

# Le centre d'accès sécurisé aux données de la statistique publique française : un nouvel outil pour les chercheurs

 Jean-Pierre Le Gléau \* et Jean-François Royer \*\*

*En 2010, l'Insee a mis en place un nouveau dispositif d'accès aux fichiers individuels de la statistique publique, essentiellement à destination des chercheurs : le CASD, « centre d'accès sécurisé distant ». Rendu possible par l'évolution récente de la loi de 1951 sur le secret en matière de statistiques, il permet d'accéder à des données individuelles selon une procédure très réglementée. Cet article en donne une présentation générale.*

La protection de la confidentialité des données recueillies à des fins d'établissement de statistiques est un des principes majeurs de la statistique publique. Dès 1951, la loi organisant la statistique publique en France<sup>1</sup> comprenait le mot « secret » dans son titre. Dans le code de bonnes pratiques de la statistique européenne publié en 2005, le secret statistique figure au principe n° 5<sup>2</sup>. Il convient en effet de pouvoir assurer aux personnes répondant à une enquête que leurs réponses ne pourront en aucun cas être utilisées de façon à leur porter préjudice. C'est pourquoi la préservation du secret statistique constitue une disposition majeure de la loi fondamentale de la statistique française.

Pendant longtemps cette exigence n'a pas été ressentie comme une forte gêne par les utilisateurs de la statistique publique. Ceux-ci se contentaient de résultats globaux des enquêtes ou de résultats désagrégés selon des catégories prédéfinies (activités, professions, lieux) sans que la diffusion de ces résultats porte atteinte au secret. Cependant, la

pression pour que ces désagrégations soient de plus en plus fines est allée en croissant au fil des années, au fur et à mesure du progrès des techniques et des outils d'analyse statistique dans de nombreux secteurs de la société. Or, au-delà d'un certain niveau de détail dans les croisements, le secret statistique ne peut plus être respecté, car les « cases de tableaux » révèlent en fait des informations individuelles. Cette contradiction est devenue particulièrement apparente dans les travaux des chercheurs. La communauté des chercheurs en sciences humaines et sociales a depuis longtemps fait valoir son besoin d'accéder à des informations individuelles pour effectuer des travaux de plus en plus complexes. Ces travaux ne nécessitent pas, en général, de connaître nommément la personne ayant répondu à l'enquête. Mais les chercheurs ne peuvent plus se contenter de disposer de résultats agrégés. Le développement de la théorie et des outils économétriques, conjugué à celui de la puissance des instruments de calcul ouvre de multiples possibilités fondées sur l'analyse de données individuelles, que les chercheurs sont impatients d'explorer.

## Accès à des données individuelles : la situation entre 1951 et 2008

Dans le cas de renseignements individuels relatifs à des « données à caractère économique et financier » -

c'est-à-dire en pratique des données sur des entreprises ou des établissements - la loi de 1951 a dès l'origine prévu la possibilité de dérogations au secret statistique. Depuis 1984, de telles dérogations sont possibles sous réserve d'un passage devant un comité spécialisé rattaché au Conseil national de l'information statistique (Cnis), le comité du secret statistique. Celui-ci examine le bien-fondé de la demande, évalue la nécessité d'avoir recours à des informations individuelles, juge le sérieux du demandeur et la capacité de l'organisme auquel il appartient et qui soutient sa demande à respecter en pratique les contraintes liées au respect du secret statistique. En cas d'avis positif, le demandeur peut avoir accès aux données, après signature individuelle par chaque bénéficiaire d'un engagement à respecter les règles du secret statistique. Il faut souligner que cette possibilité, qui en pratique bénéficie à des équipes de recherche et à des administrations publiques, ne leur est pas réservée par la loi.

Dans le cas des « données relatives aux faits et comportements d'ordre privé » - c'est-à-dire en pratique des données sur des personnes ou des ménages - la loi de 1951 ne prévoyait à l'origine aucune dérogation et, jusqu'en 2008, aucune communication de renseignements individuels n'était possible. Seule la communication de fichiers anonymisés était licite.

