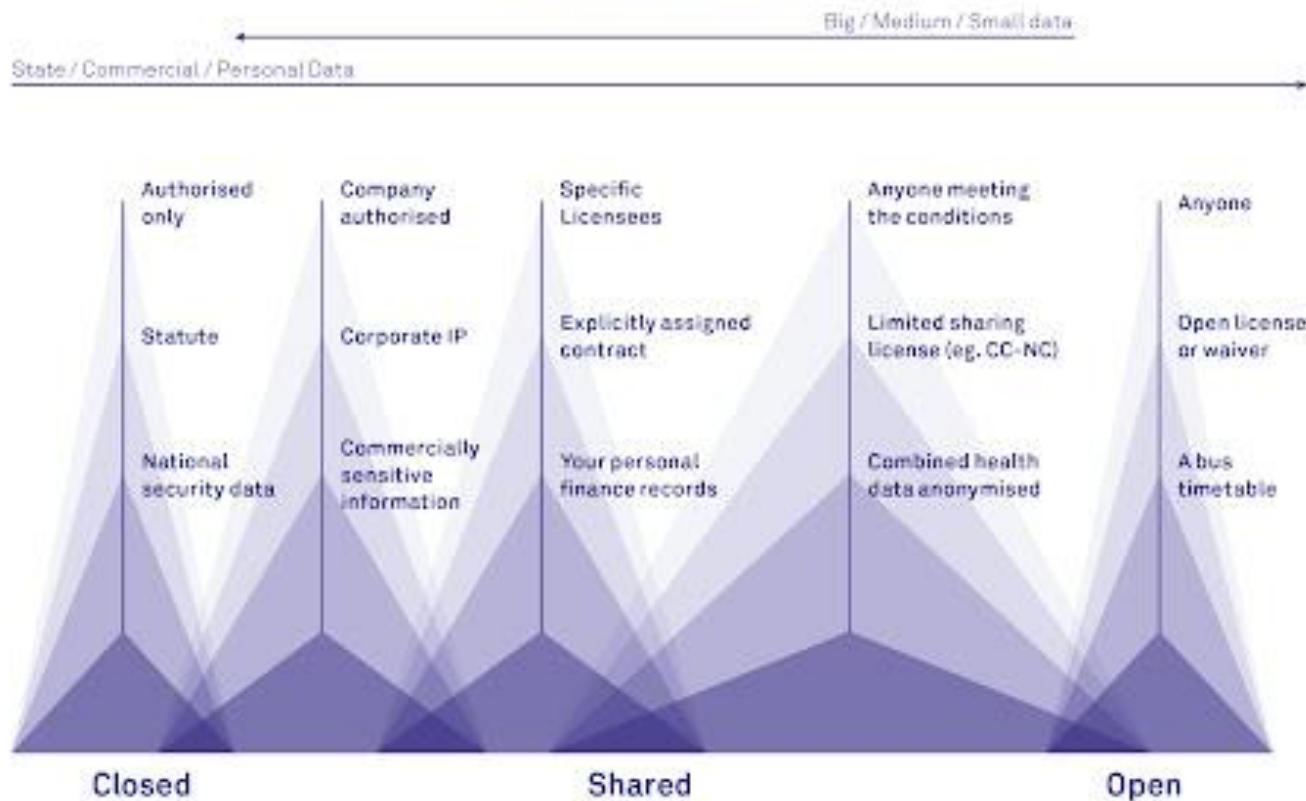
An 4

An 5

**2. Cartographier le
paysage, la qualité
et l'accès pour les
données existantes**

Groupe de travail informatique

The Data Spectrum





COOP DREAMS

Why I would raise chickens

I'm excited about the poverty-fighting power of poultry.

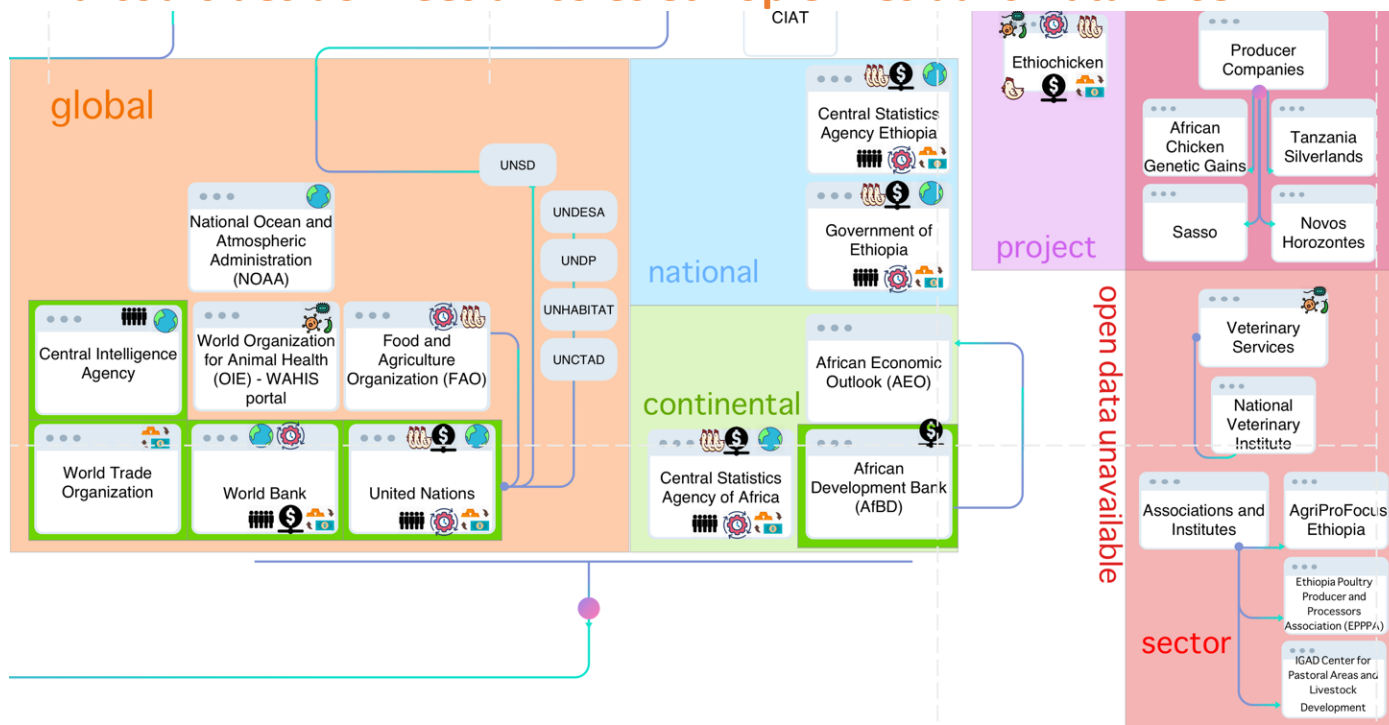
By **Bill Gates** | June 07, 2016 • 3 minute read



