



Conseil national  
de l'information statistique

Paris, le 6 mars 2018 – n° 21/H030

## ASSEMBLÉE PLÉNIÈRE

---

31 janvier 2018

---

ASSEMBLÉE PLÉNIÈRE

- 31 janvier 2018 -

---

Président : Patrice DURAN, Président du Conseil national de l'information statistique

INTRODUCTION.....	5
I. L'ACCÈS AUX DONNÉES .....	5
II. LA STATISTIQUE PUBLIQUE.....	12
III. DISCUSSION ET VOTE DES ORIENTATIONS ET DE L'AVIS SUR LE PROGRAMME STATISTIQUE 2018 .....	16
CONCLUSION.....	20

## Liste des participants

Cédric AFSA	Conseil national de l'information statistique (CNIS)
Marie-Hélène AMIEL	Ministère de l'Intérieur – Département des statistiques, des études et de la documentation
Isabelle ANXIONNAZ	Conseil national de l'information statistique (CNIS)
Brigitte BACCAÏNI	Commissariat général à l'égalité des territoires (CGET)
Catherine BELLER	Conseil national de l'information statistique (CNIS)
Monique BOURBIGOT-POGNAT	Insee - Unité affaires juridiques et contentieuses (UAJC)
Antoine BOZIO	Institut des politiques publiques
Jean-Claude BOURQUIN	Union fédérale des consommateurs (UFC – Que choisir ?)
Sébastien CANCEL	Conseil national de l'information statistique (CNIS)
Nathalie CARON	Ministère de l'Enseignement supérieur, de la recherche et de l'innovation – Direction de l'évaluation, de la prospective et de la performance (DEPP)
Chantal CASES	Insee – Direction des Statistiques démographiques et sociales (DSDS)
Christine CHAMBAZ	Ministère de la Justice – Sous-direction de la statistique et études
Christine CHOIGNOT	Union nationale interfédérale des œuvres et des organismes privés sanitaires et sociaux (UNIOPSS)
Jean-Pierre CLING	Insee – Direction de la Méthodologie et de la coordination statistique et internationale (DMCSI)
Christel COLIN	Insee – Direction des statistiques d'entreprises (DSE)
Thibaut DE SAINT POL	Institut national de la jeunesse et de l'éducation populaire (INJEP)
Nicolas DÉMOULIN	Assemblée nationale
Raoul DEPOUTOT	Ministère de l'Economie et des Finances – Direction générale des douanes et des droits indirects (DGDDI)
Nicolas DROMEL	Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation - Direction générale de la recherche et de l'innovation (DGRI)
Marie-Michèle DUBOIS	Conseil national de l'information statistique (CNIS)
Patrick DUCHATEAU	Mouvement des entreprises de France - Fédération française du bâtiment
Patrice DURAN	Ecole normale supérieure
Sylvie EGHBAL-TÉHÉRANI	Insee – Direction de la Méthodologie et de la coordination statistique et internationale (DMCSI)
Lionel FONTAGNÉ	Université Paris 1 Panthéon-Sorbonne
Jacques FOURNIER	Banque de France
Adrien FRIEZ	Ministère de l'Economie et des finances - Direction générale de l'administration et de la fonction publique (DGAFP)
Claudine GASNIER	Autorité de la statistique publique (ASP)
Laurent GASNIER	Conseil économique social et environnemental (CESE)
Pierre-Yves GEOFFARD	Ecole d'économie de Paris
Brigitte GOTTI	Chambre de commerce et d'industrie de Paris-Ile-de-France
Michel ISNARD	Insee - Inspection générale
Sylvie LAGARDE	Insee – Direction de la Méthodologie et de la coordination statistique et internationale (DMCSI)
Brice LEPETIT	Ministère de l'Economie et des finances – Direction générale des finances publiques (DGFIP)
Jean-Louis LHÉRITIER	Ministère des Solidarités et de la santé – Direction de la recherche, des études, de l'évaluation et des statistiques (DREES)
Yara MAKDESSI	Conseil national de l'information statistique (CNIS)
Michaël MALIN	Conseil national de l'information statistique (CNIS)
Françoise MAUREL	Conseil national de l'information statistique (CNIS)
Sylvain MOREAU	Ministère de la Transition écologique et solidaire - Service de la donnée et des études statistiques (SDES)
Rodolphe NARGEOT	Conseil national de l'information statistique (CNIS)
Xavier NIEL	Ministère de l'Intérieur - Direction générale des collectivités locales (DGCL)
Vladimir PASSERON	Insee – Direction des Statistiques démographiques et sociales (DSDS)

