

Green Data for Health

Data papers du Challenge Green Data for Health

Base de Connaissances sur les Produits PhytoPharmaceutiques à partir de sources Ouvertes (C3PO) – V1.0

Mots clés

Santé environnementale, Pesticides, Substances actives, Produits phytopharmaceutiques, Ventes, Usages, Agronomie, Agriculture, Toxicité, Écotoxicité, Réglementation, Base de données, API, Green Data for Health, Challenge Green Data for Health

Version du document

V2 (01/02/2024)

Version de C3PO

V1.0 (15/12/2023)

Auteur

Antoine Camus (Office Français de la Biodiversité)

Citation de ce document

Camus, A. (OFB). 2024. **Base de Connaissances sur les Produits PhytoPharmaceutiques à partir de sources Ouvertes (C3PO) – V1.0**. Data papers du Challenge Green Data for Health. Version 2.

SOMMAIRE

- Résumé
- Éléments de contexte et motivation
- Description des données
- Intérêt des données
- Collecte et traitement des données
- Plus-value et perspectives
- Bibliographie

Résumé

De nombreuses substances entrant dans la composition des produits phytopharmaceutiques (PPP) ont un impact présumé ou démontré sur la santé [1] et l'environnement (notamment déclin de la biodiversité [2, 3, 4], contamination des sols [5], des eaux souterraines et de surface [6]).

Pour mener des études s'intéressant à ces impacts, il est nécessaire de se baser, entre autres, sur des données de référence. Or, s'il existe en France et en Europe plusieurs sources de données relatives aux produits phytopharmaceutiques et aux pesticides, ces dernières, souvent gérées par des structures différentes, ne reposent pas systématiquement sur des référentiels communs. Ainsi dans la pratique il s'avère complexe de faire le lien entre les informations contenues dans ces différentes bases de données [7].

Dans le cadre du Challenge Green Data For Health, l'Office Français de la Biodiversité (OFB) a abordé cette problématique en réunissant au sein d'une base de connaissances sur les produits phytopharmaceutiques à partir de sources ouvertes (C3PO) diverses informations de référence sur les PPP et les pesticides issues des bases de données suivantes : BNV-D, Agritox, E-phy, référentiels "paramètres " et "groupes de paramètres" du Sandre, EU Pesticides Database, tableau d'entrées harmonisées disponible à l'annexe VI du règlement CLP (ATP). La liste des produits et substances contenus dans C3PO est issue d'une extraction de la dernière version en vigueur des données ouvertes de la BNV-D.

Les étapes suivies au cours du Challenge pour parvenir à la première version de la base C3PO ont été :

1. Élaboration d'un modèle et d'un dictionnaire de données décrivant le contenu de la base C3PO (schémas, tables et leurs attributs, relations entre tables)
2. Rédaction de scripts Notebook Jupyter (python) permettant d'automatiser la génération des données qui seront ensuite intégrées dans la base C3PO
3. Rédaction de scripts SQL permettant l'intégration des données générées par les scripts Notebook Jupyter dans la base de données

Grâce à ces réalisations, la base de données peut désormais être mise à jour automatiquement par déclenchement des scripts Notebook Jupyter et SQL. Il est désormais envisagé d'intégrer d'autres sources de données ouvertes dans C3PO et d'en diffuser les données par API.

Éléments de contexte et motivation

Contexte

Le Green Data for Health (GD4H) est une offre de service incubée au sein de l'ECOLAB, laboratoire d'innovation pour la transition écologique du Commissariat Général au Développement Durable (Ministère de la Transition écologique et de la Cohésion des territoires et Ministère de la Transition énergétique). Action phare du 4^o Plan National Santé Environnement, le GD4H vise à permettre une meilleure mobilisation et valorisation, par les acteurs de la recherche et de l'expertise, des données environnementales au service de la santé-environnement.

Les données environnementales existantes restent peu mobilisées pour faire émerger des associations entre expositions environnementales et pathologies humaines, le GD4H a donc lancé un Challenge permettant le développement d'outils ancrés dans la communauté de la donnée en santé-environnement afin d'adresser des problématiques partagées.

Dans le cadre de ce Challenge, l'Office Français de la Biodiversité (OFB) a porté un projet de création d'une base de connaissances sur les produits phytopharmaceutiques à partir de sources ouvertes (C3PO).

En effet, de nombreuses substances entrant dans la composition des produits phytopharmaceutiques (PPP) ont un impact présumé ou démontré sur la santé [1] et l'environnement (notamment déclin de la biodiversité [2, 3, 4], contamination des sols [5], des eaux souterraines et de surface [6]). Par ailleurs, pour mener des études s'intéressant à ces impacts, il est nécessaire de se baser, entre autres, sur des données de référence. Or, s'il existe en France et en Europe plusieurs sources de données relatives aux produits phytopharmaceutiques et aux pesticides, ces dernières, souvent gérées par des structures différentes, ne reposent pas systématiquement sur des référentiels communs. Ainsi dans la pratique il s'avère complexe de faire le lien entre les informations contenues dans ces différentes bases de données [7].

Dans le cadre du Challenge Green Data For Health, l'OFB a donc abordé cette problématique en réunissant au sein de la base C3PO des informations de référence sur les PPP et les pesticides issues des bases de données suivantes : BNV-D, Agritox, E-phy, référentiels "paramètres " et "groupes de paramètres" du Sandre, EU Pesticides Database, tableau d'entrées harmonisées disponible à l'annexe VI du règlement CLP.

Motivation

Parmi les sources de données sur les PPP, la BNV-D (Banque Nationale des Ventes de produits phytopharmaceutiques par les Distributeurs agréés), dont la diffusion des données est assurée par l'OFB, occupe une place particulière car elle est la meilleure base de données ouvertes disponible en France pour approximer la pression exercée par ces produits sur la santé et l'environnement. Les données de cette base proviennent des déclarations de ventes et achats de PPP, mais elles ne comportent pas d'informations relatives à certaines caractéristiques telles que données de

toxicité/écotoxicité, les usages, ou encore les fonctions des produits et des substances actives qui composent ces derniers.

Ces caractéristiques sont contenues dans d'autres sources de données ouvertes de référence aux niveaux français (Agritox, E-phy, référentiels "paramètres" et "groupes de paramètres" du Sandre) et européen (EU Pesticides Database, tableau d'entrées harmonisées disponible à l'annexe VI du règlement CLP).

La base de données C3PO a été créée à partir de l'enrichissement automatisé des données de la BNV-D avec ces autres sources de données. Le code source (scripts python de croisement entre les sources de données, script SQL de création de la base, dump SQL de la base de données) et la documentation attenante (notamment dictionnaire de données) sont hébergés et accessibles au public sur Gitlab : <https://gitlab.com/data-challenge-gd4h/c3po>.

Les fichiers plats (.csv) comportant les informations des tables de la base de données C3PO sont accessibles en ligne sur [data.gouv.fr](https://www.data.gouv.fr/fr/datasets/base-de-connaissances-sur-les-produits-phytopharmaceutiques-a-partir-de-sources-ouvertes-c3po/) : <https://www.data.gouv.fr/fr/datasets/base-de-connaissances-sur-les-produits-phytopharmaceutiques-a-partir-de-sources-ouvertes-c3po/>.

Il est désormais envisagé d'intégrer d'autres sources de données ouvertes dans C3PO et de rendre accessibles ses données via un service en ligne de type API.

L'objectif de ce *data paper* est de décrire ces données contenues dans la base C3PO et leur intérêt scientifique afin d'en susciter des réutilisations, et de détailler les conditions techniques et juridiques de cette réutilisation.

Description des données

Le modèle de données de la base C3PO est élaboré suivant une logique par source de données, il se décline en sept schémas :

- **bnvd**, comportant des informations relatives aux substances (années de vente, fonctions, classification RPD) et aux produits (type d'autorisation, - AMM ou PCP -, dont dispose le produit pour sa commercialisation en France, années de ventes, composition en substances, emploi autorisé dans les jardins) répertoriés dans les données ouvertes de la BNV-D ;
- **agritox**, comportant une table de jointure entre substances de Agritox et de la BNV-D, ainsi que des informations relatives à ces substances (propriétés physico-chimiques, toxicité, écotoxicité, catégories et mentions de danger) ;
- **atp**, comportant une table de jointure entre substances du tableau d'entrées harmonisées disponible à l'annexe VI du règlement CLP (ATP) et de la BNV-D , ainsi que des informations relatives à ces substances (identifiants, catégories et mentions de danger) ;
- **eupdb**, comportant une table de jointure entre substances de la base de données de l'Union Européenne sur les pesticides (EU Pesticides Database) et de la BNV-D, ainsi que des informations relatives à ces substances (statut d'autorisation dans les Etats-Membres de l'UE, dates de début et de fin d'approbation, autres indications relevant de la réglementation européenne : substances de base, candidates à substitution ou à faible risque) ;
- **sandre**, comportant une table de jointure entre substances du référentiel « Paramètres » du Sandre et de la BNV-D, ainsi que des informations relatives à ces substances (groupes de paramètres, famille et fonction auxquels elles sont affiliées) ;
- **ephy**, comportant :
 - une table de jointure entre substances de E-phy et de la BNV-D, ainsi que des informations relatives à ces substances (variants, état d'autorisation en France) ;
 - deux tables de jointure entre produits de E-phy (une pour les AMM et une pour les PCP) et de la BNV-D, ainsi que des informations relatives à ces produits (dénominations et type commerciaux, titulaire de l'AMM, détenteur du PCP, état d'autorisation, composition en substances, usages dont culture autorisées suivant la nomenclature E-phy ainsi que suivant la nomenclature du RPG, mentions de danger et phrases de risque) ;
- **c3po** : schéma comportant deux tables de métadonnées de la base C3PO

Ce modèle de données permet de faire le lien entre la BNV-D et chacune des sources de données via les tables de jointures.

Format et encodage

A chaque table correspond un fichier plat (format .csv). Chaque fichier est encodé en UTF-8 et dispose d'une ligne d'en-tête, le séparateur de colonnes est le point-virgule.

A noter

Les fichiers plats associés aux tables de chaque schéma sont accessibles sur data.gouv.fr.

