



Libération

mardi 13 avril 2021 856 words, p. 22,23

IDÉES/

Quand l'Insee se nourrit de données privées

Par ANNE-LAURE DELATTE Anne-Laure Delatte, chargée de recherches au CNRS (Leda, université Dauphine)

ÉCONOMIQUES Avec la crise sanitaire, l'institut utilise de plus en plus les big data. Comment ouvrir la statistique publique aux données collectées dans le cadre d'une activité commerciale, sans pour autant être piégé dans un rapport de dépendance ?

Il y a une phrase du directeur général de l'Insee dans la dernière note de conjoncture de mars 2021, qui n'a l'air de rien mais est une bombe truffée de sous-entendus: «Dans l'attente d'une éventuelle évolution du cadre juridique, l'Insee n'est pas fermé à contribuer financièrement à l'exploitation statistique réalisée en partenariat, sauf lorsqu'il s'agit de mettre à notre disposition des données déjà existantes au bon format ; la plupart de nos partenaires ont compris cette position.» Traduction : que les boîtes privées qui ont essayé de me faire payer pendant la crise aillent bien se faire cuire un oeuf au Kazakhstan, je ne vous passerai pas un kopeck d'argent public pour des données brutes que vous collectez déjà. Pour comprendre la guerre de tranchées qui se joue encore sur le terrain des données entre l'administration française et les grandes entreprises privées, il faut revenir au samedi soir du 14 mars 2020, veille du premier et seul vrai confinement. Tout va fermer, les entreprises, les commerces et les Français vont rester cloîtrés chez eux. Le directeur de l'Insee comprend qu'il aura besoin de données instantanées de comptabilité nationale pour diagnostiquer ce qui se passe, ce qui va se passer et pour que Bercy décide des politiques publiques de réaction à la crise. Connaître en temps réel l'ampleur des dégâts sur les différents secteurs d'activité, sur l'emploi, avoir un découpage par région pour réagir au plus près pendant le grand confinement. Or, la note de conjoncture de l'Insee qui fournit une prévision de l'activité pour les six mois à venir et qui est «the» référence pour Bercy, se base sur des enquêtes faites avant le confinement. Non seulement cette note prévue pour le 26 mars peut être jetée à la poubelle mais l'outil d'enquête traditionnelle ne pourra pas être mobilisé pour avoir les informations pertinentes après le 15 mars : les enquêteurs ne pourront plus se déplacer physiquement dans les foyers confinés ou les entreprises fermées, et surtout l'Insee va avoir besoin de données instantanées et non mensuelles pour saisir, vite, l'évolution des principaux agrégats économiques et aider à piloter la politique économique au plus près. On est samedi soir, et le directeur de l'Insee pense qu'il va avoir besoin de données à «haute fréquence» : données de cartes bancaires et données de caisse des grandes surfaces pour suivre la consommation des ménages au jour le jour ; fret ferroviaire et consommation d'électricité des entreprises pour approximer leur production branche par branche ; recherche de mots-clés dans Google comme «allocations-chômage» pour avoir une idée de l'état de l'emploi en temps réel. L'équipe conjoncture utilise déjà des big data pour établir ses prévisions mais il va falloir industrialiser la pratique tant que la crise durera. Et c'est ce qui s'est passé : l'Insee les a intégrées dans ses modèles pour produire de l'information en temps réel. Mais pas à n'importe quel prix. D'où la petite phrase du début.

On raconte qu'il y a eu des négociations âpres entre l'administration française et les propriétaires de données massives -les banques, les opérateurs téléphoniques, les plateformes de vente en ligne etc. -, que l'administration a tenu bon pour que les données massives soient transférées gracieusement ou à un coût limité. Car si ces données revêtent un intérêt général majeur en temps de crise, l'administration voulait limiter au maximum le fait de rémunérer des entreprises privées avec de l'argent public pour des données qui existent



déjà. Autrement dit: pas de profit privé grâce à la crise sur le dos des contribuables. Le principe paraît juste et il faut reconnaître un certain sangfroid et un vrai sens de l'intérêt général à ceux qui l'ont défendu dans la tempête. Sauf que la question va se reposer même après la pandémie car les épisodes de disruption massive risquent de se multiplier dans la décennie à venir (coucou la crise climatique). Il est donc urgent de réfléchir à un cadre réglementaire pour encadrer le transfert vers les instituts de statistique publique de données privées existantes, collectées dans le cadre d'une activité commerciale - comptes bancaires, courses en supermarché, téléphonie mobile. Cela fait partie des évolutions silencieuses mais clés pour piloter l'activité économique en temps de grand bouleversement. Ouvrir la statistique publique aux données massives privées sans pour autant être piégé dans un rapport de dépendance; adapter la collecte maison pour augmenter la fréquence des données et ne pas dépendre des données collectées par branches, syndicats patronaux etc. qui défendent seulement leurs intérêts. Car au coeur du pilotage économique, il y a les données ; et au coeur du contrat social, l'intérêt général. ?

Cette chronique est assurée en alternance par Ioana Marinescu, Anne-Laure Delatte, Bruno Amable et Pierre-Yves Geoffard

Pas de profit privé grâce à la crise sur le dos des contribuables. Le principe paraît juste et la question va se reposer après la pandémie car les épisodes de disruption massive risquent de se multiplier (coucou la crise climatique).

Google

Google is the world's No. 1 Internet research site. Researches are available in more than 100 languages and include translation, definition, currency exchange, and other functions. Sales break down by ...



Google

Google est le 1er site mondial de recherche sur Internet. Les recherches sont disponibles en plus de 100 langues et comprennent des fonctionnalités de traduction, de définition, de conversion de devises ...