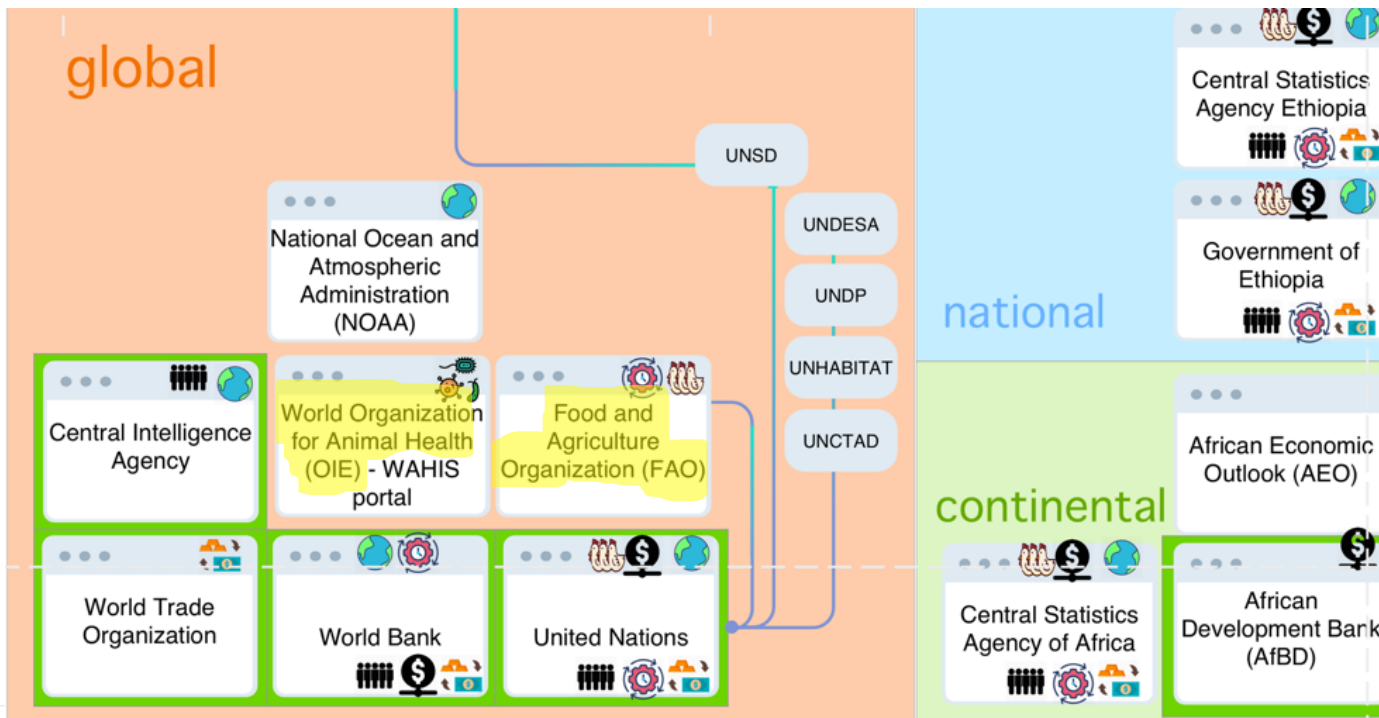
<https://animalhe...>

If you were living on \$2 a day, what would you do to improve your life?

Parcours des données avicoles éthiopiennes dans Dataverse



Parcours des données avicoles éthiopiennes dans Dataverse



Comment obtient-on les données?

- En téléchargeant des fichiers
- En grattant des sites Web
 - Les données ne sont disponibles que sur des pages Web
- Au moyen d'API (interfaces de programmation d'applications)
 - Interaction machine-machine



FAOSTAT

	Data	Selected Indicators	Compare Data	Definitions and Standards	FAQ	<input type="text" value="Search an Indicator or Commodity"/>
--	------	---------------------	--------------	---------------------------	-----	---

Food and agriculture data

FAOSTAT provides free access to food and agriculture data for over 245 countries and territories and covers all FAO regional groupings from 1961 to the most recent year available.

[Explore Data](#)

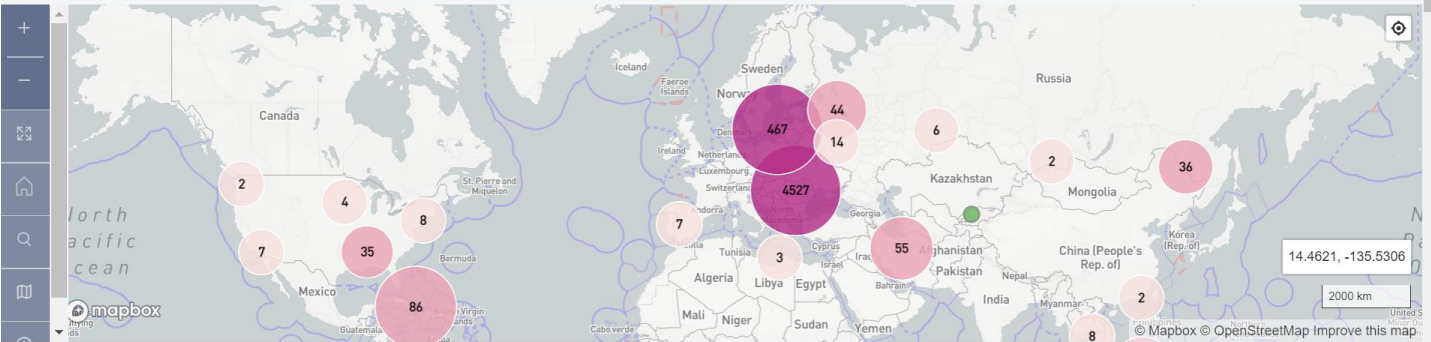


Bulk Download

All FAOSTAT Data 928 MB

Updated on Aug 20, 2021

Animal disease events



Outbreak locations have been provided by the relevant Veterinary Services and may not represent the exact location of an outbreak. OIE assumes no liability for the data displayed.

Below you can find reports (**immediate notifications and follow-up reports**) on relevant **animal disease events** for OIE-listed diseases and emerging diseases. This table allows you to search, filter, visualize and download the reports.

Consult further information

La voie de la reproductibilité

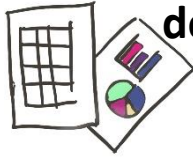
Personnes et processus



THE ROAD TO REPRODUCIBILITY

Qualité des données

Acquisition et ingestion des données



Code



Tableaux de bord



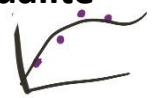
Décisions éclairées

DATA UTOPIA

Harmonisation des données

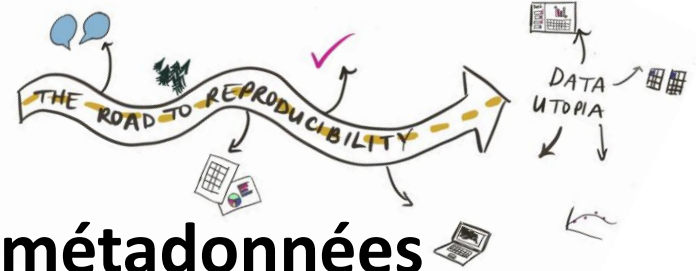


Modèles de haute qualité

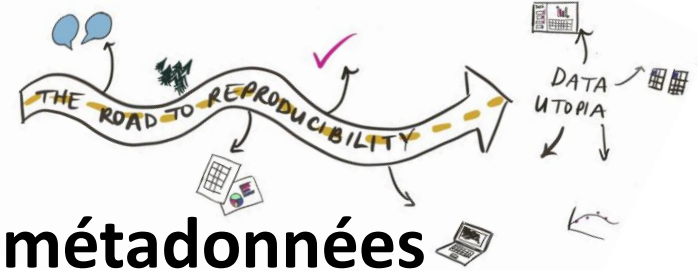


L'informatique – La voie de la reproductibilité

1. Valeurs et buts
2. Qualité des données et des métadonnées
3. Acquisition et ingestion des données
4. Personnes et processus



L'informatique – La voie de la reproductibilité



1. Valeurs et buts
2. **Qualité des données et des métadonnées**
 - 2.1. Sources de données
 - 2.2. Analyse de la qualité des données internes
 - 2.3. Analyse de la qualité des données externes
3. Acquisition et ingestion des données
4. Personnes et processus

2.2. Qualité des données

« La préparation des données représente environ 80 % du travail des scientifiques des données »

- Il y a de **NOMBREUSES** dimensions à la qualité des données.
- **Pas besoin de réinventer la roue – nous faisons le travail pour vous!**

2.2. Analyse de la qualité des données internes

- **Qualité des données internes : De quelle qualité sont les données d'une MÊME source/d'un MÊME fournisseur?**
- **Dimensions de la qualité des données internes :**
 - A. Exactitude
 - B. Crédibilité
 - C. Exhaustivité
 - D. Cohérence
 - E. Disponibilité des formats de données
 - F. Crédibilité longitudinale

2.2. Analyse de la qualité des données internes : Dimensions de la qualité des données internes

A. Exactitude : Les données présentent-elles des valeurs et des types de données réalistes?

Par exemple :

- Les métadonnées reflètent le contenu des données
- Les valeurs aberrantes dans les données sont minimales, ou quand elles sont présentes, réalistes
- **Pas de nombres à virgule flottante (décimales) dans les chiffres sur les populations animales**

2.2. Analyse de la qualité des données internes : Dimensions de la qualité des données internes

A. Exactitude : Les données présentent-elles des valeurs et des types de données réalistes?

	2005	2006	2007
Total cattle and calves	5,848	5,739	5,598
All female cattle	4,244	4,196	4,095
Aged 2 years or more	2,550	2,541	2,475

54 % de types de données incorrects
Décimales dans les données sur les
populations de bovins

Statistical data set

Structure of the agricultural industry in England and the UK at June

Detailed annual statistics on the structure of the agricultural industry at 1 June in England and the UK.