Tableau 1 : Description des caractéristiques principales de chaque fichier (chaque fichier correspond à une table de la base de données), le nombre de lignes de chaque fichier correspond à celui de la version V1.0 de C3PO (date de publication 15/12/2023)

Nom du fichier	Schéma de la base de données	Source de données	Fréquence de mise à jour de la source de données	Description	Nombre de lignes	Variables (clés primaires indiquées par « * » et rappelées entre parenthèses lorsque clés étrangères de certaines tables, champs utilisés pour l'appariement entre la BNV-D et les sources de données en gras souligné)
bnvd_substance.csv	bnvd	Données ouvertes de la BNV-D (achats et ventes), version en vigueur Producteur : OFB	Annuelle (au dernier trimestre de l'année N+1 suivant la dernière année N de ventes de produits phytopharmaceutiques)	Liste et informations uniques (fonctions, première et dernière années de vente, dernière classification RPD connue) des substances répertoriées dans la BNV-D.	714	id_substance* substance_cas substance_cas fonction_principale fonction_precision fonction_autres annee_min annee_max rpd_classification_last rpd_mention_last rpd_taux_last
bnvd_ref_classification.csv	bnvd	Données ouvertes de la BNV-D (achats et ventes), version en vigueur Producteur : OFB	Annuelle (au dernier trimestre de l'année N+1 suivant la dernière année N de ventes de produits phytopharmaceutiques)	Informations annuelles des classifications et mentions RPD de chaque substance de la BNV-D.	83	id_classif* annee rpd_classification rpd_mention rpd_classification_taux rpd_mention_taux rpd_taux
bnvd_substance_classification.csv	bnvd	Données ouvertes de la BNV-D	Annuelle (au dernier trimestre de l'année N+1 suivant la dernière année	Table de jointure entre bnvd_substance et bnvd_ref_classification.	6 902	id_substance (clé primaire associée: id_substance de la table bnvd_substance) id_classif (clé primaire associée: id_classif de la table bnvd_ref_classification)

Green Data for Health
Data papers Challenge Green Data for Health

Nom du fichier	Schéma de la base de données	Source de données	Fréquence de mise à jour de la source de données	Description	Nombre de lignes	Variables (clés primaires indiquées par « * » et rappelées entre parenthèses lorsque clés étrangères de certaines tables, champs utilisés pour l'appariement entre la BNV-D et les sources de données en gras souligné)
		(achats et ventes), version en vigueur Producteur : OFB	N de ventes de produits phytopharmaceutiques)			
bnvd_amm.csv	bnvd	Données ouvertes de la BNV-D (achats et ventes), version en vigueur Producteur : OFB	Annuelle (au dernier trimestre de l'année N+1 suivant la dernière année N de ventes de produits phytopharmaceutiques)	Liste et informations uniques (première et dernière années de vente, conditionnement) des produits ayant disposé ou disposant d'une AMM et répertoriés dans la BNV-D.	4 882	amm* annee_min annee_max conditionnement
bnvd_amm_evol_composition.csv	bnvd	Données ouvertes de la BNV-D (achats et ventes), version en vigueur Producteur : OFB	Annuelle (au dernier trimestre de l'année N+1 suivant la dernière année N de ventes de produits phytopharmaceutiques)	Informations annuelles associées à chaque produit ayant disposé ou disposant d'une AMM et répertorié au moins une fois dans l'historique de ventes de la BNV-D (composition en substances.	51 586	amm (clé primaire associée: amm de la table bnvd_amm) annee eaj id_substance
bnvd_pcp.csv	bnvd	Données ouvertes de la BNV-D (achats et ventes), version en vigueur Producteur : OFB	Annuelle (au dernier trimestre de l'année N+1 suivant la dernière année N de ventes de produits phytopharmaceutiques)	Liste et informations uniques (première et dernière années de vente, conditionnement) des produits ayant disposé ou disposant d'un PCP et répertoriés dans la BNV-D.	969	pcp* annee_min annee_max conditionnement
bnvd_pcp_evol_composition.csv	bnvd	Données ouvertes de la BNV-D (achats et ventes), version en vigueur Producteur : OFB	Annuelle (au dernier trimestre de l'année N+1 suivant la dernière année N de ventes de produits phytopharmaceutiques)	Informations annuelles associées à chaque produit ayant disposé ou disposant d'un PCP et répertorié au moins une fois dans l'historique de ventes de la BNV-D (composition en substances.	5 972	pcp (clé primaire associée: pcp de la table bnvd_pcp) annee eaj id_substance
sandre_substance_bnvd.csv	sandre	Référentiel « Paramètres » du Sandre	Continue	Table de jointure entre substances de la BNV-D et	714	id_substance_bnvd (clé primaire associée: id_substance de la table bnvd_substance)

Green Data for Health
Data papers Challenge Green Data for Health

Nom du fichier	Schéma de la base de données	Source de données	Fréquence de mise à jour de la source de données	Description	Nombre de lignes	Variables (clés primaires indiquées par « * » et rappelées entre parenthèses lorsque clés étrangères de certaines tables, champs utilisés pour l'appariement entre la BNV-D et les sources de données en gras souligné)
		Producteur : OIEau		substances du référentiel « Paramètres » du Sandre.		id_substance_sandre (clé primaire associée : cdparametre de la table sandre_parametres)
sandre_parametres.csv	sandre	Référentiel « Paramètres » du Sandre Producteur : OIEau	Continue	Liste et informations uniques (code sandre, libellé et numéro CAS) des substances du référentiel « Paramètres » du Sandre disposant d'un numéro CAS.	3 002	cdparametre* nomparametre cas
sandre_gpes_parametres.csv	sandre	Référentiel « Groupes de paramètres » du Sandre Producteur : OIEau	Continue	Liste et informations uniques (code sandre, libellé, famille chimique et fonction) des substances du référentiel « Groupes de Paramètres » pour les paramètres Sandre disposant d'un numéro CAS et affiliées au groupe de paramètres de code Sandre « 95 » (Phytoprotecteurs).	135	cdgroupeparametres* nomgroupeparametres famille fonction
sandre_parametres_gpes_parametres.csv	sandre	Référentiels « Paramètres » et « Groupes de paramètres » du Sandre Producteur : OIEau	Continue	Table de jointure entre les tables sandre_parametres et sandre_gpes_parametres.	8 493	id_substance_bnvd (clé primaire associée : cdparametre de la table sandre_parametres) cdgroupeparametres (clé primaire associée : cdgroupeparametre de la table sandre_gpes_parametres)
eupdb_substance_bnvd.csv	eupdb	Base de données de l'Union Européenne sur les pesticides (EU PDB) Producteur : Union Européenne	Continue	Table de jointure entre substances de la BNV-D et substances de la EU PDB.	714	id_substance_bnvd (clé primaire associée : id_substance de la table bnvd_substance) id_substance_eupdb (clé primaire associée : id_substance de la table eupdb_substance)
eupdb_substance.csv	eupdb	Base de données de l'Union Européenne sur les	Continue	Liste et informations uniques (statut d'autorisation dans les Etats-Membres de l'UE,	1 481	id_substance* substance_eupdb etat_reg_1107_2019 date_approbation

Green Data for Health
Data papers Challenge Green Data for Health

Nom du fichier	Schéma de la base de données	Source de données	Fréquence de mise à jour de la source de données	Description	Nombre de lignes	Variables (clés primaires indiquées par « * » et rappelées entre parenthèses lorsque clés étrangères de certaines tables, champs utilisés pour l'appariement entre la BNV-D et les sources de données en gras souligné)
		pesticides (EU PDB) Producteur : Union Européenne		dates de début et de fin d'approbation, autres indications relevant de la réglementation européenne : substances de base, candidates à substitution ou à faible risque) des substances de la EU PDB.		date_expiration_approbation substance_base candidat_substitution faible_risque autorisations_pays autorisation_france
agritox_substance_bnvd.csv	agritox	Base de données Agritox Producteur : Anses	Bimestrielle	Table de jointure entre substances de la BNV-D et substances de Agritox.	714	id_substance_bnvd (clé primaire associée: id_substance de la table bnvd_substance) id_substance_agritox (clé primaire associée: id_substance de la table agritox_substance)
agritox_substance.csv	agritox	Base de données Agritox Producteur : Anses	Bimestrielle	Liste et informations uniques (dénomination, identifiants et formule chimique, masse molaire, pureté) des substances de Agritox.	277	substance* substance cas substance cas nom_uipac ec masse_molaire formule_brute purete_minimale impuretes_pertinentes
agritox_substance_classement.csv	agritox	Base de données Agritox Producteur : Anses	Bimestrielle	Classement (classification, catégories et mentions de danger) des substances de Agritox.	871	id_substance (clé primaire associée: id_substance de la table agritox_substance) classement_ref classement_date classement_code_h classement_categorie_danger classement_mention_danger facteur_m_valeur facteur_m_origine facteur_m_date classification_simpl
agritox_substance_proprietes_pc.csv	agritox	Base de données Agritox Producteur : Anses	Bimestrielle	Propriétés physico-chimiques des substances de Agritox.	2 027	

Green Data for Health
Data papers Challenge Green Data for Health

Nom du fichier	Schéma de la base de données	Source de données	Fréquence de mise à jour de la source de données	Description	Nombre de lignes	Variables (clés primaires indiquées par « * » et rappelées entre parenthèses lorsque clés étrangères de certaines tables, champs utilisés pour l'appariement entre la BNV-D et les sources de données en gras souligné)
agritox_substance_toxicite.csv	agritox	Base de données Agritox Producteur : Anses	Bimestrielle	Données de toxicité des substances de Agritox.	1 070	id_substance (clé primaire associée: id_substance de la table agritox_substance) nature_etude dl50_cl50 espece sexe valeur unite
agritox_substance_ecotoxicite.csv	agritox	Base de données Agritox Producteur : Anses	Bimestrielle	Données d'écotoxicité des substances de Agritox.	282	id_substance (clé primaire associée: id_substance de la table agritox_substance) etudes donnee_toxicite donnee_toxicite_valeur donnee_toxicite_unite facteur_securite valeur_pnec unite_pnec
agritox_substance_vtr.csv	agritox	Base de données Agritox Producteur : Anses	Bimestrielle	Valeurs toxicologiques de référence des substances de Agritox.	876	id_substance (clé primaire associée: id_substance de la table agritox_substance) nom valeur unite source date etude_pivot_1 etude_pivot_1_noael etude_pivot_1_noael_unite etude_pivot_2 facteur_securite facteur_correction_abs_orale_incompl
atp_substance_bnvd.csv	atp	Tableau d'entrées harmonisées disponible à l'annexe VI du CLP (ATP)	Annuelle	Table de jointure entre substances de la BNV-D et substances de l'ATP.	716	id_substance_bnvd (clé primaire associée: id_substance de la table bnvd_substance) id_substance_atp (clé primaire associée: id_substance de la table atp_substance)

Green Data for Health
Data papers Challenge Green Data for Health

Nom du fichier	Schéma de la base de données	Source de données	Fréquence de mise à jour de la source de données	Description	Nombre de lignes	Variables (clés primaires indiquées par « * » et rappelées entre parenthèses lorsque clés étrangères de certaines tables, champs utilisés pour l'appariement entre la BNV-D et les sources de données en gras souligné)
		Producteur : Commission Européenne				
atp_substance.csv	atp	Tableau d'entrées harmonisées disponible à l'annexe VI du CLP (ATP) Producteur : Commission Européenne	Annuelle	Liste et informations uniques (dénomination) des substances de l'ATP.	4 742	id_substance* index ici ec cas
atp_substance_classement.csv	atp	Tableau d'entrées harmonisées disponible à l'annexe VI du CLP (ATP) Producteur : Commission Européenne	Annuelle	Classement (classification, catégories et mentions de danger) des substances de l'ATP.	20 508	id_substance (clé primaire associée : id_substance de la table atp_substance) code_h_detail code_h categorie_danger mention_danger classification_simplifiee
ephy_substance_bnvd.csv	ephy	Base de données E-phy Producteur : Anses	Hebdomadaire	Table de jointure entre substances de la BNV-D et substances de E-phy.	740	id_substance_bnvd id_substance_ephy
ephy_substance.csv	ephy	Base de données E-phy Producteur : Anses	Hebdomadaire	Liste et informations uniques (identifiants, liste de variants, état d'autorisation) des substances de E-phy. Remarque : la liste des noms de variants est stockée dans un même et unique champ, en cas d'existence de plusieurs noms de variants pour une même substance, ils sont séparés par « ».	1 326	id_substance* substance_cas_variant substance cas variant etat_autorisation
ephy_substance_variants.csv	ephy	Base de données E-phy Producteur : Anses	Hebdomadaire	Détail des variants i.e. des autres noms possibles pour la substance d'après E-phy, un variant par ligne.	1 747	id_substance variant exp cas variant