Claire PLATEAU	Insee – Direction des Etudes et synthèses économiques (DESE)
Dominique POLTON	Institut des données de santé
Gilles POUZIN	Confédération française des travailleurs chrétiens (CFTC)
Nicolas PROKOVAS	Confédération générale du travail
Corinne PROST	Ministère du Travail – Direction de l'animation de la recherche, des études et des statistiques (DARES)
Marina RANDRIAMISAINA	Banque de France
Patrick REDOR	Insee - Unité affaires juridiques et contentieuses (UAJC)
Olivier REDOULES	Mouvement des entreprises de France (MEDEF)
Benoît ROBIN	Confédération générale du travail - Force ouvrière (CGT-FO)
Isabelle SAVELLI-THIAULT	Chambre de commerce et d'industrie de Paris-Ile-de-France
Jean-Christophe SCIBERRAS	Solvay
Jean-Eric SCHOETTL	Conseil d'État
Yvon SÉRIEYX	Union nationale des associations familiales (UNAF)
Roxane SILBERMAN	Comité de concertation pour les données en SHS (CCDSHS) - Réseau Quetelet
Stéphane TAGNANI	Conseil national de l'information statistique (CNIS)
Sophie TASQUÉ	Confédération générale du travail - Force ouvrière (CGT-FO)
Jean-Luc TAVERNIER	Insee - Direction générale
Xavier TIMBEAU	Observatoire français des conjonctures économiques (OFCE)

## INTRODUCTION

**Patrice DURAN, Président du Cnis**, adresse ses meilleurs vœux pour l'année 2018 à l'ensemble des participants à cette assemblée plénière, en souhaitant que le programme du Cnis réponde à leurs attentes et leur permette d'envisager des possibles – les hommes du possible étant, selon Robert Musil, par opposition aux hommes du réel, ceux qui, à partir de la connaissance qu'ils ont d'une société donnée, en envisagent les possibles.

Parmi les changements intervenus au sein du Conseil, il mentionne la désignation de Claude Raynal, sénateur, en tant que Président de la CNERP, ainsi que celle de Xavier Timbeau, de l'OFCE, en tant que Président de la commission « Environnement et développement durable ». Il annonce également la nomination au sein du Conseil de Nicolas Démoulin (au titre de l'Assemblée nationale), Fabien Gay (au titre du Sénat), Guillaume Duval (au titre du CESE), Brigitte Gotti (au titre des Chambres de commerce et d'industrie), Isabelle Savelli-Thiault (au titre de représentante suppléante des Chambres de commerce et d'industrie) et Yvon Sérieyx (au titre de représentant suppléant de l'UNAF).

### I. L'ACCÈS AUX DONNÉES

*Participant à cette table ronde :*

- *Françoise MAUREL, Insee, Directrice de la diffusion et de l'action régionale*
- *Patrick REDOR, Insee, Secrétaire du Comité du secret statistique*
- *Dominique POLTON, Institut national des données de santé, Présidente*
- *Jacques FOURNIER, Banque de France, Directeur général des statistiques*
- *Antoine BOZIO, Institut des politiques publiques, Directeur*
- *Nicolas DROMEL, Ministère de l'enseignement supérieur, de la recherche et de l'innovation, Direction générale de la recherche et de l'innovation*

