

Atelier open **FOOD** facts

Séminaire ALIMining
29–30 septembre 2025 Toulouse



SOMMAIRE



Programme Science

Exploration des données



LE PROGRAMME “OPEN FOOD FACTS POUR LA SCIENCE”



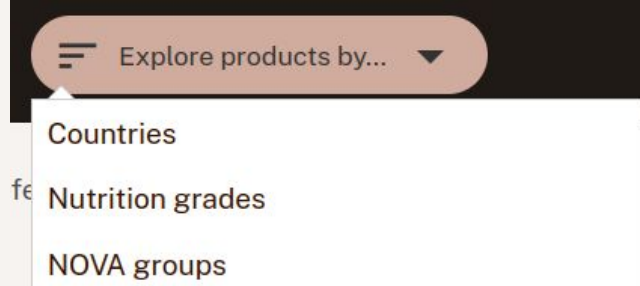
Données et outils accessibles “sur étagère” (1/2)



Une base de données mondiale ouverte

- ±4 M produits
- disponibles dans 180 pays & 40 langues

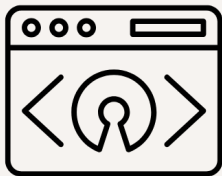
+ Assistance*



- Explorez les produits selon une 30aine de **facettes** : pays, catégories, marques, étiquettes, ingrédients, ...
- De nombreux [exports](#) : MongoDB, JSONL, CSV, Parquet
- [Mirabelle](#) : un outil pour les requêtes complexes et pour construire vos propres sélections de données, grâce au langage SQL
- Les produits avec des erreurs de qualité de données peuvent être filtrés ([218+ contrôles disponibles](#))
- Indicateur de complétude des données
- [Quick graphs](#)
- Partenariat avec [HuggingFace](#) (export Parquet quotidien)

*Dans la limite de **quelques heures** par projet. Ex : nous pouvons vous aider à répondre à des questions complexes ou à créer des outils pour vos recherches.

Données et outils accessibles “sur étagère” (2/2)



Open Source Open Contents

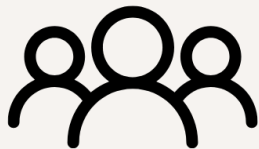
- TOUTES nos données, logiciels, contenus, etc., sont entièrement ouverts.

+ **Assistance***

- Ce qui est codé pour un projet est publié en open source et **immédiatement disponible** à toutes fins et projets
- **Docs**: 3000 pages de wiki, documentation technique, tutoriels, extraits de code disponibles ; ex : guide de [réutilisation de nos données](#)
- **API** : si vous avez besoin de données en temps réel ou si vous avez besoin de l'historique d'un produit ([doc](#))
- **SDKs** : composants logiciels prêts à l'emploi pour vos applications pour accéder aux données Open Food Facts
- Apprentissage automatique (IA) pour compléter les données des contributeurs (modèles disponibles). Ex : prédiction de catégorie

*Dans la limite de quelques heures par projet. Ex : nous pouvons vous aider à répondre à des questions complexes ou à créer des outils pour vos recherches.

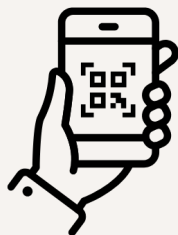
Comptez sur la communauté Open Food Facts



- 30,000+ contributeurs dans 180+ pays
- Contribuer à faire avancer la science est une grande source de motivation

- Rejoignez notre Slack **#science**
- Échangez avec ses pairs
- Rejoignez nos webinaires réguliers des chercheurs
- Publiez les données de vos propres produits sur Open Food Facts pour **bénéficier de l'effort communautaire** (complétion, corrections...)
- Organisez **des opérations spéciales** pour collecter des données en amont de votre recherche de catégories de produits particulières ou de pays d'intérêt
 - Ex : [L'opération "Plein pot sur les emballages"](#)

Comment contribuer au programme OFF pour la Science ?



Enrichir la base de données

- Ajouter des produits
- Contribuer aux taxonomies
- Convaincre les étudiants d'enrichir la base de données
- Organiser une Scan Party



Donnez-nous votre avis !

- Racontez nous comment vous utilisez OFF dans un article de blog/un court webinaire
- Complétez le formulaire Répertoire des études liées à OFF
- Faites nous vos retours pour que l'on améliore l'expérience des chercheurs



D'un tableur privé...

[illegible]

**Aidez-nous à mettre en valeur
notre impact collectif en ajoutant
votre projet !**

...Vers un annuaire public

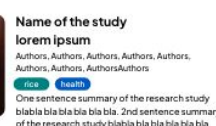


Search by keyword

Add my research project +

➔ Topic, Country, publication year,
Intensity of Open Food Facts usage

Name of the study
lorem ipsum
Authors, Authors, Authors, Authors, Authors,
Authors, Authors, AuthorsAuthors
[ham](#) [health](#)
One sentence summary of the research study
blabla bla bla bla bla bla. 2nd sentence summary
of the research study blabla bla bla bla bla bla



Opportunités de recherche de fonds ensemble



Projets de recherche en cours

- Horizon Europe : [DivinFood](#), [Foodture](#)
- [NutriSight](#) (DRG4Food)
- [FoodContact](#) (INSERM, EREN)
- U-TRANS (Institut Gustave Roussy)
- [Nutrinet-Santé](#) (EREN)

Actuellement impliqués dans 6 projets de recherche

Sujets d'intérêt :

Santé

Nutrition

Impact environnemental

Open data

Science citoyenne

Taxonomies

Machine learning

Éducation

Suivi de l'offre alimentaire

Politiques publiques

Nous sommes ouverts à collaborer !