Green Data for Health
Data papers Challenge Green Data for Health

Nom du fichier	Schéma de la base de données	Source de données	Fréquence de mise à jour de la source de données	Description	Nombre de lignes	Variables (clés primaires indiquées par « * » et rappelées entre parenthèses lorsque clés étrangères de certaines tables, champs utilisés pour l'appariement entre la BNV-D et les sources de données en gras souligné)
				Remarque: les informations contenues dans cette table ont été obtenues après extraction des modalités de variants séparées par « » dans le champ 'variant' de la table ephy_substance.		
ephy_amm_bnvd.csv	ephy	Base de données E-phy Producteur: Anses	Hebdomadaire	Table de jointure entre produits de la BNV-D et de E-phy disposant d'une AMM.	4 230	amm_bnvd (clé primaire associée: amm de la table bnvd_amm) amm_ephy (clé primaire associée: amm de la table ephy_amm)
ephy_pcp_bnvd.csv	ephy	Base de données E-phy Producteur: Anses	Hebdomadaire	Table de jointure entre produits de la BNV-D et de E-phy disposant d'un PCP.	969	pcp_bnvd (clé primaire associée: pcp de la table bnvd_pcp) pcp_ephy (clé primaire associée: pcp de la table ephy_pcp)
ephy_amm.csv	ephy	Base de données E-phy Producteur: Anses	Hebdomadaire	Liste et informations uniques (identifiants, titulaire de l'AMM, type de produit, type commercial, état d'autorisation) des produits de E-phy disposant d'une AMM.	14 697	amm* nom_produit titulaire type_commercial type_produit etat_autorisation date_retrait date_premiere_autorisation produit_ref_amm produit_ref_nom
ephy_pcp.csv	ephy	Base de données E-phy	Hebdomadaire	Liste et informations uniques (identifiants, détenteur du PCP, type de produit, type commercial, état d'autorisation) des produits de E-phy disposant d'un PCP.	1 894	pcp* nom_produit détenteur_pcp type_commercial type_produit etat_autorisation produit_ref_amm (clé primaire associée: amm de la table ephy_amm) produit_ref_nom
ephy_pcp_origine.csv	ephy	Base de données E-phy Producteur: Anses	Hebdomadaire	Informations relatives à l'origine des PCP.	2 590	pcp (clé primaire associée: pcp de la table ephy_pcp) amm_produit_importe nom_produit_importe etat_membre_origine

Green Data for Health
Data papers Challenge Green Data for Health

Nom du fichier	Schéma de la base de données	Source de données	Fréquence de mise à jour de la source de données	Description	Nombre de lignes	Variables (clés primaires indiquées par « * » et rappelées entre parenthèses lorsque clés étrangères de certaines tables, champs utilisés pour l'appariement entre la BNV-D et les sources de données en gras souligné)
ephy_amm_second_nom.csv	ephy	Base de données E-phy Producteur : Anses	Hebdomadaire	Seconds noms éventuels des produits de E-phy disposant d'une AMM.	3 000	amm (clé primaire associée: amm de la table ephy_amm) second_nom_produit
ephy_amm_composition.csv	ephy	Base de données E-phy Producteur : Anses	Hebdomadaire	Informations relatives à la composition en substances des produits de E-phy disposant d'une AMM.	21 865	amm (clé primaire associée: amm de la table ephy_amm) id_substance (clé primaire associée: id_substance de la table ephy_substance) variant_exp teneur_substance teneur_substance_unite
ephy_amm_usage.csv	ephy	Base de données E-phy Producteur : Anses	Hebdomadaire	Informations relatives aux usages des produits de E-phy disposant d'une AMM.	92 194	id_usage* amm (clé primaire associée: amm de la table ephy_amm) usage_gamme mentions_autorisees usage_restrictions_court usage_restrictions_long fonctions formulations usage usage_date_decision stade_cultural_min_bbch stade_cultural_max_bbch usage_etat dose_homologuee dose_homologuee_unite dar_jours dar_bbch nb_max_applications usage_date_fin_distribution usage_date_fin_utilisation conditions_emploi znt_aquatique znt_arthropodes_non_cibles znt_plantes_non_cibles intervalle_min

Green Data for Health
Data papers Challenge Green Data for Health

Nom du fichier	Schéma de la base de données	Source de données	Fréquence de mise à jour de la source de données	Description	Nombre de lignes	Variables (clés primaires indiquées par « * » et rappelées entre parenthèses lorsque clés étrangères de certaines tables, champs utilisés pour l'appariement entre la BNV-D et les sources de données en gras souligné)
ephy_ref_culture_rpg.csv	ephy	Base de données E-phy Producteur: Anses	Hebdomadaire	Nomenclature des cultures et groupes de cultures du RPG (version 2.1) .	168	id_culture* culture code_gpe_cult_rpg libelle_gpe_cult_rpg
ephy_usage_culture.csv	ephy	Base de données E-phy Producteur: Anses	Hebdomadaire	Table de jointure entre les usages des produits d'après E-phy (ephy_amm_usage.csv) et les cultures suivant la nomenclature du RPG (ephy_ref_culture_rpg.csv).	110 241	id_usage (clé primaire associée: id_usage de la table ephy_amm_usage) id_culture (clé primaire associée: id_culture de la table ephy_ref_culture_rpg)
ephy_amm_mention_danger.csv	ephy	Base de données E-phy Producteur: Anses	Hebdomadaire	Informations relatives aux mentions de danger associées aux produits de E-phy disposant d'une AMM.	10 002	amm (clé primaire associée: amm de la table ephy_amm) classes_mention_danger_lb_court classes_mention_danger_lb_long
ephy_amm_phrase_risque.csv	ephy	Base de données E-phy Producteur: Anses	Hebdomadaire	Informations relatives aux phrases de risque associées aux produits de E-phy disposant d'une AMM	8 494	amm (clé primaire associée: amm de la table ephy_amm) phrases_risque_lb_court phrases_risque_lb_long
c3po_source.csv	c3po	/	A chaque mise à jour de C3PO	Table comportant les métadonnées relatives à chaque source de données intégrées à la version correspondante de la base C3PO	7	source date_publication date_collecte
c3po_version.csv	c3po	/	A chaque mise à jour de C3PO	Table comportant les métadonnées relatives à la version correspondante de C3PO	1	version date_publication

Les relations entre tables sont indiquées dans le modèle physique de données (Figure 1 : Modèle physique de données de C3PO (V2)Figure 1).

Green Data for Health
Data papers Challenge Green Data for Health

Tableau 2 : Dictionnaire des variables

Fichier	Champ	Type	Description
bnvd_substance.csv	id_substance	Texte	Clé du type #_sub_bnvd (# étant un nombre).
bnvd_substance.csv	substance_cas	Texte	Concaténation des champs "Substance" et "N°CAS" de la BNV-D séparés par un <i>underscore</i> ("_").
bnvd_substance.csv	substance	Texte	Libellé de la substance d'après la BNV-D (libellé uniquement en minuscules et sans accents).
bnvd_substance.csv	cas	Texte	N° CAS de la substance d'après la BNV-D.
bnvd_substance.csv	fonction_principale	Texte	Fonction principale de la substance attribuée par le GT "mise en qualité des données de la BNV-D".
bnvd_substance.csv	fonction_precision	Texte	Précision éventuelle sur la fonction principale de la substance attribuée par le GT "mise en qualité des données de la BNV-D".
bnvd_substance.csv	fonction_autres	Texte	Autres fonctions possibles pour la substance attribuées par le GT "mise en qualité des données de la BNV-D".
bnvd_substance.csv	annee_min	Entier	Première année de vente de la substance.
bnvd_substance.csv	annee_max	Entier	Dernière année de vente de la substance.
bnvd_substance.csv	rpd_classification_last	Texte	Dernière modalité de classification RPD en vigueur pour la substance (en fonction de la dernière année de vente).
bnvd_substance.csv	rpd_mention_last	Texte	Dernière modalité de mention RPD en vigueur pour la substance (en fonction de la dernière année de vente).
bnvd_substance.csv	rpd_taux_last	Double	Dernier taux de redevance en vigueur pour la substance (en fonction de la dernière année de vente). Unité : €/kg (de substance vendue)
bnvd_ref_classification.csv	id_classif	Texte	Clé du type #_classif (# étant un nombre).
bnvd_ref_classification.csv	annee	Entier(4)	Années durant lesquelles la substance a été vendue en France.
bnvd_ref_classification.csv	rpd_classification	Texte	Classement de la substance au titre de la redevance pour pollution diffuse : <u>Nomenclature applicable pour la période 2008 – 2018 :</u> <ul style="list-style-type: none"> - N minéral : substance minérale dangereuse pour l'environnement - N Organique : substance organique dangereuse pour l'environnement - T, T+, CMR : substance toxique, très toxique, cancérogène, mutagène, toxique pour la reproduction - Autre : autre substance <u>Nomenclature applicable à partir de 2019 :</u> <ul style="list-style-type: none"> - CMR : cancérogénicité, mutagénicité sur les cellules germinales ou toxicité pour la reproduction