**Françoise MAUREL** rappelle que l'Insee s'est engagé, dès 2003, dans une politique de mise en ligne gratuite de l'ensemble de ses résultats statistiques. L'objectif était ainsi de mettre à disposition du plus grand nombre un service gratuit et de qualité en matière d'accès aux données, grâce à l'abaissement des coûts marginaux de diffusion permis par Internet. Le site internet de l'Insee, constituant le vecteur principal de cette politique de diffusion, recueille aujourd'hui près de 30 millions de visites par an et met à disposition un ensemble d'informations diffusées dans les meilleurs délais : des indicateurs de conjoncture, des études et analyses, des ouvrages de références, ainsi que des données détaillées. Cette palette de produits en libre accès vise à répondre aux besoins de différents publics : particuliers, chercheurs, professionnels spécialistes de la donnée, administrations, organismes ayant besoin de statistiques (et notamment de statistiques locales), etc. Par ailleurs, les données mises à disposition s'accompagnent toujours de métadonnées, c'est-à-dire de la documentation nécessaire à leur compréhension.

En complément des produits grand public diffusés (chiffres clés ou données de cadrage), s'agissant de fournir aux grands utilisateurs un accès autonome à des données brutes et à leurs métadonnées, l'Insee poursuit également un objectif de progression dans l'*open data*, tant au plan juridique qu'au plan technique, à travers le développement d'API ou de web services. De tels outils sont aujourd'hui proposés autour de séries chronologiques ou de données macroéconomiques. A terme, l'objectif serait d'élargir ce dispositif à d'autres types de données, pour faciliter l'accès à des statistiques localisées notamment.

Au-delà de la production des données, d'importants moyens sont ainsi mobilisés par l'Insee pour faire fonctionner et faire évoluer régulièrement les infrastructures nécessaires à leur diffusion – l'ouverture des données nécessitant de surcroît de préparer et de structurer celles-ci.

La diffusion de données individuelles demeure toutefois contrainte par les règles du secret statistique. Afin de ne pas diffuser de données susceptibles de permettre une identification directe ou indirecte des unités statistiques, l'Insee est parfois amené à réduire, *ex ante*, le degré de détail de l'information mise en ligne. La diffusion de micro-données ou de données individuelles s'articule ainsi autour de trois formats :

- des fichiers de données individuelles anonymisées (issues de l'enquête Emploi et du recensement de la population notamment), diffusés en ligne avec un niveau de détail relativement faible pour éviter tout risque de réidentification ;
- des fichiers dits de production et de recherche, recouvrant des données détaillées d'enquêtes auprès des ménages, diffusés par le biais du réseau Quetelet à destination des chercheurs en sciences sociales, après avis du Comité du secret statistique ;
- des données confidentielles, directement ou indirectement nominatives, diffusées à travers le CASD, après avis du Comité du secret statistique.

Suite à la promulgation de la loi pour une République numérique, l'Insee a par ailleurs mis fin aux redevances de réutilisation du répertoire SIRENE. La base de données du répertoire a ainsi été mise en ligne en janvier 2017, sur le site [data.gouv.fr](http://data.gouv.fr), et fait l'objet depuis d'une mise à jour quotidienne. S'agissant d'une base relativement complexe, le téléchargement d'extractions plus réduites sur le site [sirene.fr](http://sirene.fr) a également été prévu. Ces nouvelles modalités semblent avoir répondu aux attentes des utilisateurs, avec un nombre de téléchargement très élevé.

Parallèlement à la mise en œuvre de sa propre stratégie de diffusion, l'Insee accompagne également d'autres travaux relatifs à l'ouverture des données aux chercheurs, à travers la fourniture permanente de moyens au Cnis et au Comité du secret statistique notamment. L'Insee a ainsi contribué récemment aux réflexions menées par Antoine Bozio et Pierre-Yves Geoffard sur l'accès des chercheurs aux données administratives.

**Patrick REDOR** rappelle que le Comité du secret statistique, dans le cadre défini par la loi de 1951 et un décret de mars 2009, a pour mission d'émettre des avis sur les demandes d'accès aux données individuelles relatives aux ménages ou aux entreprises couvertes par le secret, qu'elles proviennent de sources statistiques, fiscales ou administratives. Le Comité se fonde pour cela sur la nature et l'intérêt des travaux envisagés, l'adéquation du volume des informations demandées, ainsi que la qualité et les garanties apportées par l'organisme demandeur.