Nos perspectives (1/3)



Open Food Facts pour la Science : un ensemble de processus et d'outils pour faciliter la réutilisation de nos données pour la recherche scientifique :

Corriger Collecter

- En amont, **des processus de collecte de données** formalisés à des usages scientifiques (protocoles de science citoyenne, ...)
- **Fonctionnalités personnalisées dans l'application** pour leurs besoins de recherche : par exemple, la possibilité de collecter de données non encore disponibles dans Open Food Facts (comme les produits ciblant les enfants, etc.)
- **Nettoyage** des données incorrectes ou erronées
- Identification des données présentant des valeurs **aberrantes** ou exceptionnelles
- Plus de **scores et de données calculés** : meilleure analyse des ingrédients (plus de langues, meilleure précision) ; symbole nutritionnel sur le devant de l'emballage (Canada) ; boîte d'information nutritionnelle (USDA) ; octogones d'Amérique latine ; Health Star Rating (Pacifique Ouest) ; Indice Nutritionnel de Santé (HENI)

Enrichir

- **Interopérabilité** avec plus de bases de données (aujourd'hui : Ciqua, Agrybalise, Eurocode2, Système harmonisé de l'Organisation mondiale des douanes... ; demain : Foundation Foods (USDA), Foodex2 (EFSA), Oqali...)
- **Pré-calculs statistiques automatiques** : pour chaque point de données, statistiques pré-calculées pour une meilleure compréhension des données — taille de l'échantillon, valeurs mini/max, moyenne, médiane, valeur la plus fréquente, etc.

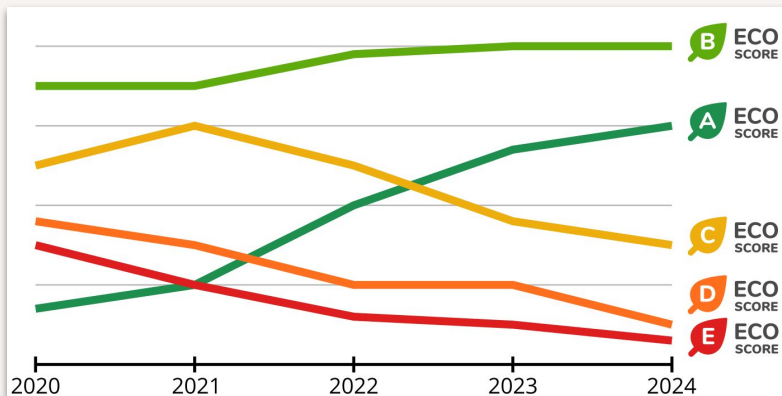
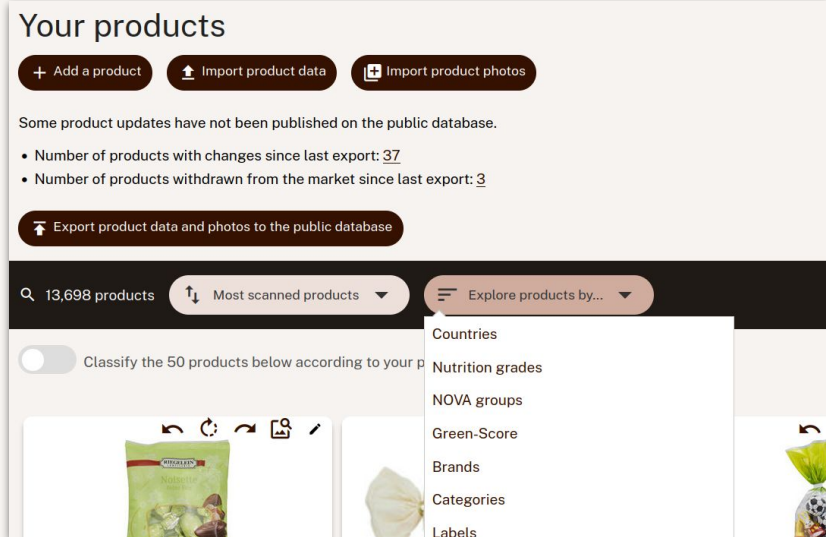
Extraire

- Extraction de données **multicritères** simplifiée



Nos perspectives (2/3)

- **Un portail dédié pour les scientifiques** : une plateforme permettant de profiter des outils d'Open Food Facts : documentation, stockage, contrôle et nettoyage, données calculées, sélection, téléchargements, etc.
- Une nouvelle plateforme pour **suivre l'évolution des produits**
 - identifier les marqueurs de l'évolution du produit (nutrition, ingrédients, scores...)
 - permettre le suivi de l'évolution : facettes, graphiques...



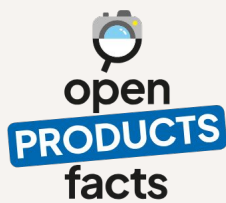
Nos perspectives (3/3)

 Open Prices



Open Prices : lancé en 2024, le projet montre déjà des résultats prometteurs

- 100,000 points de données dans le monde entier



Autres bases de données de produits :

- Open **Beauty** Facts : 54,000 produits
- Open **Products** Facts : 27,000 produits
- Open **Pet Food** Facts : 14,000 produits

Pas encore aussi réussi que OFF, mais nous y travaillons activement.



EXPLORATION DES DONNÉES





<https://github.com/openfoodfacts/openfoodfacts-exports/blob/main/tutorials/hands-on-data/open%20food%20facts%20data%20tutorial.ipynb>

Or go to

<https://github.com/openfoodfacts/openfoodfacts-exports/>

and navigate to
tutorials/hands-on-data

and open

open food facts data tutorial.ipynb



Questions – Réponses



**Merci pour votre
attention**

Contacts

- contact@openfoodfacts.org



**Rejoignez notre
répertoire**