From: [Department for Environment, Food & Rural Affairs](#)

Source des données : Royaume-Uni. Department for Environment, Food and Rural Affairs. Statistical data set: Structure of the agricultural industry in England and the UK at June, 2021 (consulté le 13 septembre 2021). [Fichier de données.] Sur Internet : https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/946266/structure-june-eng-series-22dec20.ods

2.2. Analyse de la qualité des données internes : Dimensions de la qualité des données internes

5. Exhaustivité :

Y a-t-il des valeurs manquantes ou incomplètes dans les données? La **différence entre les données non collectées/inconnues, de valeur 0, de valeur imputée, etc.** est-elle claire?

2.3. Analyse de la qualité des données externes

- **Qualité des données externes : De quelle qualité sont les données ENTRE différentes sources de données?**
- **Concordance :**
 - Les « mêmes observations » de différentes sources donnent-elles le même chiffre?
 - Comment les méthodes de collecte qui donnent les mêmes observations se comparent-elles?
 - Les valeurs manquantes concordent-elles d'une source de données à l'autre?

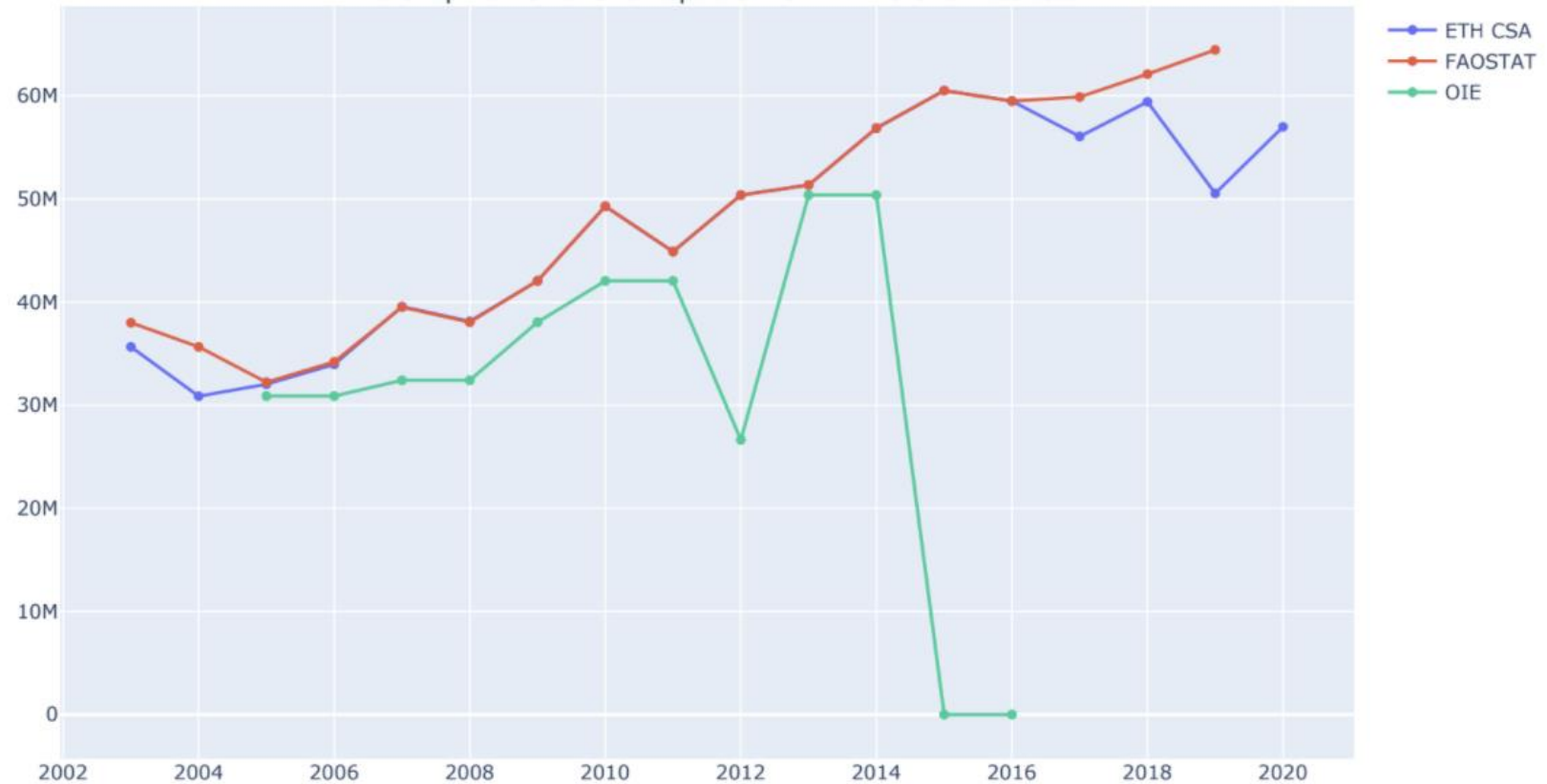
Data Quality Analysis: Population (ETH CSA, FAOSTAT, OIE)

Line Graph of Cattle Populations in Different Datasets

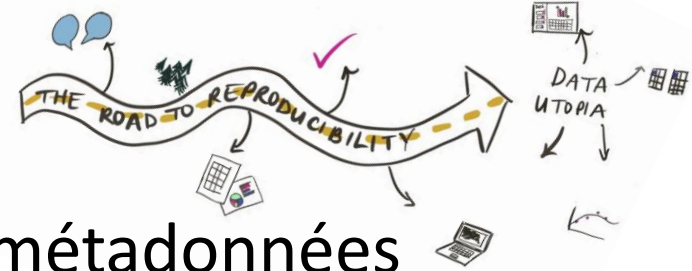


Data Quality Analysis: Population (ETH CSA, FAOSTAT, OIE)

Line Graph of Chickens Populations in Different Datasets



L'informatique – La voie de la reproductibilité



1. Valeurs et buts
2. Qualité des données et des métadonnées
- 3. Acquisition et ingestion des données**
 - 3.1. Simplification de l'ingestion des données dans GBADs
 - 3.2. Licences de données**
 - 3.3. Provenance et citation des données
 - 3.4. Code : GBADs dans GitHub
4. Personnes et processus

3.2. Licences de données

Les contacts personnels sont insuffisants!