\* Chef du département de la coordination statistique à l'Insee

\*\* Chercheur au Crest-Insee

1. Loi n° 51-711 du 7 juin 1951 sur l'obligation, la coordination et le secret en matière de statistiques

2. « Le respect de la vie privée ou du secret des affaires des fournisseurs de données (ménages, entreprises, administrations et autres répondants), la confidentialité des informations qu'ils communiquent et l'utilisation de celles-ci à des fins strictement statistiques doivent être absolument garantis ».

Pour répondre à la demande de plus en plus pressante des chercheurs, l'Insee a créé en 2006 les « fichiers de production et de recherche » (FPR), fichiers individuels, présentant une observation par individu enquêté, mais où l'on a au préalable supprimé toutes les variables permettant l'identification directe ou indirecte des personnes : nom, adresse, profession et nationalité détaillées, identification de l'employeur pour les salariés, etc. Ces FPR sont mis à la disposition des chercheurs via le réseau Quetelet, avec lequel l'Insee a signé une convention précisant les modalités de mise en œuvre du dispositif. Cette disposition a permis de satisfaire un grand nombre de demandes des chercheurs pour l'accès à des données individuelles, et elle reste d'une grande utilité.

Cependant, il existe un ensemble de demandes, légitimes du point de vue scientifique, qui ne peuvent trouver leur satisfaction dans la mise à disposition des FPR. Ce sont les projets de recherche qui nécessitent un détail géographique, par nationalité ou par profession non disponible dans les FPR<sup>3</sup>, ou qui font appel à un appariement avec d'autres fichiers, par exemple en couplant, grâce à l'identifiant de l'employeur, des caractéristiques de l'enquêté et de l'entreprise dans laquelle il travaille. L'accès à ces informations rompt nécessairement l'anonymat et était, de ce fait, impossible jusqu'en 2008.

### Les évolutions issues de la loi de 2008 concernant les données sur les ménages

En 2008, la modification de la loi statistique de 1951 par un article de la loi sur les archives<sup>4</sup> a ouvert la possibilité d'accéder à des renseignements individuels relatifs aux faits et comportements d'ordre privé sous deux conditions :

- l'une liée à la finalité : cette communication n'est possible qu'à des fins de statistique publique ou de recherche scientifique ou historique ;
- l'autre liée à la procédure : celle-ci exige en effet un avis du comité du

secret statistique, dont la compétence a donc été étendue.

La procédure suppose, après l'avis du comité du secret statistique, l'accord du service dont émanent les documents avant la décision formelle de communication par la direction des archives. Pour ce qui le concerne, l'Insee a informé le comité du secret statistique, lors de sa séance du 6 octobre 2009, des conditions qu'il mettrait pour donner cet accord :

- pour les services statistiques ministériels (SSM), qui ont dans leur mission de gérer des dossiers couverts par le secret statistique, les fichiers seront transmis sous réserve de la production par eux d'un document attestant qu'ils se sont bien dotés des moyens pour faire respecter le secret statistique et que leur équipement informatique présente bien l'étanchéité suffisante par rapport à celui du reste du ministère ;

- pour les autres demandeurs, l'accès aux données se fera par le biais d'un centre d'accès sécurisé ;

- dans ces deux cas, les fichiers auront été, si nécessaire, transformés pour perdre leur caractère directement nominatif ; c'est-à-dire qu'on y aura retiré, le cas échéant, le nom du répondant et le NIR (numéro d'inscription au répertoire - le numéro de la Sécurité sociale) ;

- pour les demandes d'accès à des documents directement nominatifs, l'accord ne pourrait être donné que dans le cas de la constitution d'un échantillon pour une enquête de la statistique publique ayant obtenu un visa<sup>5</sup> sous réserve de l'accord explicite, au cas par cas, du comité de direction de l'Insee.