En cas d'avis favorable, c'est-à-dire dans la grande majorité des cas, un engagement nominatif de confidentialité est signé par le demandeur. Un accord est également donné formellement par le producteur des données. *In fine*, une autorisation d'accès est délivrée, par la DGFIP pour les données fiscales et par les Archives pour toutes autres données.

Les modalités techniques d'accès aux données peuvent ensuite être variables. La majorité des données nécessitent d'être consultées *via* le CASD, s'agissant notamment des données issues de sources statistiques ou fiscales. Certains producteurs, dont le service statistique des douanes, font néanmoins le choix de mettre à disposition directement leurs données.

Au cours des trois années écoulées, le Comité du secret statistique a enregistré une forte hausse de son activité – les travaux du Comité étant menés au rythme de 4 séances plénières (consacrées aux demandes initiales) et 8 consultations numériques intermédiaires (réservées aux demandes complémentaires) par an. Près de 300 demandes d'accès ont ainsi été formulées en 2017, dont un tiers pour de nouveaux projets, soit une augmentation globale de 5 %, après deux augmentations successives de 17 % et 20 % en 2016 et 2015. Cette évolution, si elle témoigne d'un développement de l'accès aux données et d'un intérêt accru pour les sources statistiques ou fiscales, impacte le fonctionnement du Comité du secret statistique, dont le secrétariat est assuré par l'Unité des affaires juridiques et contentieuses de l'Insee, ainsi que les services producteurs, fortement impliqués dans l'instruction préalable des demandes d'accès.

Dans ce contexte, un allègement des procédures d'accès pour les chercheurs a été engagé, en vue notamment de réduire le nombre des demandes complémentaires. Un accès dès la première demande à l'ensemble des millésimes d'une source a ainsi été ouvert. La durée initiale d'habilitation des chercheurs a également été augmentée d'un an. Enfin, la nécessité de fournir une lettre d'appui pour les doctorants a été supprimée.

En parallèle, l'accent a été mis sur la dématérialisation des procédures, dans l'optique de faciliter l'instruction des demandes et d'alléger les charges correspondantes. Fin novembre 2017, le Comité du secret statistique a été doté d'un site internet autonome, intégrant une rubrique consacrée à l'actualité de l'instance, un référentiel des sources (construit en lien avec le CASD et la DDAR de l'Insee), ainsi qu'une foire aux questions. Une collaboration a également été engagée avec le CASD et la DGFIP, en vue d'aboutir, à l'horizon du 1<sup>er</sup> semestre 2018, à la mise en place d'un portail d'accès aux données, reposant

sur des procédures entièrement dématérialisées (pour la saisie des formulaires de demande notamment) et un système de comptes d'accès permanents pour les chercheurs (impliquant un engagement de confidentialité unique pour l'ensemble des sources, qu'elles soient statistiques, fiscales ou mixtes).

Dans le cadre de la mise en œuvre de la loi pour une République numérique, des travaux ont également été engagés autour de l'accès aux données administratives (à l'occasion d'un séminaire spécifique organisé en janvier 2017 notamment), ainsi qu'autour de l'accès aux fichiers de production et de recherche.

Vis-à-vis des sources administratives, le choix a été fait de permettre aux organismes en faisant la demande d'obtenir un avis unique pour la consultation de leurs données par le biais du Comité du secret statistique. L'Acoss s'est notamment inscrite dans cette démarche. Par ailleurs, toujours au titre de l'accès aux bases administratives, le RICA et les DMMO ont été insérés dans le périmètre de compétence du Comité du secret statistique.

Vis-à-vis des fichiers de production et de recherche, considérés comme des fichiers à diffusion restreinte et ayant vocation à intégrer le périmètre du Comité du secret statistique, une procédure est en cours de définition, qui devrait être opérationnelle dans le courant du 1<sup>er</sup> semestre 2018 et reposer sur une habilitation permanente des demandeurs et un accès *via* le réseau Quetelet ou les services producteurs.