Image : <https://www.gida-global.org/care>

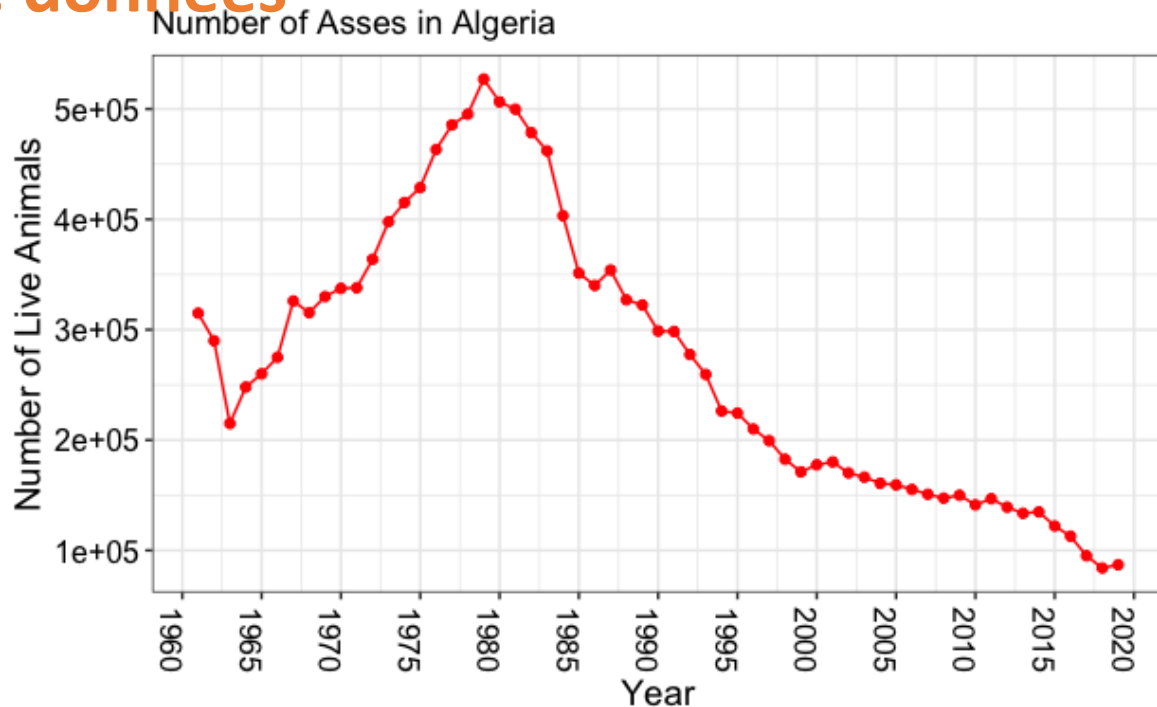
Buts de l'octroi de licences de données :

1. Les licences de données constituent souvent une obligation légale
2. Protéger et respecter les contributeurs de données (principes CARE)
3. Informer les utilisateurs des données (qui peut utiliser les données, comment et quand?)
4. Éclairer la vue générale (données privées, données accessibles à un groupe)

Pour plus de détails sur les licences de données pour GBADs : <http://www.gbadske.org/Documentation/DataGovernanceHandbook/dataOwnership.html>

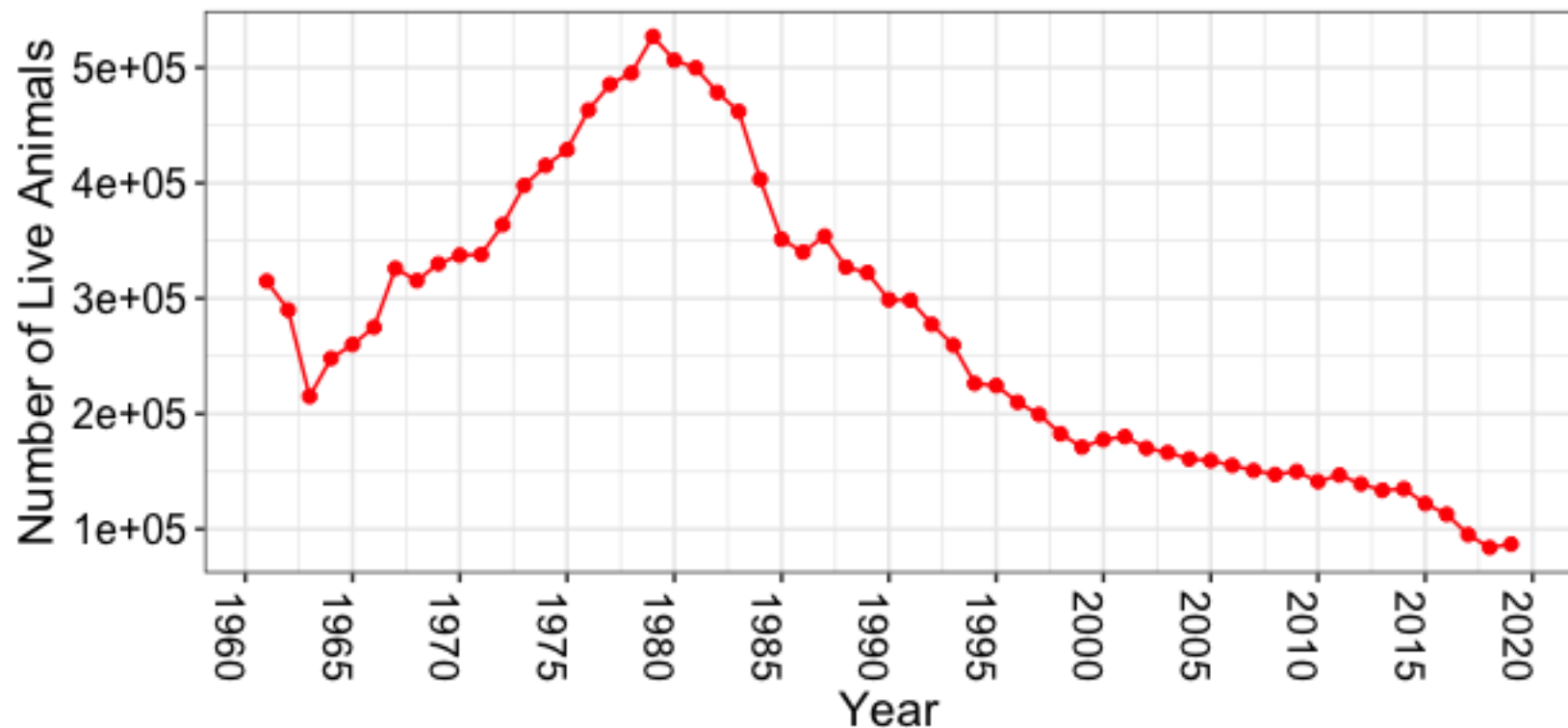
3.3. Citations de données

- Tout comme les communications écrites, les données doivent être citées!



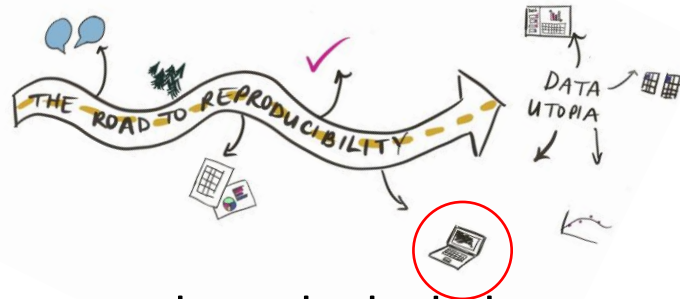
Mais d'où viennent ces données!?

Number of Asses in Algeria

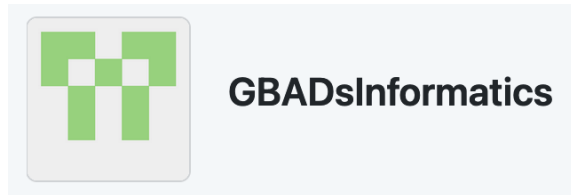
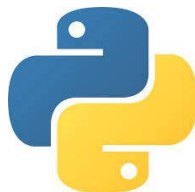


Data Source: Food and Agriculture Organization of the United Nations. (2021).
FAOSTAT, Crop and Livestock Products, Live Animals. [Data file].
Retrieved August 30, 2021 from <http://www.fao.org/faostat/en/#data/QCL>

3.4. Code : GBADs dans GitHub

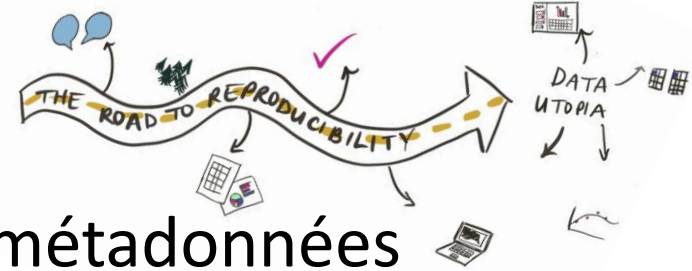


- Un lieu de **partage de code** pour les modèles comme les calculs de la biomasse
- Permet de convertir un code en outils, de simplifier les calculs et de contribuer à la « science ouverte » et à la reproductibilité
- Pratiques exemplaires pour le code R dans le Guide de gouvernance des données