Green Data for Health
Data papers Challenge Green Data for Health

Fichier	Champ	Type	Description
			<ul style="list-style-type: none"> - Santé A : toxicité aiguë de catégorie 1, 2 ou 3, ou toxicité spécifique pour certains organes cibles, de catégorie 1, à la suite d'une exposition unique ou après une exposition répétée, soit en raison de leurs effets sur ou via l'allaitement - Env A : toxicité aiguë pour le milieu aquatique de catégorie 1 ou toxicité chronique pour le milieu aquatique de catégorie 1 ou 2 - Env B : toxicité chronique pour le milieu aquatique de catégorie 3 ou 4 - Autre : autre substance
bnvd_ref_classification.csv	rpd_mention	Texte	<p><u>Nomenclature applicable à partir de 2019 :</u></p> <ul style="list-style-type: none"> - Exclusion : substance ne répondant pas aux critères des paragraphes 3.6 et 3.7 de l'annexe II au règlement (CE) n° 1107/2009 - Substitution : substance dont on envisage la substitution au sens de l'article 24 du règlement (CE) n° 1107/2009
bnvd_ref_classification.csv	rpd_classification_taux	Double	<p>Composante de la redevance liée à la classification pour le calcul du taux de redevance pour pollutions diffuses (RPD) applicable à la substance pour l'année considérée.</p> <p>Unité : €/kg (de substance vendue)</p>
bnvd_ref_classification.csv	rpd_mention_taux	Double	<p>Composante de la redevance liée à la mention pour le calcul du taux de redevance pour pollutions diffuses (RPD) applicable à la substance pour l'année considérée.</p> <p>Unité : €/kg (de substance vendue)</p>
bnvd_ref_classification.csv	rpd_taux	Double	<p>Taux de redevance pour pollutions diffuses (RPD) applicable à la substance pour l'année considérée.</p> <p>Unité : €/kg (de substance vendue)</p>
bnvd_substance_classification.csv	id_substance	Texte	Clé du type #_sub_bnvd (# étant un nombre).
bnvd_substance_classification.csv	id_classif	Texte	Clé du type #_classif (# étant un nombre).
bnvd_amm.csv	amm	Texte	Numéro d'autorisation de mise sur le marché (AMM) du produit phytopharmaceutique.
bnvd_amm.csv	annee_min	Entier	Première année de vente de l'AMM.
bnvd_amm.csv	annee_max	Entier	Dernière année de vente de l'AMM.
bnvd_amm.csv	conditionnement	Texte	<p>Unité de conditionnement de l'AMM.</p> <p>Modalités possibles : "l" (litre) ; "kg" (kilogramme)</p>
bnvd_amm_evol_composition.csv	amm	Texte	Numéro d'autorisation de mise sur le marché (AMM) du produit phytopharmaceutique.
bnvd_amm_evol_composition.csv	annee	Entier(4)	Années de ventes de l'AMM.
bnvd_amm_evol_composition.csv	eaj	Texte	<p>AMM avec emploi autorisé dans les jardins.</p> <p>Modalités possibles : "Oui" ; "Non"</p>

Green Data for Health
Data papers Challenge Green Data for Health

Fichier	Champ	Type	Description
bnvd_amm_evol_composition.csv	id_substance	Texte	Clé du type #_sub_bnvd (# étant un nombre), permettant d'identifier les substances entrant dans la composition de l'AMM.
bnvd_pcp.csv	pcp	Texte	Numéro de Permis de Commerce Parallèle (PCP) du produit phytopharmaceutique. NB : Dans la BNV-D, les distributeurs ne sont pas censés déclarer de PCP (uniquement des AMM), néanmoins des erreurs de déclaration sont constatées, avec des n°PCP déclarés
bnvd_pcp.csv	annee_min	Entier	Première année de vente du produit PCP.
bnvd_pcp.csv	annee_max	Entier	Dernière année de vente du produit PCP.
bnvd_pcp.csv	conditionnement	Texte	Unité de conditionnement du produit PCP. Modalités possibles : "l" (litre) ; "kg" (kilogramme)
bnvd_pcp_evol_composition.csv	pcp	Texte	Numéro de Permis de Commerce Parallèle (PCP) du produit phytopharmaceutique.
bnvd_pcp_evol_composition.csv	annee	Entier(4)	Années de ventes du PCP.
bnvd_pcp_evol_composition.csv	eaj	Texte	PCP avec autorisé dans les jardins. Modalités possibles : "Oui" ; "Non"
bnvd_pcp_evol_composition.csv	id_substance	Texte	Clé du type #_sub_bnvd (# étant un nombre), permettant d'identifier les substances entrant dans la composition du PCP.
ephy_substance.csv	id_substance	Texte	Clé du type #_sub_ephy (# étant un nombre)
ephy_substance.csv	substance_cas_variant	Texte	Concaténation des champs "Substance", "N°CAS" et "Variant" de E-phy séparés par un <i>underscore</i> ("_") permettant d'attribuer un identifiant unique aux substances recensées dans E-phy.
ephy_substance.csv	substance	Texte	Libellé de la substance d'après E-phy.
ephy_substance.csv	cas	Texte	N° CAS de la substance d'après E-phy.
ephy_substance.csv	variant	Texte	Autres noms possibles pour la substance d'après E-phy (variants séparés par " ").
ephy_substance.csv	etat_autorisation	Texte	Etat d'autorisation de la substance d'après E-phy. Modalités possibles : "NON_INSCRITE";"INSCRITE";"AUTRE_CAS";"INSCRIPTION_EN_COURS"
ephy_substance_bnvd.csv	id_substance_bnvd	Texte	Clé du type #_sub_bnvd (# étant un nombre).
ephy_substance_bnvd.csv	id_substance_ephy	Texte	Clé du type #_sub_ephy (# étant un nombre).
ephy_substance_variants.csv	id_substance	Texte	Clé du type #_sub_ephy (# étant un nombre).
ephy_substance_variants.csv	variant_exp	Texte	Autres noms possibles pour la substance d'après E-phy, un variant par ligne (extraction de chaque modalité entre les « » des enregistrements du champ 'variant' de la table ephy.substance).
ephy_substance_variants.csv	cas_variant	Texte	N°CAS éventuel associé au(x) variant(s) dans E-phy.
ephy_amm.csv	amm	Texte	N°AMM (E-phy) du produit.

Green Data for Health
Data papers Challenge Green Data for Health

Fichier	Champ	Type	Description
ephy_amm.csv	nom_produit	Texte	Nom commercial associé au N°AMM.
ephy_amm.csv	titulaire	Texte	Titulaire détenteur de l'autorisation de mise sur le marché (AMM).
ephy_amm.csv	type_commercial	Texte	Type commercial du produit. Modalités possibles : "Générique" ; "Produit de référence" ; "Deuxième gamme" ; "Produit de revente"
ephy_amm.csv	type_produit	Texte	Type de produit. Modalités possibles : "PRODUIT-MIXTE" ; "MELANGE" ; "MFSC" ; "PPP" ; "ADJUVANT"
ephy_amm.csv	etat_autorisation	Texte	Indication si l'AMM est autorisé (à date d'extraction de la base E-phy). Modalités possibles : "AUTORISE" ; "RETIRE"
ephy_amm.csv	date_retrait	Date	Date de retrait de l'AMM (format JJ/MM/AAAA), pour les AMM pour lesquels etat_autorisation="RETIRE".
ephy_amm.csv	date_premiere_autorisation	Date	Date de première autorisation de l'AMM (format JJ/MM/AAAA), pour les AMM pour lesquels celle-ci est connue.
ephy_amm.csv	produit_ref_amm	Texte	AMM française de référence auquel se rapporte éventuellement l'AMM.
ephy_amm.csv	produit_ref_nom	Texte	Nom associé à l'AMM de référence.
ephy_amm_bnvd.csv	amm_bnvd	Texte	Numéro d'autorisation de mise sur le marché (AMM) du produit phytopharmaceutique dans la BNV-D.
ephy_amm_bnvd.csv	amm_ephy	Texte	N°AMM (E-phy) du produit.
ephy_pcp.csv	pcp	Texte	N°PCP (E-phy) du produit.
ephy_pcp.csv	nom_produit	Texte	Nom commercial associé au N°PCP.
ephy_pcp.csv	detenteur_pcp	Texte	Titulaire détenteur du permis de commerce parallèle (PCP).
ephy_pcp.csv	type_commercial	Texte	Type commercial du produit. Modalités possibles : "Produit de commerce parallèle"
ephy_pcp.csv	type_produit	Texte	Type de produit. Modalités possibles : "PRODUIT-MIXTE" ; "MELANGE" ; "MFSC" ; "PPP" ; "ADJUVANT"
ephy_pcp.csv	etat_autorisation	Texte	Indication si le PCP est autorisé (à date d'extraction de la base E-phy). Modalités possibles : "AUTORISE" ; "RETIRE"
ephy_pcp.csv	produit_ref_amm	Texte	AMM française de référence à laquelle se rapporte le PCP.
ephy_pcp.csv	produit_ref_nom	Texte	Nom associé à l'AMM de référence.
ephy_pcp_origine.csv	pcp	Texte	N°PCP (E-phy) du produit.
ephy_pcp_origine.csv	amm_produit_importe	Texte	N°AMM attribuée par le pays d'origine au PCP.
ephy_pcp_origine.csv	nom_produit_importe	Texte	Nom commercial attribué par le pays d'origine au PCP.

Green Data for Health
Data papers Challenge Green Data for Health

Fichier	Champ	Type	Description
ephy_pcp_origine.csv	etat_membre_origine	Texte	Pays (état-membre de l'UE) d'origine du PCP.
ephy_pcp_bnvd.csv	pcp_bnvd	Texte	Numéro de Permis de Commerce Parallèle (PCP) du produit phytopharmaceutique dans la BNV-D.
ephy_pcp_bnvd.csv	pcp_ephy	Texte	N°PCP (E-phy) du produit.
ephy_amm_second_nom.csv	amm	Texte	N°AMM (E-phy) du produit.
ephy_amm_second_nom.csv	second_nom_produit	Texte	Seconds noms commerciaux associés au N°AMM.
ephy_amm_composition.csv	amm	Texte	N°AMM (E-phy) du produit.
ephy_amm_composition.csv	id_substance	Texte	Clé du type #_sub_ephy (# étant un nombre).
ephy_amm_composition.csv	variant_exp	Texte	Autres noms possibles pour la substance d'après E-phy, un variant par ligne (extraction de chaque modalité entre les « » des enregistrements du champ 'variant' de la table ephy.substance).
ephy_amm_composition.csv	teneur_substance	Double	Teneur de la substance active dans l'AMM.
ephy_amm_composition.csv	teneur_substance_unite	Texte	Unité de la teneur de la substance active dans l'AMM.
ephy_amm_usage.csv	id_usage	Texte	Clé du type #_usage (# étant un nombre).
ephy_amm_usage.csv	amm	Texte	N°AMM (E-phy) du produit.
ephy_amm_usage.csv	usage_gamme	Texte	Gamme d'usage de l'AMM. Modalités possibles : "Amateur / emploi autorisé dans les jardins" ; "Professionnel"
ephy_amm_usage.csv	mentions_autorisees	Texte	Mention(s) autorisée(s) pour l'AMM. Modalités possibles : "Production biologique amateur" ; "Produit à faible risque article 47" ; "Liste biocontrôle" ; "Emploi autorisé dans les jardins" ; " Utilisable en agriculture biologique"
ephy_amm_usage.csv	usage_restrictions_court	Texte	Restriction(s) d'usage potentielle(s) (libellé court) pour l'AMM. Modalités possibles : "Restriction jardins amateurs" ; "Restriction néonicotinoïdes"
ephy_amm_usage.csv	usage_restrictions_long	Texte	Restriction(s) d'usage potentielle(s) (libellé long) pour l'AMM. Chaque modalité est obligatoirement associée à une valeur du champ <i>usage_restrictions_court</i> , il s'agit des suivantes : <ul style="list-style-type: none"> • Pour usage_restrictions_court = "Restriction jardins amateurs" : "A compter du 01/01/2019, la mise sur le marché, la délivrance, l'utilisation et la détention de ce produit sont réservées aux utilisateurs professionnels, conformément à l'article L. 253-7-III du Code rural et de la pêche maritime" • Pour usage_restrictions_court = "Restriction néonicotinoïdes" : - Modalité 1 : "A compter du 01/09/2018, l'utilisation de ce produit contenant une substance de la famille des néonicotinoïdes est interdite pour les usages qui ne font pas l'objet d'une dérogation par arrêté ministériel, conformément à l'article L. 253-8 du Code rural et de la pêche maritime"