3. Afin de rendre les FPR totalement anonymes, certaines variables géographiques, sur la nationalité ou la profession sont supprimées (code commune) ou regroupées en un petit nombre de modalités.

4. Loi n°2008-696 du 15 juillet 2008 relative aux archives, article 2

5. Ce visa est attribué par le ministre chargé de l'économie et le ministre à la compétence duquel ressortissent les intéressés.

6. Pour un panorama, on pourra consulter dans le *Courrier des statistiques* n°121-122 de décembre 2007 : « L'accès aux micro-données et la gestion de la confidentialité dans quelques INS européens » par Dominique Rouault.

### Le recours à un centre d'accès sécurisé

Depuis 2008, la loi rend donc possible la communication de données statistiques sur les personnes ou les ménages à des chercheurs. Mais l'impératif de protection des répondants continue à s'imposer à tous, et particulièrement aux institutions statistiques. C'est pourquoi l'Insee a choisi de recourir à un « centre d'accès sécurisé » pour mettre en œuvre les nouvelles dispositions permises par la loi.

Qu'est-ce qu'un centre d'accès sécurisé ? C'est un dispositif technique permettant à des utilisateurs autorisés de travailler sur des données confidentielles tout en s'assurant que ces données ne risquent pas d'être détournées pour d'autres fins, accidentellement ou intentionnellement. En matière de données statistiques, la création des premiers centres d'accès sécurisés destinés à la recherche remonte aux années 1990 aux USA et au Canada. En Europe, des pays du Nord (Danemark, Pays-Bas) ont créé de tels centres au début des années 2000. Ils ont été rejoints par la Grande-Bretagne, puis l'Allemagne, au milieu de cette décennie<sup>6</sup>.

Les premiers centres d'accès sécurisés à des données statistiques ont été des centres « physiques », c'est-à-dire des locaux gérés par une institution statistique (éventuellement à l'intérieur d'une université, comme au Canada) et contenant un serveur de données. Dans ces locaux, l'institution statistique donne au demandeur un accès physique directement sur le serveur. Cela implique que le demandeur se déplace pour travailler en dehors de son environnement habituel. Cette solution exige aussi l'aménagement de locaux spécifiques (éventuellement en plusieurs points du territoire) matériellement sécurisés (portes, fenêtres) dont la surveillance est confiée à une personne assermentée. C'est la solution retenue au Canada et en Allemagne. Elle est très coûteuse, aussi bien pour l'Institut de statistique que pour les chercheurs.

Grâce aux progrès de l'informatique, on a vu apparaître depuis 2005 des « centres d'accès sécurisés distants » (CASD) n'imposant plus aux chercheurs de se déplacer dans un endroit particulier. Ce sont des systèmes informatiques dans lesquels les données restent stockées sur un serveur unique, géré par l'institution statistique, mais sur lequel les chercheurs peuvent travailler à distance, sans risque de dissémination. Il peut s'agir seulement de pouvoir soumettre des travaux à distance (« remote submission »), les résultats étant récupérés de façon différée ; ou bien il peut s'agir d'un véritable accès « en ligne » (« remote access ») permettant au chercheur une forte interactivité avec les données. Dans ce dernier cas, le chercheur distant travaille réellement « sur le serveur » de l'institution statistique, mais le logiciel de sécurisation des données l'empêche de les exporter.

### La solution retenue par l'Insee

Pour les demandeurs autres que les SSM, l'Insee a donc choisi la solution du passage par un centre d'accès sécurisé distant en « remote access », où les données, les métadonnées et les programmes d'utilisation restent donc sur un serveur central unique géré par l'institution statistique ou sous son contrôle. Techniquement, cette solution aurait pu être mise en œuvre en installant un logiciel gérant la confidentialité (interdisant entre autres toute copie ou impression des données examinées) sur chaque poste de demandeur. C'est la solution retenue aux Pays-Bas. Elle implique une adaptation de ce logiciel à chaque configuration des ordinateurs ; elle doit être modifiée lorsque cette configuration change.