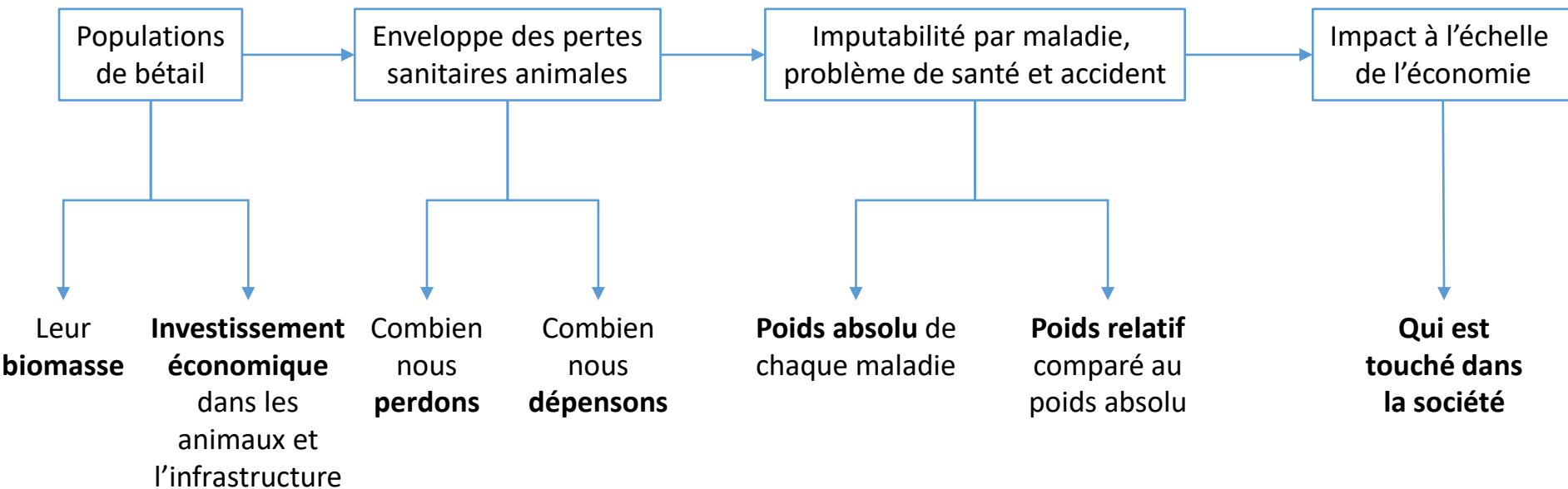


L'informatique – La voie de la reproductibilité

1. Valeurs et buts
2. Qualité des données et des métadonnées
3. Acquisition et ingestion des données
4. **Personnes et processus**
 - 4.1. Travailler avec d'autres domaines
 - 4.2. Ce que le domaine de l'informatique peut offrir aux autres

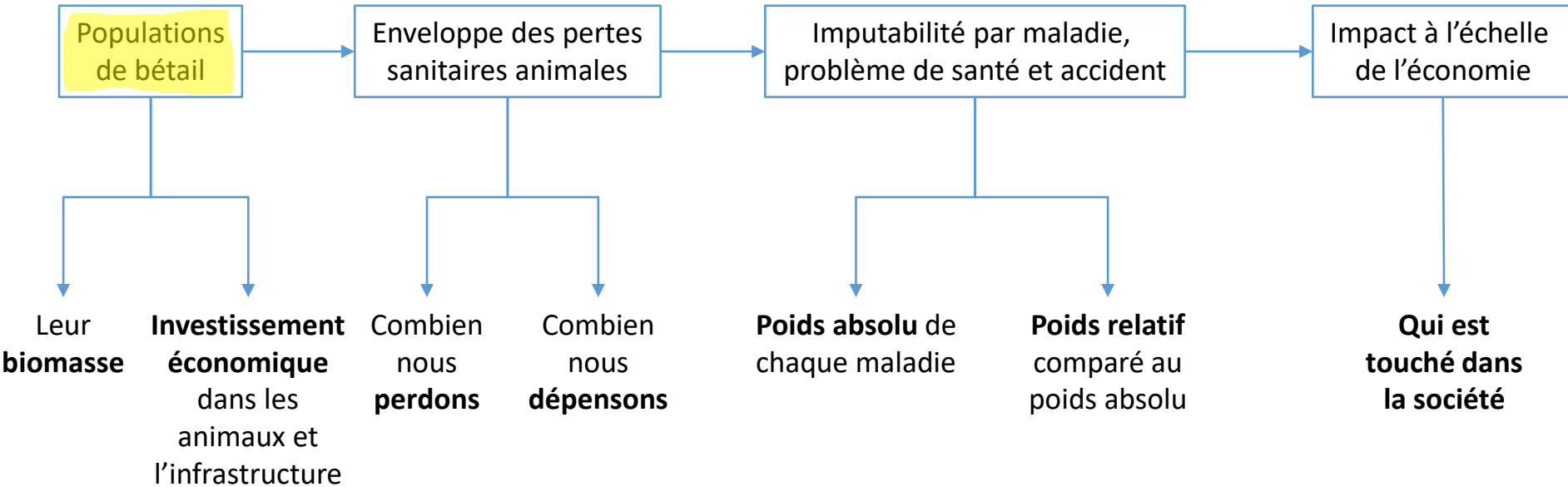


GBADs - Structure analytique



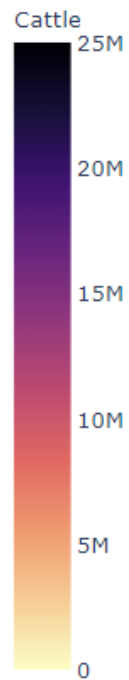
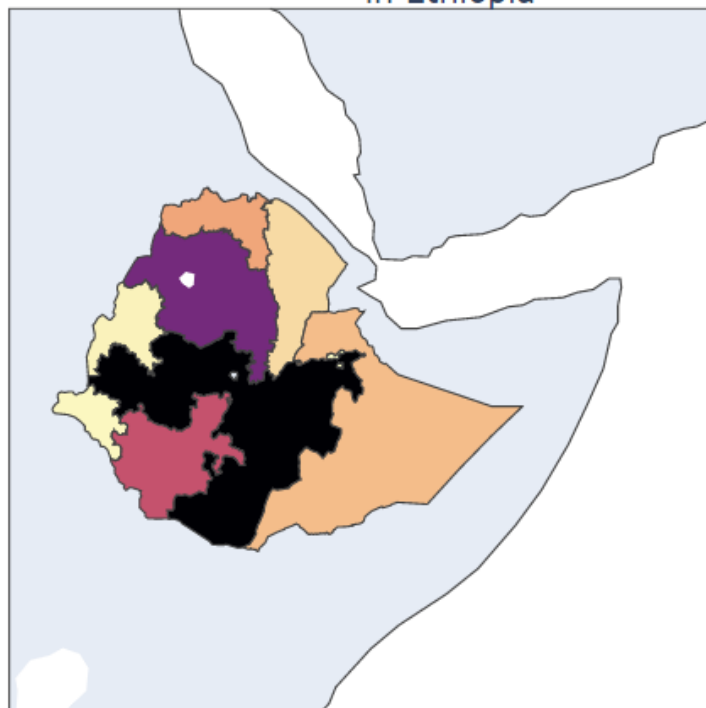
Rushton *et al.* 2021