Green Data for Health
Data papers Challenge Green Data for Health

Fichier	Champ	Type	Description
			- Modalité 2 : "A compter du 01/09/2018, l'utilisation de ce produit contenant une substance de la famille des néonicotinoïdes est interdite, conformément à l'article L. 253-8 du Code rural et de la pêche maritime"
ephy_amm_usage.csv	fonctions	Texte	Fonction(s) rattachée(s) à l'AMM.
ephy_amm_usage.csv	formulations	Texte	Formulation(s) de l'AMM.
ephy_amm_usage.csv	usage	Texte	<p>Intitulé de l'usage conformément au catalogue des usages (référentiel national sur la base duquel les AMM des produits sont délivrées, et apportant les informations pratiques sur la portée des différents usages en termes de cultures et de bio-agresseurs). Les usages phytopharmaceutiques sont définis en opérant différents niveaux de regroupements de cultures et/ou de cibles. On peut avoir des usages définis :</p> <ul style="list-style-type: none"> - par un groupe de cultures et une fonction : exemple « Céréales * désherbage » ; - par un groupe de cultures, un mode d'application et une fonction : exemple « Céréales*traitement des semences*insecticides » ; - par un groupe de culture plus réduit, un mode d'application et un groupe de bio-agresseurs : exemple « pommier (qui inclut pommettes, poirier, nashi, cognassier et nêfle)* traitement des parties aériennes * chenilles foreuses des fruits (qui inclut différentes espèces de carpocapse et la teigne des pommes) ». <p>Les libellés des usages sont « simplifiés ». Ainsi, sous un terme spécifique, plusieurs espèces végétales peuvent être couvertes : on parle alors de la portée de l'usage. Par exemple, dans l'usage « cassissier » sont inclus : le cassissier, le myrtillier, le groseillier, le sureau noir, la mûre (Morus sp.), l'airelle, le cynhorodon, l'azerolier.</p>
ephy_amm_usage.csv	usage_date_decision	Date	Date de décision de l'autorisation de l'usage (format JJ/MM/AAAA).
ephy_amm_usage.csv	stade_cultural_min_bbch	Entier	Stade cultural minimal de l'échelle BBCH pour lequel l'usage est autorisé.
ephy_amm_usage.csv	stade_cultural_max_bbch	Entier	Stade cultural maximal de l'échelle BBCH pour lequel l'usage est autorisé.
ephy_amm_usage.csv	usage_etat	Texte	<p>Etat d'autorisation de l'usage.</p> <p>Modalités possibles : "Autorisé";"Autorisé (Provisoire)";"Retrait"</p>
ephy_amm_usage.csv	dose_homologuee	Double	Dose homologuée pour l'usage.
ephy_amm_usage.csv	dose_homologuee_unite	Texte	Unité de la dose homologuée pour l'usage.
ephy_amm_usage.csv	dar_jours	Entier	Délai de traitement avant récolte (exprimé en jours) à respecter pour l'usage.
ephy_amm_usage.csv	dar_bbch	Entier	Délai de traitement avant récolte (exprimé en échelle BBCH) à respecter pour l'usage.
ephy_amm_usage.csv	nb_max_applications	Entier	Nombre maximum d'applications autorisé pour l'usage.
ephy_amm_usage.csv	usage_date_fin_distribution	Date	Date de fin de distribution du produit pour l'usage (format JJ/MM/AAAA).
ephy_amm_usage.csv	usage_date_fin_utilisation	Date	Date de fin d'utilisation du produit pour l'usage (format JJ/MM/AAAA).
ephy_amm_usage.csv	conditions_emploi	Texte	Les limitations de la portée sont indiquées dans les conditions de l'usage. Par exemple, pour un produit autorisé pour l'usage « Cassissier*Trt Part.Aer.*Chenilles phytophages » avec comme conditions d'usage « Uniquement sur cassissier et groseilliers », le produit

Green Data for Health
Data papers Challenge Green Data for Health

Fichier	Champ	Type	Description
			n'est pas autorisé sur les autres cultures incluses dans la portée de l'usage (le myrtille, le sureau noir, la mûre (Morus sp.), l'airelle, le cynhorodon, l'azerolier).
ephy_amm_usage.csv	znt_aquatique	Double	Largeur de la zone de non-traitement (exprimée en m) autour des milieux aquatiques à respecter pour l'usage.
ephy_amm_usage.csv	znt_arthropodes_non_cibles	Double	Largeur de la zone de non-traitement (exprimée en m) pour protéger les arthropodes non ciblés par le produit à respecter pour l'usage.
ephy_amm_usage.csv	znt_plantes_non_cibles	Double	Largeur de la zone de non-traitement (exprimée en m) pour protéger les plantes non ciblées par le produit à respecter pour l'usage.
ephy_amm_usage.csv	intervalle_min	Entier	Délai minimum (exprimé en jours) à respecter entre deux applications du produit pour l'usage considéré.
ephy_ref_culture_rpg.csv	id_culture	Texte	Clé du type #_culture (# étant un nombre).
ephy_ref_culture_rpg.csv	culture	Texte	Listes des groupes de cultures (nomenclature E-phy) sur lesquelles le produit est autorisé (liste des cultures issues de l'usage).
ephy_ref_culture_rpg.csv	code_gpe_cult_rpg	Texte	Listes des codes des groupes de cultures (nomenclature RPG en 28 classes en vigueur pour le millésime 2021) sur lesquelles le produit est autorisé. NB : La correspondance entre les cultures E-phy (issues de l'usage) et groupes de cultures du RPG a été réalisée par l'OFB.
ephy_ref_culture_rpg.csv	libelle_gpe_cult_rpg	Texte	Listes des groupes de cultures (nomenclature RPG en 28 classes en vigueur pour le millésime 2021) sur lesquelles le produit est autorisé. NB : La correspondance entre les cultures E-phy (issues de l'usage) et groupes de cultures du RPG a été réalisée par l'OFB.
ephy_usage_culture.csv	id_usage	Texte	Clé du type #_usage (# étant un nombre).
ephy_usage_culture.csv	id_culture	Texte	Clé du type #_culture (# étant un nombre).
ephy_amm_mention_danger.csv	amm	Texte	N°AMM (E-phy) du produit.
ephy_amm_mention_danger.csv	classes_mention_danger_lb_court	Texte	Liste des classes de mention de danger associées au produit (libellé court).
ephy_amm_mention_danger.csv	classes_mention_danger_lb_long	Texte	Liste des classes de mention de danger associées au produit (libellé long).
ephy_amm_phrase_risque.csv	amm	Texte	N°AMM (E-phy) du produit.
ephy_amm_phrase_risque.csv	phrases_risque_lb_court	Texte	Liste des phrases de risque associées au produit (libellé court).
ephy_amm_phrase_risque.csv	phrases_risque_lb_long	Texte	Liste des phrases de risque associées au produit (libellé long).
agritox_substance.csv	id_substance	Texte	Clé du type #_sub_agritox (# étant un nombre).
agritox_substance.csv	substance_cas	Texte	Concaténation des champs "Substance" et "N°CAS" d'AGRITOX séparés par un <i>underscore</i> ("_").
agritox_substance.csv	substance	Texte	Libellé de la substance d'après AGRITOX (libellé sans accents).
agritox_substance.csv	cas	Texte	N° CAS de la substance d'après AGRITOX.
agritox_substance.csv	nom_uipac	Texte	Libellé de la substance d'après UIPAC.
agritox_substance.csv	ec	Texte	N° EC de la substance d'après AGRITOX.

Green Data for Health
Data papers Challenge Green Data for Health

Fichier	Champ	Type	Description
agritox_substance.csv	masse_molaire	Texte	Masse molaire de la substance.
agritox_substance.csv	formule_brute	Texte	Formule chimique brute de la substance.
agritox_substance.csv	purete_minimale	Texte	Pureté minimale de la substance d'après AGRITOX.
agritox_substance.csv	impuretes_pertinentes	Texte	Impureté(s) pertinente(s) susceptible(s) d'être associée(s) à la substance d'après AGRITOX.
agritox_substance_bnvd.csv	id_substance_bnvd	Texte	Clé du type #_sub_bnvd (# étant un nombre).
agritox_substance_bnvd.csv	id_substance_agritox	Texte	Clé du type #_sub_agritox (# étant un nombre).
agritox_substance_classement.csv	id_substance	Texte	Clé du type #_sub_agritox (# étant un nombre).
agritox_substance_classement.csv	classement_ref	Texte	Référence à partir de laquelle ont été établis les classements associés à chaque substance répertoriée dans Agritox.
agritox_substance_classement.csv	classement_date	Date	Date (format JJ/MM/AAAA) à laquelle ont été établis les classements associés à chaque substance répertoriée dans Agritox.
agritox_substance_classement.csv	classement_code_h	Texte	Codes H associés à chaque substance répertoriée dans Agritox.
agritox_substance_classement.csv	classement_categorie_danger	Texte	Libellés clairs des catégories de danger associées à chaque substance répertoriée dans Agritox. Attention : certains codes H peuvent correspondre à plusieurs catégories de danger. Pour chaque substance associée à un code H de ce type, Agritox indique le détail de la catégorie (ex. pour le code H300 correspondant aux catégories "Toxicité aiguë (par voie orale), catégories 1, 2" la substance sera soit en catégorie 1, soit en catégorie 2).
agritox_substance_classement.csv	classement_mention_danger	Texte	Libellés clairs des mentions de danger (aussi dites "phrases de risque") associées à chaque substance répertoriée dans Agritox.
agritox_substance_classement.csv	facteur_m_valeur	Entier	Valeur du facteur M éventuel associé au classement de la substance. Modalités possibles : 1, 10, 100, 1000, 10000, 100000, 1000000 Les facteurs M sont utilisés pour la classification des mélanges contenant des substances classées comme très toxiques pour l'environnement (catégorie 1). Le concept de facteur M a été établi pour donner plus de poids à des substances très toxiques lors de la classification des mélanges. Ainsi, un mélange contenant une substance ayant un facteur M élevé sera plus facilement classé dangereux pour l'environnement (source : https://www.sicadae.eu/ufags/test/).
agritox_substance_classement.csv	facteur_m_origine	Texte	Origine du facteur M éventuel associé au classement de la substance. Modalités possibles : UE, ANSES
agritox_substance_classement.csv	facteur_m_date	Date	Valeur du facteur M éventuel associé au classement de la substance.
agritox_substance_classement.csv	classification_simpl	Texte	Liste des classifications simplifiées associées à la substance dans Agritox.
agritox_substance_proprietes_pc.csv	id_substance	Texte	Clé du type #_sub_agritox (# étant un nombre).

