

Google : ne pas confondre épidémie et questions

Google a créé un service gratuit, Google Flu Trends (GFT), www.google.org/flutrends, basé sur l'idée que certains termes de recherche sont de bons indicateurs de la propagation de la grippe ; en comptant le nombre des requêtes utilisant ces termes, GFT pourrait détecter les épidémies de grippe et suivre leur évolution.

Considéré initialement comme un complément de la surveillance épidémiologique classique, GFT a été rapidement présenté comme une alternative (gratuite) à ces réseaux.

Après plusieurs années d'expérience, l'apport de GFT est plus modeste que prévu : incapacité à prédire l'arrivée de la pandémie de 2009 puis à la suivre, surestimation importante de l'ampleur épidémique en décembre 2012, etc. De plus, les résultats fournis par GFT ne peuvent être considérés comme scientifiques, car les chercheurs sans lien avec Google ne peuvent pas vérifier comment ces résultats sont obtenus : il est impossible de connaître l'algorithme de calcul pour cause de secret commercial ; de plus, cet algorithme change fréquemment pour s'adapter au mieux à l'objectif prioritaire de Google, le ciblage publicitaire. Enfin, le nombre des requêtes est biaisé par des sociétés qui envoient des requêtes en très grand nombre pour se trouver en tête de liste lors de l'affichage des résultats des requêtes des internautes.

Finalement, quand Google est consulté par les internautes qui se posent des questions sur la grippe, GFT permet surtout de suivre les variations de l'intérêt de la population pour cette maladie.

Source : *The Parable of Google Flu : Traps in Big Data Analysis.* Lazer D et al, *Science* vol 343, 14 mars 2014.

Consultable sur www.sciencemag.org/content/343/6176.toc

Le Dico du doc



Big Data



Surnom donné aux bases de données collectées systématiquement à grande échelle pour gérer ou évaluer des grands organismes ou des grands problèmes : actualisation des moteurs de recherche sur internet, gestion des hôpitaux, liquidation des prestations d'assurance-maladie, enregistrement des décès...

Il est toujours tentant d'utiliser ces bases de données pour mener des analyses non prévues au départ :

- détection d'épidémies (de maladie, de mode ou de rumeur), à partir des requêtes faites par les internautes,
- estimation de la surmortalité due aux épidémies, à partir de la hausse du nombre des hospitalisations en période épidémique,
- évaluation de l'efficacité d'un vaccin, à partir des baisses de la mortalité succédant à une campagne de vaccination...

La pertinence et la validité de ce type d'analyse posent de nombreux problèmes scientifiques, légaux et humains.

Source : Open Rome

Météo antibio

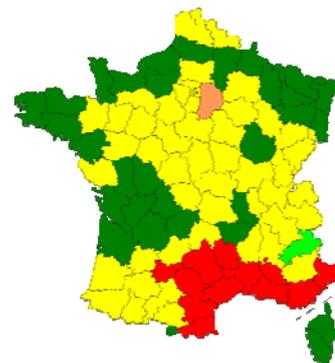
Risques

- Grippe faible
- Bronchiolite très faible
- Inf respiratoire en baisse
- Gastro-entérite faible
- Allergies pollens élevé

Sources : <http://www.grog.org>
et <http://www.pollens.fr>

Pollens

Risque très élevé dans la région méditerranéenne et le rebord sud-ouest du Massif Central. La pollution aérienne aggrave les troubles respiratoires.



Rouge très élevé Orange élevé Jaune moyen
Vert foncé faible Vert clair très faible

Source : RNSA <http://www.pollens.fr>