**Dominique POLTON** souligne qu'avant la loi pour une République numérique, la loi de modernisation du système de santé de janvier 2016, en impulsant la création du Système national des données de santé (SNDS), a profondément transformé l'accès aux données administratives de santé.

Le SNDS, destiné à rassembler différentes bases administratives, pour les enrichir et leur appliquer des modalités de gouvernance et d'accès uniques, a ainsi été structuré autour d'une base relative à la médecine de ville et d'une base hospitalière – ces deux bases, une fois chaînées, permettant de décrire très finement les parcours de soins de 67 millions de Français, avec un historique de 10 ans. Les causes médicales de décès ont ensuite été intégrées au dispositif. A terme, deux autres sources devraient également être chaînées : l'une relative au handicap et l'autre regroupant, au moins sur un échantillon, les données de l'assurance maladie obligatoire et de l'assurance maladie complémentaire.

En parallèle, trois modalités d'accès aux données de santé ont été définies :

- un accès en *open data* pour les données considérées comme non-identifiantes ;
- un accès permanent limité à un échantillon au 100<sup>ème</sup> pour un certain nombre d'organismes exerçant des missions de service public (l'Inserm, les équipes de recherche des CHU et des Centres de lutte contre le cancer, etc.) ;
- un accès sur projet.

Afin de faciliter les accès sur projet, en permettant le déploiement de procédures simplifiées pour les travaux de recherche, d'étude et d'évaluation, une fusion des chapitres IX et X de la loi informatique et libertés a été opérée. L'utilisation du NIR, requérant précédemment un décret en Conseil d'Etat, devrait également être facilitée, à travers la mise en place d'un processus reposant sur des tirs d'appariement, en cours de discussion avec la Cnil.

En vue de raccourcir les délais d'accès sur projet, pour les chercheurs et, au-delà, pour l'ensemble des utilisateurs aux finalités considérées comme légitimes, un dispositif piloté par la Drees et reposant sur un guichet unique et une procédure totalement dématérialisée a été mis en place le 28 août 2017. L'Institut national des données de santé (INDS) dispose ainsi d'un délai de 7 jours pour vérifier la complétude des dossiers de demande d'accès, avant leur transmission au Comité d'expertise pour les recherches, les études et les évaluations dans le domaine de la santé (CEREES). Celui-ci dispose ensuite d'un délai d'1 mois pour rendre un avis à la Cnil, tenant compte du rapport entre les données demandées et la finalité des travaux envisagés, de la méthodologie de ceux-ci, voire de leur qualité scientifique (pour les travaux à caractère scientifique). A défaut de réponse dans le délai prévu, cet avis est réputé positif. La Cnil dispose ensuite d'un délai de 2 mois renouvelable pour accorder ou non une autorisation d'accès – des réflexions étant en cours pour raccourcir ce délai.

L'INDS, en tant que GIP associant l'ensemble des familles d'utilisateurs, a ainsi pour rôle de faciliter les accès aux données de santé, en accompagnant les utilisateurs dans leurs demandes et en étant à l'écoute de leurs besoins (en termes d'appariements notamment). Il contribue également à l'expression des besoins en matière d'accès en *open data*. Le cas échéant, sur saisine de sa propre gouvernance, de la Cnil ou du

ministère en charge de la santé, il est par ailleurs appelé à émettre un avis sur le caractère d'intérêt public des projets qui lui sont soumis, en s'appuyant sur un comité d'experts (dont la présidence a été confiée à Jean Gaeremynck).