Green Data for Health
Data papers Challenge Green Data for Health

Fichier	Champ	Type	Description
agritox_substance_proprietes_pc.csv	propriete	Texte	Propriété(s) physico-chimique(s) mesurée(s) sur la substance. Modalités possibles : Constante de Henry ; Log P ; pka ; Pression de vapeur ; Solubilité dans l'eau
agritox_substance_proprietes_pc.csv	temperature	Texte	Température à laquelle a été déterminée la valeur de la propriété physico-chimique.
agritox_substance_proprietes_pc.csv	ph	Texte	pH auquel a été déterminée la valeur de la propriété physico-chimique.
agritox_substance_proprietes_pc.csv	observation	Texte	Observation(s) éventuelle(s) quant à la valeur de la propriété physico-chimique (ex: isomère auquel se rapporte la valeur).
agritox_substance_proprietes_pc.csv	valeur	Texte	Valeur associée à la propriété chimique, en fonction des conditions de température et de pH dans lesquelles celle-ci a été déterminée.
agritox_substance_proprietes_pc.csv	unite	Texte	Unité de mesure de la propriété chimique. Modalités possibles : Constante de Henry : Pa.m ³ /mol ; Log P : (vide) ; pka : (vide) ; Pression de vapeur : µPa, mmHg, mPa, nPa, Pa ; Solubilité dans l'eau : µg/L, g/L, mg/L, ug/L, (vide)
agritox_substance_toxicite.csv	id_substance	Texte	Clé du type #_sub_agritox (# étant un nombre).
agritox_substance_toxicite.csv	nature_etude	Texte	Nature de l'étude de toxicité réalisée sur la substance. Modalités possibles : "Toxicité aiguë cutanée" ; "Toxicité aiguë inhalation" ; "Toxicité aiguë VO"
agritox_substance_toxicite.csv	dl50_cl50	Texte	Variable mesurée dans le cadre de l'étude de toxicité. Modalités possibles (associées à chaque nature d'étude) : "CL50" (Nature de l'étude="Toxicité aiguë inhalation"); "DL50" (Nature de l'étude="Toxicité aiguë cutanée" ou "Toxicité aiguë VO")
agritox_substance_toxicite.csv	espece	Texte	Espèce sur laquelle a été mesurée la variable (DL50 ou CL50) dans le cadre de l'étude.
agritox_substance_toxicite.csv	sexe	Texte	Sexe(s) de l'espèce sur le(s)quel(s) a été mesurée la variable (DL50 ou CL50) dans le cadre de l'étude.
agritox_substance_toxicite.csv	valeur	Texte	Valeur de la variable mesurée dans le cadre de l'étude visée, en fonction de l'espèce et du sexe sur lesquels cette variable a été mesurée.
agritox_substance_toxicite.csv	unite	Texte	Unité de mesure associée à la variable de toxicité aiguë mesurée. Modalités possibles : DL50 : « mg/kg » ; CL50 : « mg/L »
agritox_substance_ecotoxicite.csv	id_substance	Texte	Clé du type #_sub_agritox (# étant un nombre)
agritox_substance_ecotoxicite.csv	etudes	Texte	Nature de la (ou des) étude(s) d'écotoxicité réalisée(s) sur la substance pour déterminer la PNEC (Predicted no-effect concentration).
agritox_substance_ecotoxicite.csv	donnee_toxicite	Texte	Variable mesurée dans le cadre de l'étude d'écotoxicité réalisée sur la substance pour déterminer la PNEC.

Green Data for Health
Data papers Challenge Green Data for Health

Fichier	Champ	Type	Description
			<p>Modalités possibles : "NOAEC";"NOEC";"CE50";"CL50";"HC5";"EAC";"NOEAEC";"CE10"</p> <ul style="list-style-type: none"> • La CL50 est la Concentration Létale pour 50% des organismes exposés, les poissons ainsi que quelques invertébrés aquatiques. • La CE50 est la Concentration d'Effet pour 50% des organismes exposés, les daphnies et quelques autres invertébrés aquatiques, les algues et les plantes aquatiques. Pour les invertébrés, l'effet est l'immobilisation. Pour les algues et plantes, sont mesurés les effets sur la biomasse (b), le taux de croissance (r) ou la densité de fronde (df) (exemple : CEr50). • La CSEO est la Concentration Sans Effet Observé pour la totalité des organismes exposés (NOEC en anglais). • En cas d'évaluation probabiliste, une HC5 (Hazard Concentration) est déterminée : c'est la concentration limite pour laquelle 5% des organismes sont affectés (exemple de critère : 5% des organismes ont une CL50 inférieure à la HC5). Un facteur de sécurité (FS) approprié prenant en compte les recommandations du document guide aquatique est appliqué à cette valeur de HC5. • Pour un microcosme ou un mésocosme, plusieurs valeurs de toxicité finale peuvent être déterminées, en particulier : <ul style="list-style-type: none"> o CSEO ou NOEC o NOEAEC (concentration sans effet écologique néfaste observé). • L'EAC (Ecologically Acceptable Concentration) est définie en considérant l'ensemble des études sur les organismes aquatiques. Elle est comparable à la PNEC.
agritox_substance_ecotoxicite.csv	donnee_toxicite_valeur	Texte	Valeur de la variable mesurée dans le cadre de l'étude d'écotoxicité réalisée sur la substance pour déterminer la PNEC (Predicted no-effect concentration). Pour une même combinaison de substances/étude(s)/donnée de toxicité plusieurs valeurs de donnée de toxicité peuvent exister.
agritox_substance_ecotoxicite.csv	donnee_toxicite_unite	Texte	Unité de mesure associée à la variable d'écotoxicité mesurée. Modalités possibles : "mg/kg sed";"ug/L";"mg/L"
agritox_substance_ecotoxicite.csv	facteur_securite	Texte	<p>Facteur utilisé par les évaluateurs du risque pour calculer une dose de référence considérée comme sûre ou en dessous de laquelle un effet nocif est improbable. La valeur du facteur de sécurité dépend de l'effet toxique, de la taille et du type de population à protéger, ainsi que de la qualité des données d'exposition et des données (éco)toxicologiques disponibles.</p> <p>- Lorsque la PNEC est définie sur la base d'études standards, elle est calculée à partir des données de toxicité aiguë (CL50 ou CE50 pour les poissons et invertébrés aquatiques, respectivement), à court terme (CE50 pour les algues et les plantes) ou à long terme (CSEO pour les poissons et invertébrés aquatiques) sur différents groupes taxonomiques</p>

Green Data for Health
Data papers Challenge Green Data for Health

Fichier	Champ	Type	Description
			et d'un facteur de sécurité (FS) de 100 (étude de toxicité aiguë) et de 10 (étude de toxicité à court terme et à long terme). Le facteur de sécurité permet de tenir compte de la variabilité intra et inter espèces. - Si la PNEC est basée sur des études de niveau supérieur (microcosmes, mésocosmes), la valeur du facteur de sécurité approprié est déterminée selon la qualité de l'information fournie et la nature des effets (2 à 4).
agritox_substance_ecotoxicite.csv	valeur_pnec	Texte	Valeur de la PNEC (Predicted no-effect concentration) associée à l'étude d'écotoxicité réalisée sur la substance. - C'est la valeur toxique de référence utilisée pour évaluer les risques pour les organismes aquatiques. - La PNEC désigne la concentration pour, ou en-dessous, de laquelle il n'est pas attendu d'effet inacceptable sur l'ensemble des organismes aquatiques. - La PNEC est déterminée en prenant en compte l'ensemble des données de toxicité sur les organismes aquatiques issues du dossier européen de la substance active et de ses métabolites, et des dossiers d'Autorisation de Mise sur le Marché (AMM) ; elle couvre les effets potentiels de la substance active et de ses métabolites pertinents pour l'évaluation du risque pour les organismes aquatiques. Elle est exprimée en µg par litre, ou en µg par kg de sédiment. Pour une même combinaison de substances/étude(s) plusieurs valeurs de PNEC peuvent exister
agritox_substance_ecotoxicite.csv	unite_pnec	Texte	Unité de mesure associée à la valeur de PNEC pour l'étude d'écotoxicité réalisée sur la substance. Modalités possibles : "ng/L";"ug/L";"mg/L";"ug/kg sed"
agritox_substance_vtr.csv	id_substance	Texte	Clé du type #_sub_agritox (# étant un nombre).
agritox_substance_vtr.csv	nom	Texte	Nom de la valeur toxicologique de référence (VTR) mesurée. Les VTR sont calculées à partir (1) d'une dose sans effet néfaste ou NOAEL (No Observed Adverse Effect Level) et (2) d'un facteur de sécurité (FS). La NOAEL choisie pour le calcul est issue de l'étude la plus appropriée sur une espèce animale sensible et représentative. Le facteur de sécurité tient compte de la variabilité intra et inter-espèce et de la nature des effets de la substance. Modalités possibles : • "DJA" (Dose journalière acceptable) : quantité de substance qui peut être quotidiennement ingérée par le consommateur, pendant toute la vie, sans effet néfaste pour sa santé. • "ARFD" (Dose de référence aiguë) : quantité maximum de substance active qui peut être ingérée par le consommateur pendant une courte période (c'est à dire au cours d'un repas ou d'un jour, dans la nourriture ou l'eau de boisson), sans effet dangereux