La solution retenue par l'Insee présente les mêmes fonctionnalités, mais a l'avantage de n'impliquer aucune intervention dans le système informatique des établissements de recherche. Elle a été mise au point et brevetée par l'équipe « Organisation du système d'information » du Genes (Groupe des écoles nationales

d'économie et statistique, établissement public qui était jusqu'à fin 2010 une direction de l'Insee). Elle consiste à mettre à disposition de l'organisme demandeur, un terminal appelé « SD Box » pour lui permettre de se connecter et de travailler sur le serveur installé au Genes. Cette « SD Box » permet au chercheur de voir les données sur lesquelles il travaille, de les intégrer dans ses calculs et d'en tirer des tableaux, graphiques, analyses. Mais à aucun moment il ne peut ni les imprimer, ni les enregistrer sur un support quelconque : clef USB, CD Rom... Ainsi, les données ne quittent jamais le serveur du Genes. La communication entre cette « SD Box » et le serveur est assurée par une liaison sécurisée et cryptée.

Cette solution a commencé à fonctionner début 2010, et à ce jour une trentaine de projets de recherche en bénéficient. Les deux paragraphes qui suivent décrivent plus en détail le cheminement d'un projet de recherche qui a recours à ce dispositif.

### Phase préparatoire : connaître les données, explorer les moyens d'accès, faire les formalités administratives

Concrètement, un chercheur qui souhaite accéder à des données individuelles sur les ménages produites par l'Insee ou par un SSM doit tout d'abord prendre connaissance de la source statistique dont ces données peuvent être tirées, et des différents produits de diffusion de cette source. Pour cela, ce chercheur dispose de deux ressources documentaires principales : le site Internet de l'Insee ([www.insee.fr](http://www.insee.fr)) (ou celui du SSM), riche en informations générales sur la source, en données agrégées et qui met à disposition un fichier individuel anonymisé d'usage public s'il a été possible d'en établir un ; le site Internet du réseau Quetelet ([www.reseau-quetelet.cnrs.fr](http://www.reseau-quetelet.cnrs.fr)) qui donne accès à un dessin de fichier, et au « FPR » s'il en existe un.

Beaucoup de recherches, on l'a dit, peuvent être conduites avec ces matériaux d'accès aisés et rapides. Le

chercheur qui estime qu'il ne peut pas s'en contenter et qu'il a besoin des fichiers complets doit à ce stade bien préciser pourquoi certaines variables lui manquent encore : c'est l'information cruciale qu'il devra fournir au comité du secret statistique. Pour cette phase d'élaboration, le chercheur devra impérativement prendre contact avec le service producteur, de façon à bien cerner sa demande et à vérifier qu'il a pris connaissance de tous les fichiers auxquels il peut avoir accès, de leurs possibilités et aussi de leurs limites.

Le chercheur présente ensuite sa demande au comité du secret statistique : le secrétariat du comité ([comite-secret@cnis.fr](mailto:comite-secret@cnis.fr)) peut lui fournir toutes les indications nécessaires. Le comité du secret statistique se réunit une fois par trimestre et donne un avis ; il opère comme cela a été décrit précédemment pour les données sur les entreprises. S'ajoutent cependant, dans les critères à prendre en compte, ceux de la finalité du traitement qui ne peut être, pour les données sur les ménages, que la statistique publique ou la recherche scientifique ou historique. La présence du chercheur à la séance du comité est nécessaire pour que ses membres puissent recevoir les éclaircissements et les assurances nécessaires.

En cas d'avis favorable du comité du secret statistique, le chercheur peut consulter les données sur le centre d'accès sécurisé distant. Un certain délai est cependant nécessaire pour d'une part mettre en place les données que le chercheur est autorisé à traiter, d'autre part accomplir les formalités requises par la Commission nationale de l'informatique et des libertés (Cnil).

