

Le 3 avril 2024

Délibéré sur les relations entre le service statistique public et le Centre d'accès sécurisé aux données (CASD) dans un contexte en mutation

L'Autorité de la statistique publique (ASP) a pris connaissance des conclusions de la mission conduite par l'Inspection générale de l'Insee concernant « le positionnement de l'Insee vis-à-vis du Centre d'accès sécurisé aux données (CASD) dans un contexte en évolution ».

Sans directement appartenir au service statistique public (SSP), le CASD, dont l'objectif était lors de sa création de faciliter la diffusion aux chercheurs des micro-données relevant de la statistique publique, fait partie des « autres autorités nationales¹ », reconnues au niveau européen. À ce titre, l'ASP a donc aussi procédé à l'audition, le 14 mars 2024, du directeur du CASD, accompagné de la directrice générale du Groupe des écoles nationales d'économie et de statistique (Genes), qui préside l'Assemblée générale de cet organisme, constitué sous forme de groupement d'intérêt public (GIP).

Ni le rapport de l'inspection générale, ni l'audition n'avaient pour objet de se pencher sur le fonctionnement interne du CASD, dont l'importance des réalisations est largement reconnue, y compris au plan international, mais d'éclairer les effets et interrelations entre sa stratégie de développement et le service statistique public. Ils ont permis de mettre en évidence les mutations très importantes intervenues ou en cours dans son rôle et dans son environnement, qui concernent d'une part la prise en compte des besoins du SSP en termes de capacités de stockage sécurisé et d'appariement de micro-données, et d'autre part les conditions d'ouverture des données aux chercheurs.

Concernant les besoins des diverses composantes du SSP, les demandes d'espace sécurisé provenant de certains services statistiques ministériels se sont développées de façon incrémentale, sans stratégie d'ensemble, au fur et à mesure que les possibilités d'appariement de diverses sources statistiques et administratives, notamment socio-fiscales, offraient de nouvelles opportunités aux statisticiens et que les solutions techniques et les performances du CASD leur offraient un environnement de travail adapté et facilitateur.

¹ *Other national Authorities*- ONA, ensemble qui comprend en France, outre le CASD, les services statistiques ministériels et le Centre d'épidémiologie sur les causes médicales de décès de l'Inserm (CépiDc-Inserm).

Or, la maîtrise générale de l'évolution des activités du CASD du point de vue du service statistique public dans son ensemble est essentielle pour apprécier la pertinence et les priorités de la réponse à apporter aux besoins formulés, au-delà des choix directement effectués par la direction du GIP.

Par ailleurs, le CASD accueille aujourd'hui des volumes massifs de micro-données d'origine administrative dont l'importance est essentielle et qui sont susceptibles d'être utilisées à la fois à des fins statistiques, de recherche, d'étude-évaluation ou de pilotage. Il s'agit non seulement des fichiers fiscaux, mais également des données issues du système national des données de santé (SNDS), de la déclaration sociale nominative (DSN) ou des fichiers d'étudiants de l'enseignement supérieur. L'élargissement de cette offre est un axe stratégique important, qui donne au CASD un rôle central dans la politique d'ouverture et d'utilisation des micro-données. Cet élargissement peut également faciliter des mutualisations et des coopérations, à la fois entre les membres du SSP et entre ces derniers et certaines équipes de recherche ou d'études.

L'ASP a dans ce cadre, et eu égard à la réussite du CASD dans l'accomplissement de ses missions, systématiquement encouragé les différentes composantes du SSP à favoriser par son entremise, après avis du Comité du secret de la statistique publique, l'accès des chercheurs ou des « évaluateurs » à ce type de données contrôlées et enrichies par leurs soins. Ce type d'accès devrait dans la mesure du possible être préféré à l'usage de données brutes à la qualité plus incertaine, et peut aussi servir de « pivot » à l'utilisation de ces données à des fins de pilotage.

La mission a néanmoins mis en évidence que plusieurs canaux de mise à disposition des micro-données, assortis de règles de confidentialité, mais aussi de tarifications différentes, existaient et étaient en cours de développement à l'intention du monde de la recherche (équipement d'excellence IDEE, *Hubs* sectoriels...). Ces canaux pluriels peuvent être un atout pour répondre aux besoins et aux modalités de financement différenciées de projets de recherche. Cependant, l'augmentation des tarifs du CASD intervenue ces dernières années, du fait de l'arrivée à échéance du financement de l'Agence nationale de la recherche (ANR) au titre des Equipex, de la stagnation de la contribution des membres du GIP et du développement limité des activités de valorisation technologique, peut poser question quant à l'équilibre et à la viabilité du « modèle économique » de ces différents modes de réponse aux besoins des chercheurs.

L'Autorité de la statistique publique estime donc nécessaire qu'une stratégie coordonnée soit établie par le service statistique public, sous l'égide de l'Insee, concernant les demandes prioritaires adressées au CASD pour ses propres besoins ou celui de certaines fonctions d'aide au pilotage exercées par les SSM (Depp, Dares concernant en particulier la DSN). Elle pense indispensable que ces priorités soient portées au sein des instances décisionnaires du CASD, dont le fonctionnement devrait permettre une visibilité plus claire et un dialogue plus approfondi concernant la programmation des activités et les équilibres financiers de l'organisme.

L'ASP souhaite par ailleurs que la relation du SSP avec le CASD s'inscrive dans le cadre d'une politique coordonnée de diffusion des données statistiques aux chercheurs. Celle-ci devrait

donner lieu, en lien avec les ministères chargés de la recherche et de son financement, à une réflexion renouvelée sur son contenu comme sur son « modèle économique », sachant que certains projets de recherche peinent à obtenir des financements intégrant les coûts d'accès sécurisé aux données et que ces voies d'accès se rapprochent d'infrastructures de recherche dont les chercheurs pourraient souhaiter une mise à disposition davantage mutualisée.