Green Data for Health
Data papers Challenge Green Data for Health

Fichier	Champ	Type	Description
			pour sa santé. <ul style="list-style-type: none"> • "AOEL" (Niveau d'exposition acceptable pour l'opérateur, les personnes présentes, le travailleur et les résidents) : valeur de référence par rapport à laquelle les expositions non alimentaires aux produits phytopharmaceutiques sont actuellement évaluées. • "AOEC" (Concentration acceptable pour l'opérateur, les personnes présentes, le travailleur et les résidents) • "AAOEL" (Niveau d'exposition aigu acceptable pour l'opérateur, les personnes présentes, le travailleur et les résidents) : valeur de référence par rapport à laquelle les expositions aiguës non alimentaires (c'est-à-dire celles qui pourraient être encourues en une seule journée) sont évaluées. Cela ne s'applique aux produits phytopharmaceutiques pour lesquels de telles expositions pourraient produire une toxicité significative.
agritox_substance_vtr.csv	valeur	Texte	Valeur de la VTR déterminée d'après les étude(s) pivot(s) décrite(s) dans les champs Etude(s) pivot(s) 1 et 2. Modalités possibles : nombre décimal avec nombre variables de chiffres significatifs, "non pertinent" ou "non fixée par manque de données"
agritox_substance_vtr.csv	unite	Texte	Unité associée à la valeur de VTR. Modalités possibles : "mg/kg pc"; "mg/kg/j"; "mg/m3"; "ppm"; (vide)
agritox_substance_vtr.csv	source	Texte	Source de la VTR.
agritox_substance_vtr.csv	date	Date	Date de détermination de la VTR au format JJ/MM/AAAA.
agritox_substance_vtr.csv	etude_pivot_1	Texte	Libellé de la première étude pivot ayant permis la détermination de cette valeur de VTR.
agritox_substance_vtr.csv	etude_pivot_1_noael	Texte	Valeur de la dose sans effet néfaste ou NOAEL (No Observed Adverse Effect Level) associée à l'étude pivot 1.
agritox_substance_vtr.csv	etude_pivot_1_noael_unite	Texte	Unité associée à la valeur de la dose sans effet néfaste ou NOAEL (No Observed Adverse Effect Level) associée à l'étude pivot 1.
agritox_substance_vtr.csv	etude_pivot_2	Texte	Libellé de la seconde étude pivot éventuelle ayant permis la détermination de cette valeur de VTR.
agritox_substance_vtr.csv	facteur_securite	Texte	Facteur de sécurité appliqué à la valeur de NOAEL de(s) (l')étude(s) pivot(s) pour obtenir la VTR. Le facteur de sécurité tient compte de la variabilité intra et inter-espèce et de la nature des effets de la substance.
agritox_substance_vtr.csv	facteur_correction_abs_orale_incompl	Texte	Facteur de correction appliqué le cas échéant pour tenir compte d'une absorption orale incomplète. Exprimé sous forme de texte en "xx%". Remarque : ce facteur ne peut s'appliquer qu'aux VTR de type AOEL ou AAOEL.
atp_substance.csv	id_substance	Texte	Clé du type #_sub_atp (# étant un nombre).
atp_substance.csv	index	Texte(12)	Référence à partir de laquelle ont été établis les classements associés à chaque substance répertoriée dans ATP.
atp_substance.csv	ici	Texte	Nom ATP de la substance.

Green Data for Health
Data papers Challenge Green Data for Health

Fichier	Champ	Type	Description
atp_substance.csv	ec	Texte	N° EC de la substance d'après ATP.
atp_substance.csv	cas	Texte	N° CAS de la substance d'après ATP.
atp_substance_bnvd.csv	id_substance_bnvd	Texte	Clé du type #_sub_bnvd (# étant un nombre).
atp_substance_bnvd.csv	id_substance_atp	Texte	Clé du type #_sub_atp (# étant un nombre).
atp_substance_classement.csv	id_substance	Texte	Clé du type #_sub_atp (# étant un nombre).
atp_substance_classement.csv	code_h_detail	Texte	Codes H associés à chaque substance répertoriée dans ATP, avec détails : présence possible à la suite du code H de caractères types "*" (ex. "H201 ****", "H361f ***", "H372 *") et/ou de la spécification entre parenthèses des organes et/ou voies d'exposition (ex. "H373 (gastro-intestinal tract) (oral)").
atp_substance_classement.csv	code_h	Texte	Codes H associés à chaque substance répertoriée dans ATP, sans détails (nomenclature type "H####" sauf pour les codes "H350i"; "H360D"; "H360Df"; "H360F"; "H360FD"; "H360Fd"; "H361d"; "H361f"; "H361fd" qui peuvent comporter une ou deux lettres à la suite des 3 chiffres).
atp_substance_classement.csv	categorie_danger	Texte	Libellés clairs en français des catégories de danger associées à chaque substance répertoriée dans ATP. Attention : certains codes H peuvent correspondre à plusieurs catégories de danger. Pour chaque substance associée à un code H de ce type, ATP indique le détail de la catégorie (ex. pour le code H300 correspondant aux catégories "Toxicité aiguë (par voie orale), catégories 1, 2" la substance sera soit en catégorie 1, soit en catégorie 2).
atp_substance_classement.csv	mention_danger	Texte	Libellés clairs des mentions de danger (aussi dites "phrases de risque") associées à chaque substance répertoriée dans ATP.
atp_substance_classement.csv	classification_simplifiee	Texte	Liste des classifications simplifiées associées à la substance dans ATP.
eupdb_substance.csv	id_substance	Texte	Clé (identifiant numérique) de la substance d'après la EU pesticide database.
eupdb_substance.csv	substance_eupbd	Texte	Libellé de la substance (en anglais) d'après la EU pesticide database.
eupdb_substance.csv	etat_reg_1107_2019	Texte	Statut de la substance d'après le Reg. (EC) No 1107/2009. Modalités possibles : - "Not approved" ("Non approuvée") - "Approved" ("Approuvée") - "Pending" ("En attente d'évaluation") - "Not yet assessed at EU level" ("Substance pas encore évaluée au niveau européen")
eupdb_substance.csv	date_approbation	Date	Date d'approbation de la substance (format JJ/MM/AAAA).
eupdb_substance.csv	date_expiration_approbation	Date	Date d'expiration de l'approbation de la substance (format JJ/MM/AAAA).
eupdb_substance.csv	substance_base	Texte(3)	Indication si la substance est une substance de base (substances pouvant être utilisées à des fins phytopharmaceutiques en agriculture alors que cela n'est pas leur vocation première), en vertu du règlement (CE) no 1107/2009 (annexe C).

Green Data for Health
Data papers Challenge Green Data for Health

Fichier	Champ	Type	Description
eupdb_substance.csv	candidat_substitution	Texte	Indication si la substance est une substance dont on envisage la substitution, en vertu du règlement (CE) no 1107/2009 (annexe E). Modalités possibles : - "endocrine disrupting properties" ("propriétés de perturbation endocrinienne") - "endocrine disrupting properties, low ADI / ARfD / AOEL, two PBT criteria" ("propriétés de perturbation endocrinienne + faible DJA / ARfD / AOEL + deux critères PBT (persistant, bioaccumulable et toxique)") - "endocrine disrupting properties, toxic for reproduction 1A / 1B, two PBT criteria" ("propriétés de perturbation endocrinienne + reprotoxique 1A / 1B + deux critères PBT (persistant, bioaccumulable et toxique)") - "endocrine disrupting properties, two PBT criteria" ("propriétés de perturbation endocrinienne + deux critères PBT (persistant, bioaccumulable et toxique)") - "low ADI / ARfD / AOEL" ("faible DJA / ARfD / AOEL") - "low ADI / ARfD / AOEL, toxic for reproduction 1A / 1B" ("faible DJA / ARfD / AOEL + reprotoxique 1A / 1B") - "low ADI / ARfD / AOEL, two PBT criteria" ("faible DJA / ARfD / AOEL + deux critères PBT (persistant, bioaccumulable et toxique)") - "non-active isomers" ("isomères non-actifs") - "toxic for reproduction 1A / 1B" ("reprotoxique 1A / 1B") - "two PBT criteria" ("deux critères PBT (persistant, bioaccumulable et toxique)")
eupdb_substance.csv	faible_risque	Texte(3)	Indication si la substance est une substance à faible risque, en vertu du règlement (CE) no 1107/2009 (annexe D).
eupdb_substance.csv	autorisations_pays	Texte	Etats-membres dans lesquels la substance est autorisée.
eupdb_substance.csv	autorisation_france	Texte(3)	Indique si la substance est autorisée en France.
eupdb_substance_bnvd.csv	id_substance_bnvd	Texte	Clé du type #_sub_bnvd (# étant un nombre).
eupdb_substance_bnvd.csv	id_substance_eupdb	Texte	Clé (identifiant numérique) de la substance d'après la EU pesticide database.
sandre_parametres.csv	cdparametre	Texte	Clé (code paramètre Sandre associé au N° CAS de la substance) de la substance d'après le référentiel Sandre – paramètres.
sandre_parametres.csv	nomparametre	Texte	Libellé associé au code paramètre Sandre.
sandre_parametres.csv	cas	Texte	Code CAS (de la substance chimique) associé au code paramètre Sandre.
sandre_substance_bnvd.csv	id_substance_bnvd	Texte	Clé du type #_sub_bnvd (# étant un nombre).
sandre_substance_bnvd.csv	id_substance_sandre	Texte	Clé (code paramètre Sandre associé au N° CAS de la substance) de la substance d'après le référentiel Sandre – paramètres.
sandre_gpes_parametres.csv	cdgroupeparametres	Texte	Liste des codes Sandre des groupes de paramètres associés au code Paramètre Sandre de la substance.
sandre_gpes_parametres.csv	nomgroupeparametres	Texte	Liste des libellés Sandre des groupes de paramètres associés aux codes Groupes de Paramètres (Sandre).

Green Data for Health
Data papers Challenge Green Data for Health

Fichier	Champ	Type	Description
sandre_gpes_parametres.csv	famille	Texte	Famille chimique associée au code Paramètre Sandre de la substance (une seule famille chimique possible par code Paramètre). La famille chimique se retrouve via les groupes de paramètres ayant comme groupes de paramètres pères : "Chimique" (code=41) ou "Micropolluants organiques" (code=50).
sandre_gpes_parametres.csv	fonction	Texte	Fonction associée au code Paramètre Sandre de la substance (plusieurs fonctions possibles par code Paramètre). Les fonctions se retrouvent via les groupes de paramètres ayant comme groupes de paramètres pères : "Phyosanitaires" (code=95).
sandre_parametres_gpes_parametres.csv	id_substance_bnvd	Texte	Clé (code paramètre Sandre associé au N° CAS de la substance) de la substance d'après le référentiel Sandre – paramètres.
sandre_parametres_gpes_parametres.csv	cdgroupeparametres	Texte	Liste des codes Sandre des groupes de paramètres associés au code Paramètre Sandre de la substance.
c3po_source.csv	source	Texte	Libellé de la source de données.
c3po_source.csv	date_publication	Date	Date de publication de la source (si connue).
c3po_source.csv	date_collecte	Date	Date de collecte des données de la source pour intégration dans C3PO.
c3po_version.csv	version	Texte	N° de version de la base de données C3PO.
c3po_version.csv	date_publication	Date	Date de publication de la version de la base de données C3PO.

Intérêt des données

Intérêt scientifique des données

Cette nouvelle base de données C3PO permettra de faciliter les valorisations et l'exploitation des données de la BNV-D par rapport à ce qu'il est possible de faire aujourd'hui. Par exemple, C3PO permettra de répondre immédiatement à la question « sur quelle culture telle substance de la BNV-D est-elle autorisée ? ».

En couplant C3PO avec les données de la BNV-D spatialisée au code postal acheteur, les valorisations potentielles sont multiples. Par exemple, il sera désormais possible de cartographier les achats de substances actives en fonction de la toxicité de ces dernières. Ou encore, il sera possible d'étudier l'effet des périodes d'interdiction au niveau européen de certaines substances sur les achats de ces dernières.

Autres sources d'intérêt

Outre la réappropriation par des études scientifiques, la base de données C3PO permettra de faciliter l'appui aux politiques publiques.