GBADs - Structure analytique



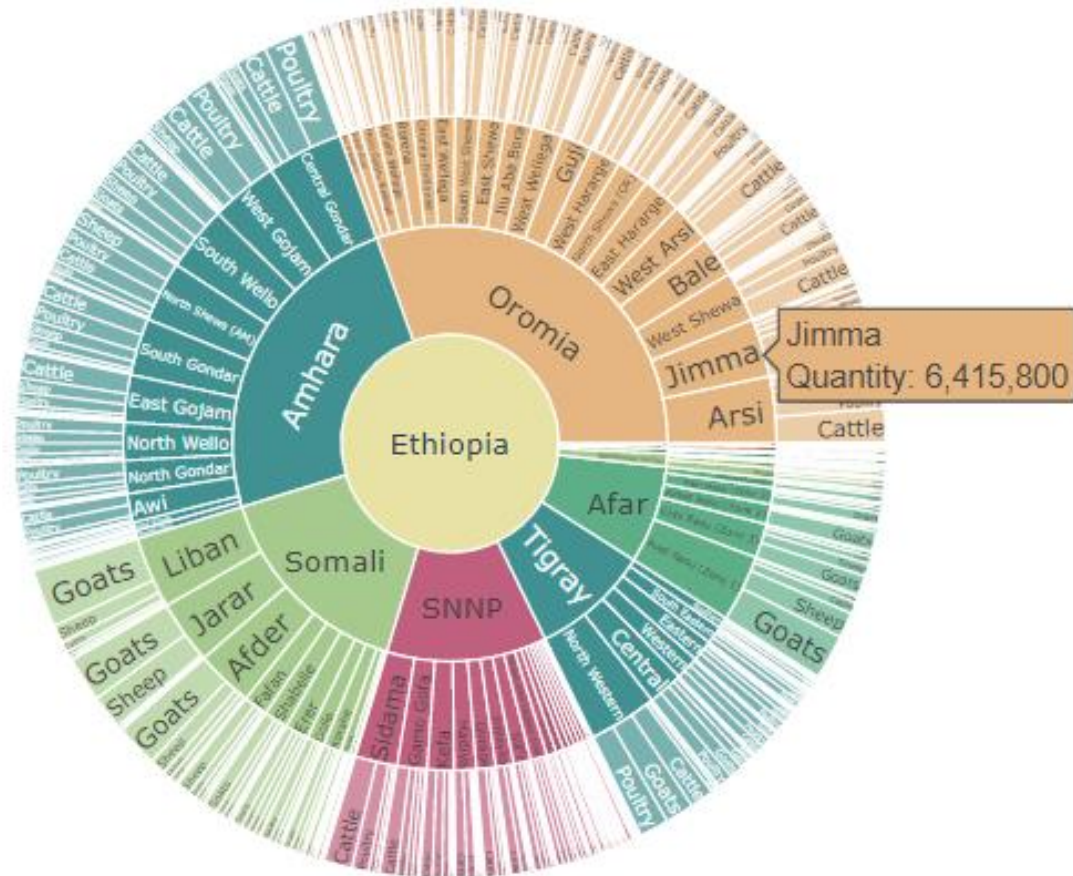
Rushton *et al.* 2021

Number of Cattle by Region in Ethiopia

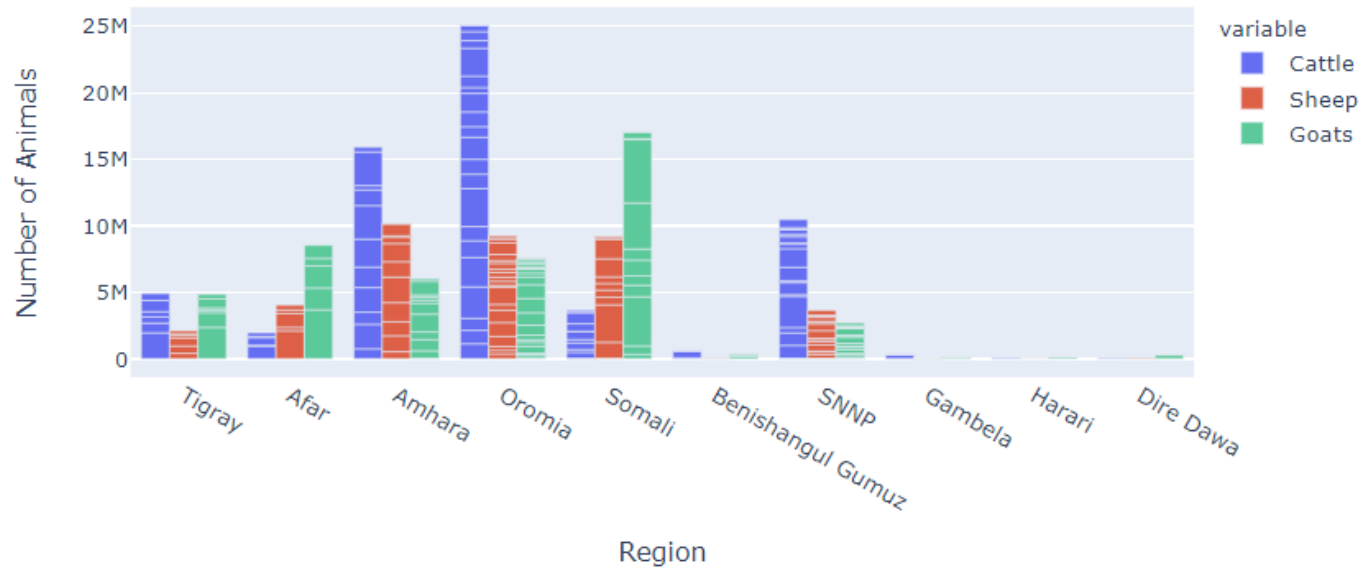


Source: Ethiopia Central Statistical Agency

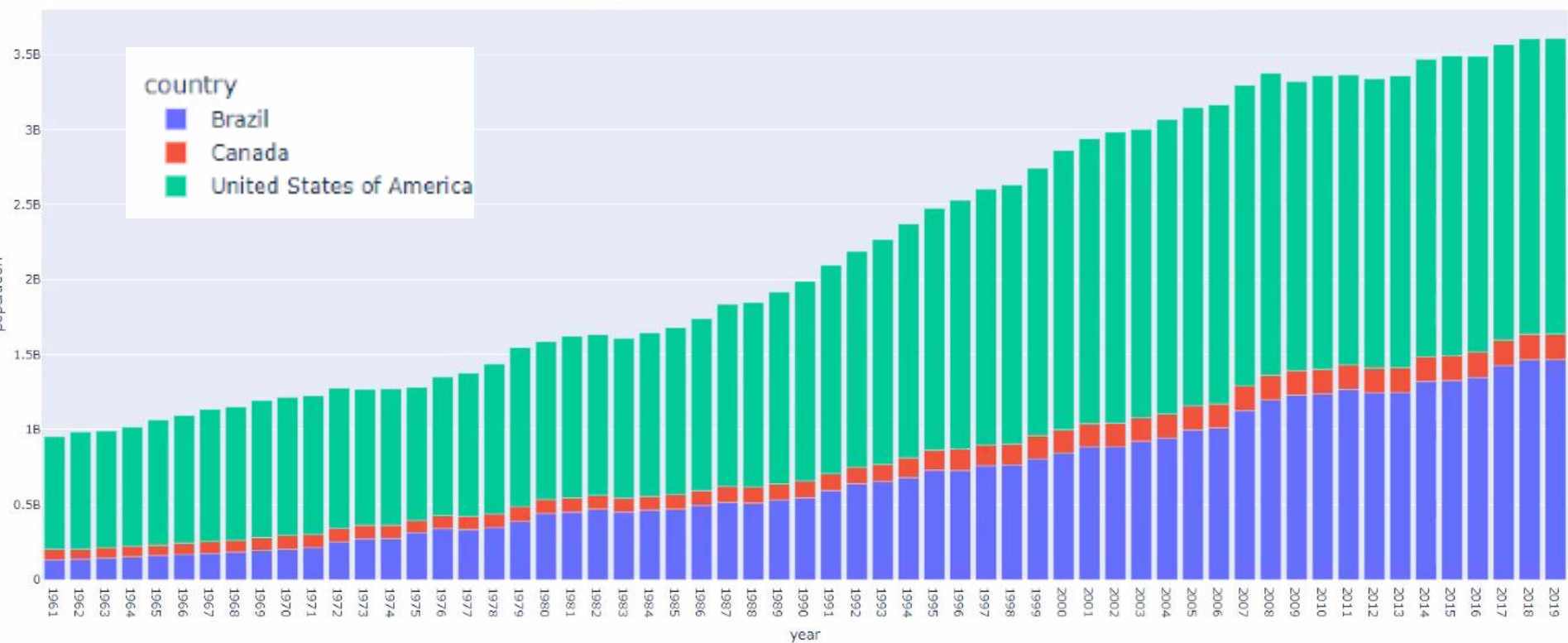
Population by Species across Ethiopia



Number of Animals by Species by Region



Source: Ethiopia Central Statistical Agency

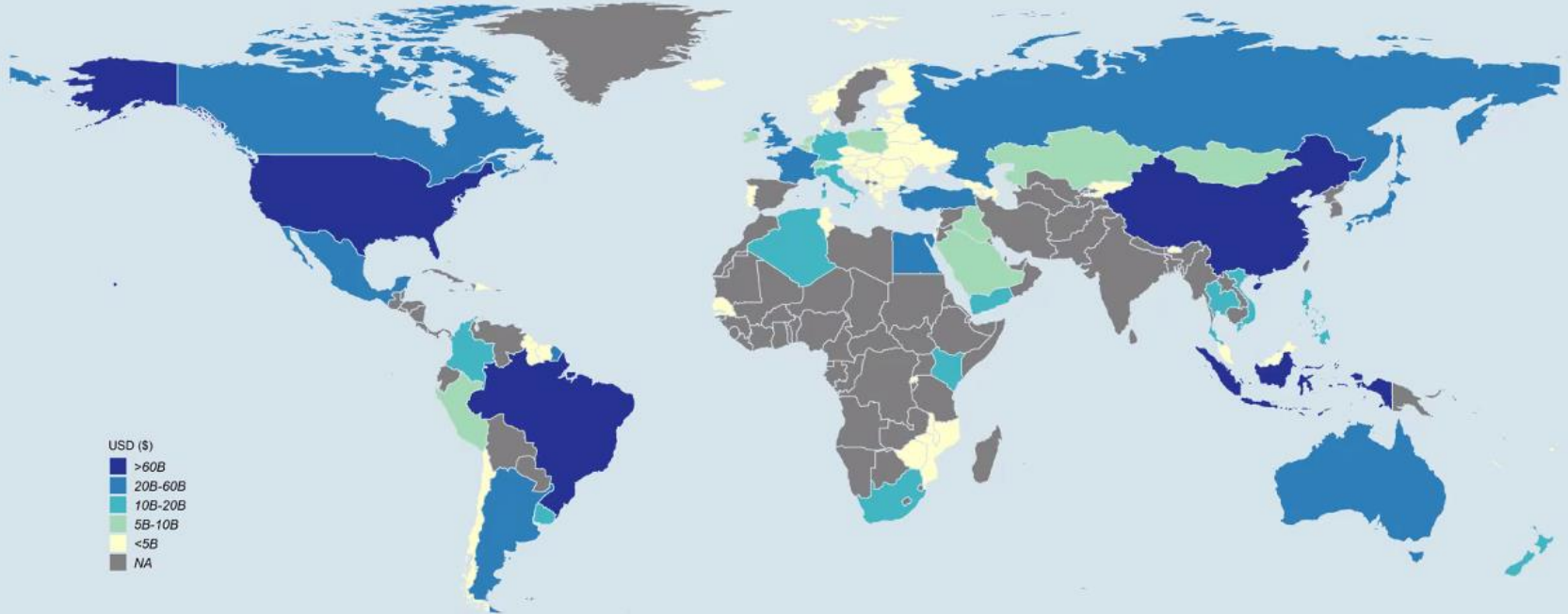


Scenario AS2

What is the global distribution of the value of animal stock in 2015?

This estimate includes cattle, chickens, pigs, sheep, pigs, camels, mules, and horses.
All values are reported in current US dollars.

Preliminary results



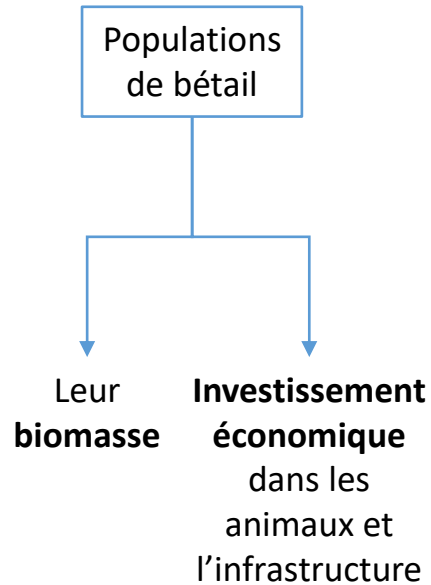
Data: FAO (2021) Notes: AS2 does not include live aquatic animals due to unavailable data. For NA quantity and/or price/value data is missing.

