

Les bases de données alimentaires, des outils à la fiabilité contestable

Dans le sillage de Yuka, de nombreuses bases de données alimentaires ont vu le jour, permettant aux consommateurs de mieux se renseigner sur ce qu'ils achètent. Du moins le pensent-ils car les erreurs inhérentes au modèle de collecte des informations plombent la fiabilité de ces outils.

Temps de lecture : minute

28 août 2020

Des noix de Saint-Jacques classées comme produit contenant des fruits à coque ? Sur les applications pour repérer la valeur nutritionnelle ou les ingrédients indésirables des produits alimentaires transformés, ce type d'erreur est plus courant qu'on ne croit, et cela vient souvent des producteurs eux-mêmes. Créée en 2012 et alimentée par les consommateurs, la base de données française sur les produits de grande consommation Open Food Facts a permis, grâce à sa gratuité et au nombre conséquent d'articles référencés, à de nombreuses applications de voir le jour : par exemple Scanup et Yuka pour des conseils sur les valeurs nutritionnelles ou les additifs, des coachs nutritionnels numériques, ou encore Isla Food pour suivre un régime halal.



À lire aussi

Peut-on (vraiment) faire confiance aux applis pour « mieux consommer » ?

Au début réticents voire opposés à la collecte de leurs données par des tiers, les industriels ont depuis compris l'intérêt de les partager et commencé à transmettre leurs informations à Open Food Facts. Mais ils ont parfois été agacés de voir des anonymes écraser leurs données et reprochent à la base d'être peu mise à jour alors qu'ils changent régulièrement leurs recettes. Conscient de l'enjeu, Open Food Facts a créé en octobre 2019 une passerelle spéciale pour que les producteurs puissent rentrer leurs informations régulièrement. Elles sont vérifiées et ensuite bloquées pour qu'aucun autre contributeur ne puisse les changer.

5 à 40% d'erreurs

Soucieuse de garder le contrôle face aux bases de données collaboratives, la filière agroalimentaire développe de nouvelles bases de données avec des informations (ingrédients, valeurs nutritionnelles, allergènes, etc.) prises à la source, et donc en théorie plus fiables. Ainsi

l'association nationale des industries alimentaires (Ania) a lancé le projet Num'alim, qui recueillera toutes les données sur les aliments (de l'agriculteur au vendeur) et leur consommation afin d'alimenter des études sur les comportements et de conseiller les professionnels. Cette plateforme s'appuiera sur Codeonline Food, la base de données de l'organisme de standardisation GS1 France, chargé par l'Ania de compiler les informations-produits directement auprès de ses membres (industriels mais aussi petits producteurs). Selon Gaëlle Auffret, directrice du marché des Produits grande consommation (PGC) pour GS1 France, cette collecte d'informations à la source assure *"une base de données fiables et sûres, qui ne peuvent pas être modifiées et qui donne la possibilité aux marques de redevenir maîtres de leurs données"* .

Mais les informations fournies par les producteurs sont-elles forcément plus fiables que celles renseignées par les consommateurs ? Julien Doyen, fondateur de l'application Allergobox qui détecte les allergies, en doute : sa plateforme Consotrust, qu'il a créée pour vérifier, corriger et certifier les données grâce à des algorithmes, a constaté 5% à 40% d'erreurs selon les producteurs. Les erreurs qu'il a constatées sont de tous types : des fautes d'orthographe (qui empêchent les algorithmes des applications de détecter un additif ou un allergène), des incohérences sur les valeurs nutritionnelles (par exemple 120 grammes de sucre dans une portion de 100 grammes de nourriture), des inversions de données (pourcentage de lipides confondu avec le pourcentage d'acide gras saturé), des erreurs d'analyse notamment dans l'identification des allergènes (noix de Saint-Jacques assimilée à un fruit à coque par exemple).

Ces inexactitudes peuvent avoir des conséquences pour les personnes qui surveillent leur consommation de gluten ou de sel, ou souffrent d'allergies ou de diabète. *"On ne pourrait pas créer Allergobox en se basant sur les données de Codeonline ou d'Open Food Facts telles quelles, sinon on tuerait quelqu'un le premier jour"* , explique Julien Doyen à l'AFP. *"Je ne conseille pas forcément l'utilisation de ces applications à mes patients,*

confirme à l'AFP le diététicien du sport David Guyonnet. *C'est au cas par cas : pour le diabète ça peut être utile, pour le reste je préfère leur apprendre à lire les étiquettes* , témoigne-t-il. D'autant que selon lui, l'affichage de notes basées sur le taux de sucre ou de gras peuvent *"induire en erreur"* les utilisateurs et focaliser leur attention sur un élément non pertinent.

"Pas de gestion centralisée ni d'outil"

L'économiste de l'alimentation Pierre Combris a participé à la création en 2008 de l'observatoire de la qualité de l'alimentation et s'est donc frotté très tôt à l'anarchie des données alimentaires. *"Le problème essentiel, c'est que la gestion des données est d'une telle complexité que beaucoup d'exploitants ont du mal à contrôler la qualité de leurs données en interne, ils n'ont pas forcément de gestion centralisée ni d'outils à leur disposition"* , démontre t-il à l'AFP.

Un manque d'outils qu'Alkemics tente de combler : les producteurs rentrent leurs informations sur les produits sur la plateforme de cette startup française, afin qu'elles soient partagées avec les distributeurs, mais aussi s'ils le souhaitent avec les applications destinées aux consommateurs. La directrice marketing d'Alkemics, Florence Di Nicola, y voit *"une chaîne automatisée des données produits"* venant répondre au *"challenge énorme"* de la réglementation dite Inco, entrée en vigueur en 2014 et censée rendre les étiquettes plus lisibles en les uniformisant et obligeant tous les sites de commerce en ligne à fournir les mêmes informations sur les produits alimentaires qu'en magasin.

Codeonline Food, Open Food Facts et Alkemics utilisent des algorithmes de contrôle des données pour éviter notamment les incohérences. L'application Yuka, qui a créé en 2018 sa propre base de données, contrôle assidûment les données d'utilisateurs et fait davantage confiance aux professionnels, même si elle leur applique des contrôles. Alkemics

estime que la *"quasi totalité des produits atteignent 98% de qualité"*, c'est-à-dire que 98% des champs ne contiennent pas d'erreur.

"Aujourd'hui il y a certainement des problèmes de fiabilité, mais ça dépend des secteurs. Plus le produit est brut, plus la fiabilité est élevée" , concède Jérôme François, le directeur général de Num'Alim qui réfléchit à créer un indicateur de fiabilité. Car dans l'attente d'une évaluation concrète, il est *"impossible de savoir le taux de fiabilité"* des différentes bases de données.

Maddyness avec AFP

Article écrit par Maddyness avec AFP