Les données disponibles sur le CASD sont aujourd'hui celles auxquelles donne accès le comité du secret statistique, dans sa formation « ménages ». Il s'agit donc d'enquêtes statistiques auprès des ménages, récentes ou plus anciennes, y compris le recensement de la population. Mais il s'agit aussi de données

administratives déjà mises à la disposition de l'Insee ou des SSM au titre de l'article 7<sup>bis</sup> de la loi de 1951<sup>7</sup>. À titre d'exemple, les déclarations annuelles de données sociales (DADS) figurent dans cette catégorie. D'autres données peuvent également être introduites dans l'espace réservé au projet du chercheur sur le CASD, si leur utilisation est nécessaire dans le cadre de l'étude pour laquelle il a demandé l'accès aux informations confidentielles. Il peut s'agir par exemple de données apportées par le chercheur lui-même ; elles sont alors introduites sur le CASD par un agent habilité à la gestion du centre.

Comme on l'a vu plus haut, avant de se connecter, le chercheur doit également avoir accompli les formalités nécessaires auprès de la Cnil : en général une simple déclaration (pour les fichiers ne comportant pas de données sensibles), exceptionnellement, une demande d'autorisation. De son côté, l'Insee effectue progressivement toutes les formalités auprès de la Cnil pour déclarer cette nouvelle forme de diffusion de données à caractère personnel. Il s'agit de modifier toutes les déclarations (ou demandes d'autorisation) effectuées par l'Insee pour les sources statistiques depuis la création de la Cnil (1978). Ce processus est long et plusieurs mois seront encore nécessaires avant que toutes les sources concernées aient été traitées. Il se peut donc que l'accès à certains fichiers soit temporairement retardé, le temps que ces formalités se fassent. Pour diminuer au maximum l'impact sur les utilisateurs du CASD, l'Insee commence évidemment en priorité par les fichiers les plus demandés sur celui-ci.

## Phase active : la formation, et le travail sur le CASD

Dans les semaines qui suivent son autorisation à accéder à certaines données confidentielles, le chercheur est invité par les gestionnaires du CASD à participer à une séance dite « d'enrôlement », au cours de laquelle lui sont rappelés les principes généraux du secret statistique, la responsabilité personnelle qui lui sera confiée et les modalités pratiques d'accès au centre. Ces dernières consistent en :

- l'attribution d'une « SD Box », si l'organisme dont il dépend n'en est pas encore pourvue ;

- l'attribution d'une carte individuelle de connexion contenant un code obtenu par traitement de ses empreintes digitales, relevées au cours de cette séance.

Le chercheur peut avoir accès aux fichiers qu'il a demandés, sur la « SD Box » qui lui a été attribuée. La reconnaissance de son identité se fait grâce à sa carte et à la reconnaissance de ses empreintes digitales par le lecteur intégré à la SD Box. Il faut noter que ce dispositif ne comporte aucun stockage des empreintes digitales ailleurs que sur la carte qui reste en possession du chercheur.

Plusieurs chercheurs peuvent être inscrits sur un même projet : le dossier présenté au comité du secret statistique mentionne en ce cas les noms de tous les chercheurs associés à ce projet. Ceux-ci ont alors accès à un espace spécifique sur le CASD, dédié à leur projet. Le gestionnaire du centre y installe les bases de données qui leur sont spécifiques, et leur ouvre l'accès aux fichiers de l'Insee ou des SSM pour lesquels l'autorisation d'accès leur a été donnée. Personne d'autre qu'eux n'a accès à cet espace. Ils peuvent également utiliser la large gamme de logiciels de calcul statistique et d'édition mis à la disposition de tous les chercheurs sur le CASD.