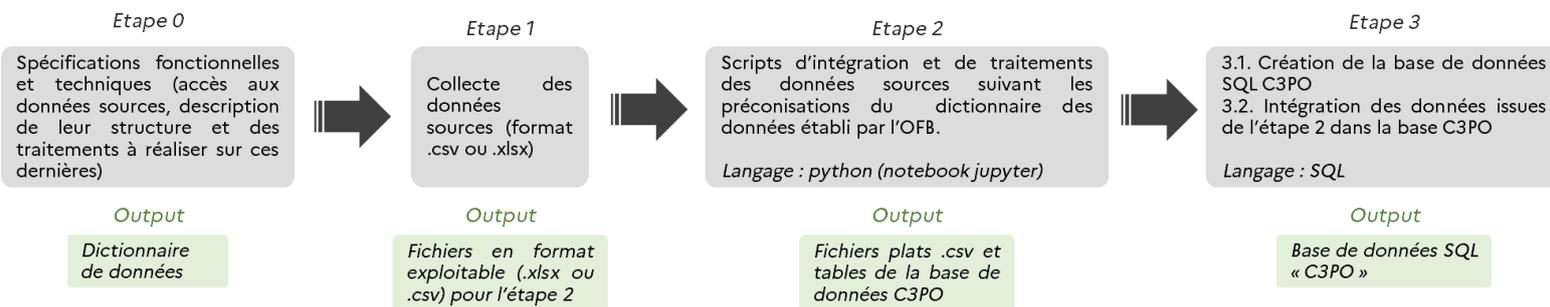
Par exemple, elle permettra d'automatiser la préparation de l'arrêté annuel établissant la liste des substances définies à l'article L. 213-10-8 du code de l'environnement relatif à la redevance pour pollutions diffuses, travail nécessitant de croiser les données de la BNV-D, d'Agritox, de la EU PDB et de l'ATP.

Un autre exemple d'appui aux politiques publiques est le croisement entre données de ventes et de toxicité des substances actives pesticides opéré dans le cadre de l'étude amont à la révision des arrêtés surveillance et évaluations, avec le calcul du Potentiel toxique (PoTox), à savoir la quantité vendue divisée par la valeur de la PNEC. C3PO permettrait d'automatiser le calcul de cet indicateur.

Collecte et traitement des données

Collecte et processus de traitement mis en œuvre

Le processus de traitement mis en œuvre dans le cadre du défi a été le suivant :



Pour l'étape 1, les données sources ont été collectées aux liens suivants :

Source	Lien	Format	Commentaire
BNV-D : achats de produits France entière (depuis 2013)	https://www.data.gouv.fr/fr/datasets/achats-de-pesticides-par-code-postal/	.csv empaquetés dans .zip	Nomenclature des fichiers "BNVD_TRACABILITE_{date de production}_ACHAT_FR_PRODUIIT_{année de vente}.csv"
BNV-D : achats de substances France entière (depuis 2013)	https://www.data.gouv.fr/fr/datasets/achats-de-pesticides-par-code-postal/	.csv empaquetés dans .zip	Nomenclature des fichiers "BNVD_TRACABILITE_{date de production}_ACHAT_FR_SUBSTANCE_{année de vente}.csv"
BNV-D : ventes de produits France entière (depuis 2008)	https://www.data.gouv.fr/fr/datasets/ventes-de-pesticides-par-departement/	.csv empaquetés dans .zip	Nomenclature des fichiers "BNVD_TRACABILITE_{date de production}_VENTE_FR_PRODUIIT_{année de vente}.csv"
BNV-D : ventes de substances France entière (depuis 2008)	https://www.data.gouv.fr/fr/datasets/ventes-de-pesticides-par-departement/	.csv empaquetés dans .zip	Nomenclature des fichiers "BNVD_TRACABILITE_{date de production}_VENTE_FR_SUBSTANCE_{année de vente}.csv"
API Sandre Référentiel - paramètres validés avec N°CAS	https://api.sandre.eaufrance.fr/referentiels/v1/par.json?outputSchema=SANDREv4&filter=\$StParametre=%27Valid%C3%A9%27%20and%20\$ParametreChimique.CdCASSubstanceChimique!=%27%27	.json (.csv et .xml possibles également)	
API Sandre Référentiel - groupes de paramètres validés	https://api.sandre.eaufrance.fr/referentiels/v1/gpr.json?outputSchema=SANDREv4&filter=\$StGroupeParametres=%27Valid%C3%A9%27%20	.json (.csv et .xml possibles également)	
20230516_referentiel_fonction_substances_V6.1.zip	https://backend-ventes-phytopharmaceutiques.eaufrance.fr/tracabilite-api/a-propos/piece-jointe/20230516_referentiel_fonction_substances_V6.1.zip	.csv empaqueté dans .zip	Prendre en compte le fichier "20230516_referentiel_fonction_substances_V6.1.1_detail.csv"
EU Pesticide database (Active substances, safeners and synergists)	https://ec.europa.eu/food/plant/pesticides/eu-pesticides-database/start/screen/active-substances	.xlsx	Cliquer sur "Export Active substances"

Green Data for Health
Data papers Challenge Green Data for Health

Source	Lien	Format	Commentaire
Agritox	https://www.data.gouv.fr/fr/datasets/base-de-donnees-agritox/	.csv empaquetés dans .zip	
ATP	https://echa.europa.eu/fr/information-on-chemicals/annex-vi-to-clp	.xlsx	Prendre en compte uniquement la dernière version de l'annexe VI disponible
E-phy	https://www.data.gouv.fr/fr/datasets/donnees-ouvertes-du-catalogue-e-phy-des-produits-phytopharmaceutiques-matieres-fertilisantes-et-supports-de-culture-adjuvants-produits-mixtes-et-melanges/	.csv empaquetés dans .zip	Prendre en compte les fichiers au format UTF-8 (nomenclature : "decisionamm-intrant-format-csv-{date de production}-utf8.zip")
RPG - référentiel des cultures et groupes de cultures	https://geoservices.ign.fr/documentation/donnees/vecteur/rpg	.csv	Cliquer sur le lien hyper-texte "Table référentielle des cultures et des groupes de cultures"

Conditions d'accès et de réutilisation

L'accès aux données de la base C3PO est ouvert. Les fichiers plats (.csv) décrits dans ce document sont disponibles sur [data.gouv.fr](https://www.data.gouv.fr) (licence ouverte Etalab 2.0). Le code source est également ouvert à tous et accessible sur [Gitlab](https://www.gitlab.com) (licence Cecill-B v1).

Plus-value et perspectives

Le challenge GD4H a permis de mettre en évidence les apports de C3PO pour faciliter les études (et leur répliquabilité) croisant diverses informations sur les pesticides. Un cas d'usage (croisant données de ventes et de toxicité/écotoxicité, et données de ventes et d'autorisation des substances) a notamment été réalisé, il est accessible au lien suivant : https://qonfluens.shinyapps.io/shiny_basec3po/.

Cette version du data paper porte sur les tables (et fichiers plats correspondant) du schéma **substance** de la base de données. Dans la prochaine version, les jeux de données relatifs aux produits dans lesquels sont contenues ces substances (schéma **amm** de la base de données) seront également rendus accessibles.

D'autres sources de données ouvertes pourront également alimenter à terme la base C3PO.

Bibliographie

[1] Inserm. 2021. **Pesticides et effets sur la santé : Nouvelles données.** Collection Expertise collective. Montrouge : EDP Sciences, 1009 pages.

[2] Sánchez-Bayo F., Wyckhuys K. A.G.. 2019. **Worldwide decline of the entomofauna: A review of its drivers.** *Biological Conservation*, Volume 232 : Pages 8-27, 0006-3207. <https://doi.org/10.1016/j.biocon.2019.01.020>

[3] Rigal S., Dakosa V., Alonso H., Auniņš A., Benkő Z., Brotons L., Chodkiewicz T., Chylarecki P., de Carlil E., del Moral J. C., Domşa C., Escandell V., Fontaine B., Foppen R., Gregory R., Harriss S., Herrando S., Husby M., Ieronymidou Ch., Jiguet F., Kennedy J., Klvaňová A., Kmecl P., Kuczyński L., Kurlavičius P., Kålås J. A., Lehikoinen A., Lindström Å., Lorrillier R., Moshøj Ch., Nellis R., Noble D., Palm Eskildsen D., Paquet J.-Y., Pelissié M., Pladevall C., Portolou D., Reif J., Schmid H., Seaman B., Szabo Z. D., Szép T., Tellini Florenzano G., Teufelbauer N., Trautmann S., van Turnhout Ch., Vermouzek Z., Vikstrøm T., Voříšek P., Weiserbs A. and Devictor V. 2023. **Farmland practices are driving bird population decline across Europe.** *PNAS*. <https://www.pnas.org/doi/10.1073/pnas.2216573120>

[4] Laure Mamy (coord.), Stéphane Pesce (coord.), Wilfried Sanchez (coord.), Marcel Amichot, Joan Artigas, Stéphanie Aviron, Carole Barthélémy, Rémy Beaudouin, Carole Bedos, Annette Bérard, Philippe Berny, Cédric Bertrand, Colette Bertrand, Stéphane Betoulle, Eve Bureau-Point, Sandrine Charles, Arnaud Chaumot, Bruno Chauvel, Michael Coeurdassier, Marie-France Corio-Costet, Marie-Agnès Coutellec, Olivier Crouzet, Isabelle Doussan, Jean-Paul Douzals, Juliette Faburé, Clémentine Fritsch, Nicola Gallai, Patrice Gonzalez, Véronique Gouy, Mickael Hedde, Alexandra Langlais, Fabrice Le Bellec, Christophe Leboulanger, Christelle Margoum, Fabrice Martin-Laurent, Rémi Mongruel, Soizic Morin, Christian Mouglin, Dominique Munaron, Sylvie Nélieu, Céline Pelosi, Magali Rault, Nicolas Ris, Sergi Sabater, Sabine Stachowski-Haberkorn, Elliott Sucré, Marielle Thomas, Julien Tournebize, Anne-Laure Achard, Morgane Le Gall, Sophie Le Perchec, Estelle Delebarre, Floriane Larras, Sophie Leenhardt (coord.). 2022. **Impacts des produits phytopharmaceutiques sur la biodiversité et les services écosystémiques.** Rapport d'ESCO, INRAE - Ifremer (France), 1408 pages.

[5] European Environment Agency. 2022. **Pesticides in rivers, lakes and groundwater in Europe.** URL : <https://www.eea.europa.eu/ims/pesticides-in-rivers-lakes-and>

[6] Claire Froger, Claudy Jolivet, Hélène Budzinski, Manon Pierdet, Giovanni Caria, Nicolas P. A. Saby, Dominique Arrouays, Antonio Bispo. **Pesticide Residues in French Soils: Occurrence, Risks, and Persistence.** *Environmental Science & Technology*, 57 (20), 7818-7827. [DOI: 10.1021/acs.est.2c09591](https://doi.org/10.1021/acs.est.2c09591)

[7] Générations Futures. 2023. **Comparaison de la toxicité et de l'écotoxicité des pesticides utilisés en agriculture biologique et en agriculture conventionnelle.** URL : <https://www.generations-futures.fr/wp-content/uploads/2023/05/bio-non-bio-vf.pdf>