GBADs - Structure analytique

Populations
de bétail

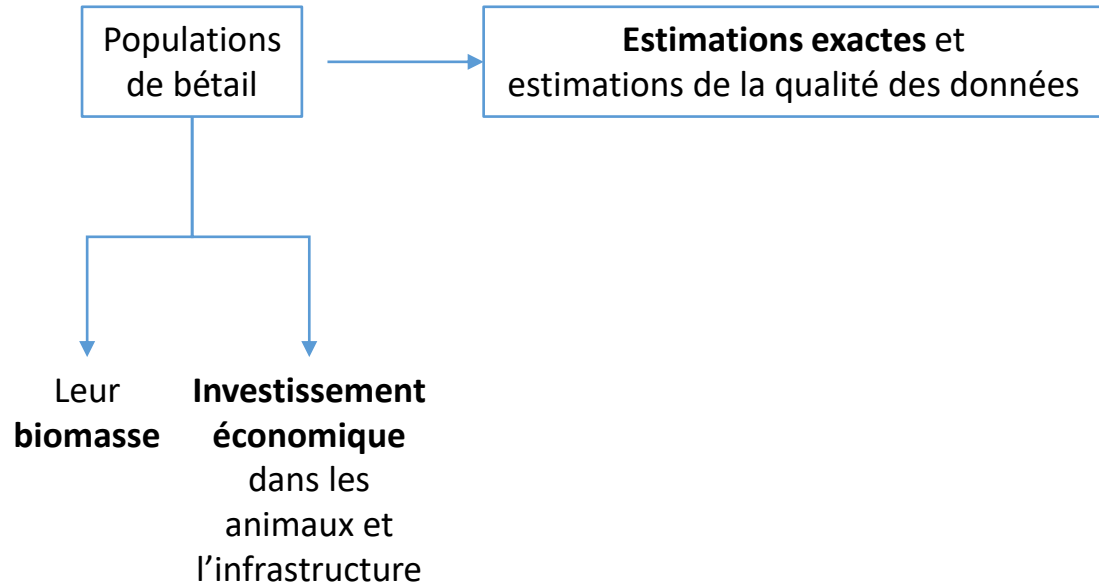
Rushton *et al.* 2021

GBADs - Structure analytique



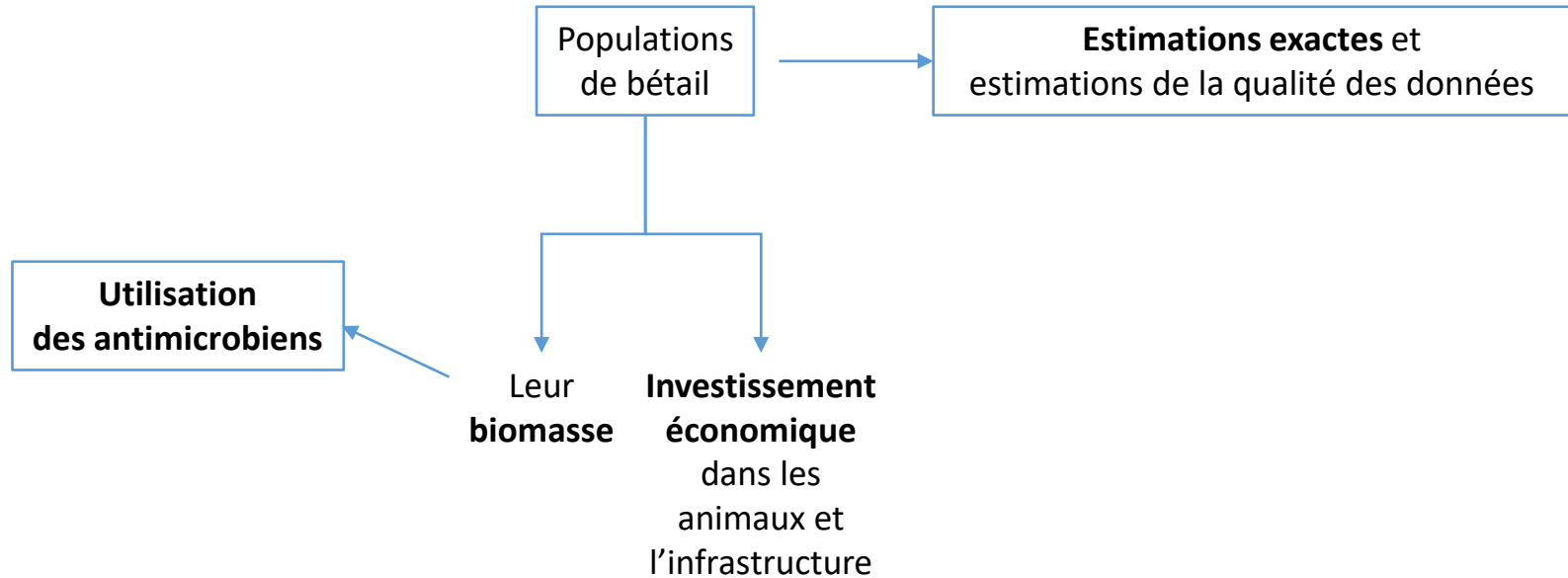
Rushton *et al.* 2021

GBADs - Structure analytique



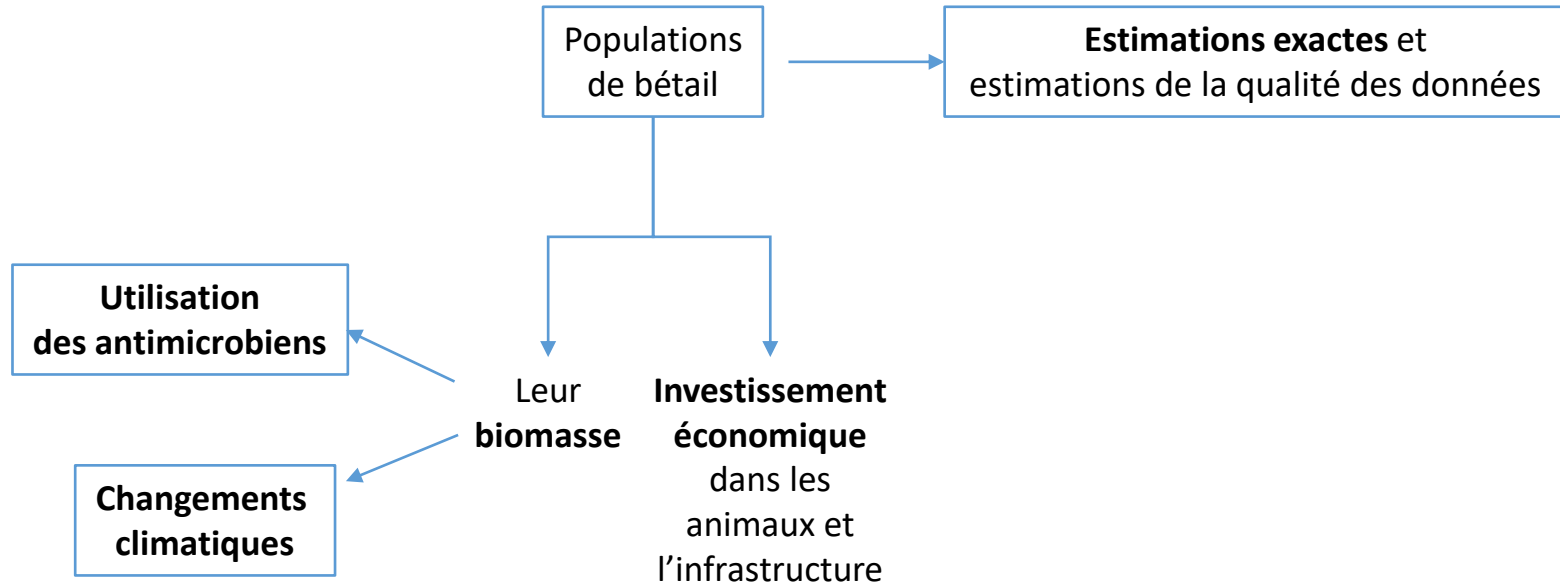
Rushton *et al.* 2021

GBADs - Structure analytique

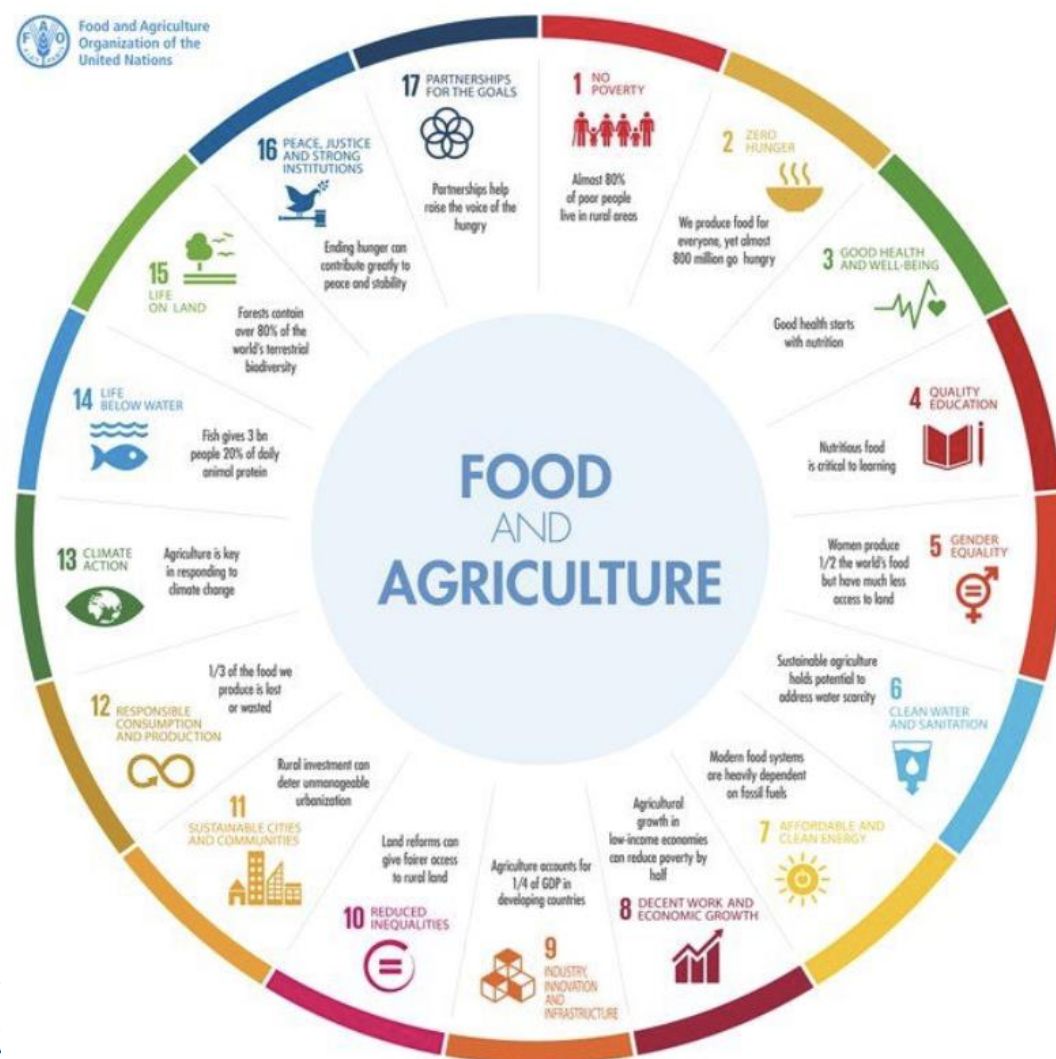


Rushton *et al.* 2021

GBADs - Structure analytique



Rushton *et al.* 2021





GBD20







GBADs



GBADS

GBD 20

 Global Burden
of Crop Loss



GBADs