Lorsque le chercheur a terminé un travail sur les données, il crée un fichier de résultats (tableaux, projet de publication,...) qu'il glisse dans

une « boîte aux lettres » virtuelle à laquelle il a accès. Ce fichier est examiné par des personnes habilitées au secret statistique. Elles s'assurent que le fichier ne contient pas d'éléments pouvant constituer une rupture du secret statistique. Si tel est le cas, elles avertissent le chercheur et lui demandent de rectifier son fichier de sortie, afin qu'il soit conforme aux règles de la confidentialité. S'il ne comporte aucun élément de nature à dévoiler une information confidentielle, il est renvoyé, par simple messagerie, au chercheur.

Aujourd'hui, cette vérification est exhaustive sur tous les fichiers de résultats produits dans le CASD dont un chercheur demande la sortie. C'est une solution coûteuse, car elle demande une grande expertise de la part des personnes habilitées. Son coût augmentera avec celui de la fréquentation du centre. Si le nombre de personnes habilitées qui se consacrent à ce contrôle était insuffisant, cela pourrait de plus entraîner des retards dans le feu vert donné aux chercheurs. C'est pourquoi on pourrait envisager à l'avenir de n'effectuer cette vérification que par sondage. Dans tous les cas, la responsabilité de la rupture du secret statistique reste de la responsabilité de celui qui aurait produit le fichier fautif. L'absence d'un contrôle exhaustif pourrait mettre en avant ce point et rappeler le chercheur à sa responsabilité personnelle, alors qu'un contrôle exhaustif peut lui donner l'illusion (fausse) qu'il est « couvert » par la vérification effectuée par un tiers.

Lorsque le chercheur a mené ses travaux jusqu'au bout, ce qui doit intervenir avant l'expiration de son habilitation, il en avertit les gestionnaires du CASD. Son accès au CASD est alors coupé. Cependant, son espace de travail est archivé pour une durée de trois ans renouvelable une fois. Cette disposition rend possibles les éventuels « retours vers les données », nécessaires après une publication par exemple. Au terme de cette période, tous les fichiers propres au chercheur et présents sur le CASD seront détruits.

7. « Sur demande du ministre chargé de l'économie, après avis du Conseil national de l'information statistique, et sauf disposition législative contraire, les informations (...) recueillies dans le cadre de sa mission, par (...) une personne morale de droit public (...) sont cédées, à des fins exclusives d'établissement de statistiques, à l'Insee ou aux services statistiques ministériels. »



## Vers une collaboration accrue entre chercheurs et statisticiens

L'accès à des données individuelles de sources statistiques pour mener des recherches est une nécessité pour la plupart des chercheurs contemporains. Un premier résultat des progrès récents pour faciliter cet accès - FPR, CASD - doit être une plus grande présence et visibilité des données françaises dans des recherches de niveau international, qu'elles soient menées par des chercheurs français ou par des chercheurs étrangers habilités. Il faut préciser en effet que ni la loi, ni la pratique administrative n'interdisent la communication de ces données à des chercheurs étrangers : ceux-ci peuvent parfaitement s'engager dans le processus de demande décrit plus haut. Et le système informatique permet d'installer des « SD Box » en tout lieu. Cependant, l'autorisation ne pourra être donnée que si les chercheurs concernés donnent toutes les garanties nécessaires, y compris sur le plan juridique, en ce qui concerne la sanction des ruptures éventuelles du secret statistique.

Un deuxième résultat de ces progrès sera une intensification de la collaboration entre statisticiens publics et chercheurs autour des sources statistiques. Déjà, des formations « sources » sont organisées à l'Insee pour les chercheurs bénéficiaires des autorisations : elles sont l'occasion d'échanges fructueux. Certaines difficultés rencontrées par les chercheurs en cours de travail devront être résolues en faisant appel aux producteurs, ce qui devrait enrichir la documentation...et dans certains cas permettre de corriger des erreurs. Au fur et à mesure que des recherches utilisant des données individuelles se termineront, on attend des chercheurs des « retours sur les sources » : appréciations, critiques, suggestions, etc. Comme beaucoup de ces sources sont renouvelées périodiquement, ce processus peut contribuer fortement à leur amélioration progressive.

## Le cas des appariements

Les chercheurs ont fréquemment besoin d'effectuer des appariements de données, pour enrichir les informations recueillies lors d'une enquête, par des données administratives, ou issues d'une autre enquête. Cet appariement se fait souvent sur un numéro d'identification (par exemple le NIR pour les personnes physiques et le numéro Sirène pour les personnes morales). Dans certains cas, le chercheur pourra mener lui-même cet appariement à l'intérieur du CASD. Dans d'autres cas, ce ne sera pas possible, en particulier si c'est le NIR qui est en jeu. Ce numéro, on l'a vu, est retiré, s'il existait dans le fichier initial, au moment de la mise sur le CASD. Si un appariement non faisable dans le CASD est cependant nécessaire, et est autorisé par la Cnil, le chercheur peut demander à l'Insee de réaliser cet appariement hors du CASD et de mettre le fichier issu de l'appariement (sans les numéros d'identification) sur le CASD. Il s'agit là bien évidemment d'une prestation spécifique, qui ne peut se faire qu'avec l'accord du gestionnaire du CASD et moyennant une facturation du service. Le fichier apparié restera sur le CASD, sans possibilité pour le chercheur de le sortir, ni d'en tirer des résultats qui ne respecteraient pas les règles du secret statistique. Ce travail peut aussi bien être effectué sur des fichiers disponibles sur le centre, qu'avec un ou des fichiers apportés par le chercheur, comme résultats d'une enquête qu'il a lui-même effectuée.

## L'avenir du CASD

Aujourd'hui, les capacités du CASD sont encore contraintes par les limites en matériel et en personnel qualifié nécessaire à son fonctionnement. Après une inévitable période de rodage, on pourra envisager d'élargir encore les missions qui lui sont attribuées, par exemple en se reposant sur lui également pour la mise à disposition de fichiers de données

individuelles sur les entreprises. Lorsqu'un chercheur a aujourd'hui accès à ces données, suite à un avis du comité du secret statistique, il obtient ces données sur un CD crypté qu'il doit conserver pendant un temps dans des conditions de sécurité garantissant la confidentialité des données, puis détruire. On conçoit facilement que le CASD apportera des garanties plus solides de ce point de vue.

En 2010, l'accès au CASD était gratuit. En effet certains aspects liés à la mise en route (disponibilité des fichiers, déclaration à la Cnil,...) ont parfois retardé la mise en œuvre du service pour les chercheurs, et il aurait pu paraître anormal de facturer un service qui n'avait pas encore atteint la qualité voulue. Il est clair cependant que, au-delà de l'investissement initial, le fonctionnement du CASD représente une charge pour l'institution qui en a la gestion : gestion des demandes, récupération des fichiers, parfois anciens, documentation, mise en ligne, formalités Cnil, attribution d'une « SD Box », licences pour les logiciels, vérification des sorties, etc. À l'avenir, l'accès au CASD devra donc couvrir les frais entraînés par ces activités. À ce jour, ni les modalités, ni le niveau de cette tarification n'ont encore été complètement définis.

Le CASD a été créé à l'Insee, au sein du Genes, qui en était une composante. Depuis le 1<sup>er</sup> janvier 2011, le Genes est devenu un établissement public d'enseignement supérieur et de recherche, distinct de l'Insee. En ce qui concerne le CASD, cela a entraîné le transfert au Genes de tâches incombant auparavant à l'Insee : habilitation des gestionnaires, transfert des fichiers, opérations de vérification en sortie, etc. Le Genes définira, en son sein, une structure spécifique de gouvernance pour le CASD. Ces modifications d'ordre administratif n'ont pas eu de conséquences sur l'accès au CASD par les chercheurs, pour lesquels cette évolution a été « transparente ». ■