A ce jour, 194 projets ont été déposés sur la plateforme de l'INDS, dont 55 sollicitant un accès à des données du SNDS – l'INDS ayant également vocation à examiner les demandes d'accès ou de réutilisation ayant trait à d'autres sources (parmi lesquelles les dossiers médicaux). La majorité des demandes formulées par des équipes de recherche l'ont été par des CHU ou des centres de lutte contre le cancer ; par comparaison, l'Inserm, le CNRS, les universités et les écoles ont formulé des demandes beaucoup moins nombreuses – l'accès à l'échantillon généraliste des bénéficiaires dont bénéficiaient précédemment les équipes du CNRS n'ayant par ailleurs pas été maintenu dans le nouveau cadre réglementaire, faute d'avoir fait l'objet, jusqu'ici, d'une réelle manifestation d'intérêt. Deux tiers des avis rendus par le CEREES se sont avérés favorables, le cas échéant avec des recommandations assorties ; les avis réservés ou défavorables sont ainsi demeurés minoritaires – un avis réservé laissant la possibilité au demandeur de déposer son dossier en l'état auprès de la Cnil ou de retravailler son dossier pour solliciter un nouvel examen par le CEREES. 17 autorisations ont d'ores et déjà été délivrées par la Cnil, dont la moitié portant sur des données du SNDS, dans des délais raisonnables (de 30 à 76 jours). L'accès aux données en tant que tel a ensuite vocation à s'effectuer à travers l'ouverture d'espaces dédiés (le cas échéant sur le site de la CNAM) ou *via* le CASD, avec un impact potentiel sur les délais d'accès.

A terme, la question des appariements nécessitera d'être approfondie. Une plateforme est également en cours de mise en place par l'Inserm – la recherche ayant intérêt à mutualiser un certain nombre de compétences autour de ces bases. Des procédures simplifiées sont par ailleurs à l'étude, qui permettraient à l'INDS, pour certaines bases telles que l'échantillon au 100<sup>ème</sup> et pour certains projets déjà examinés par des comités scientifiques, d'ouvrir directement des accès sans solliciter des autorisations multiples auprès de la Cnil.

**Jacques FOURNIER** explique qu'en matière d'accès aux données, la Banque de France poursuit deux objectifs principaux : d'une part, assurer la plus grande transparence possible sur les données relatives au système financier et à l'économie et, d'autre part, soutenir la recherche (et notamment la recherche française, parfois moins riche que la recherche américaine dans le domaine financier).

Dans cette optique, trois niveaux principaux d'information ont été définis. Au niveau macro, la Banque de France diffuse par le biais de son site internet des « Stat info », correspondant à des statistiques générales très agrégées (sur le crédit à l'économie, l'épargne, etc.). Au niveau semi-agrégé, la Banque de France diffuse un certain nombre de séries sur son site baptisé « Webstat Banque de France ». En novembre 2016, la Banque de France a également mis en place une Open Data Room, permettant la diffusion gratuite aux chercheurs de données détaillées issues de près de 700 millions de séries.

A ce jour, une trentaine d'équipes de recherche françaises et européennes, issues par exemple du Think Tank Bruegel, du CEPII, d'HEC ou encore de l'École Polytechnique, ont été accueillies dans l'Open Data Room.

En application du régime juridique européen des Banques Centrales, les données mises à disposition dans l'Open Data Room font l'objet d'une anonymisation. Elles proviennent de différentes sources statistiques recouvrant un périmètre assez vaste, avec des données relatives aux établissements de crédit, aux marchés monétaires et aux titres (avec l'ensemble des transactions réalisées quotidiennement par les banques françaises notamment), aux entreprises (avec les bilans de l'ensemble des entreprises au chiffre d'affaires supérieur à 50 000 euros notamment) et aux ménages (avec des données sur les paiements, le surendettement, etc.).

En termes de procédures, les demandes d'accès à l'Open Data Room, saisies en ligne, sont examinées dans des délais très courts par un comité rassemblant des experts internes et externes à la Banque de France – l'objectif étant de veiller à ce que les accès soient autorisés à des fins de recherche, en vue d'une publication, et non au profit d'organismes commercialisant des données.

Au sein de l'Open Data Room, un accompagnement gratuit est ensuite proposé aux équipes de recherche, avec une assistance dans le paramétrage des outils informatiques, ainsi qu'une assistance technique de la part d'un spécialiste des statistiques du domaine. A ce jour, l'accueil est assuré sur site, au sein des locaux de la Banque de France, avec la mise à disposition de l'ensemble des logiciels nécessaires à l'utilisation des données.

A terme, un accès à distance pourrait également être envisagé, dans des conditions juridiques et de sécurité adaptées. Le dispositif pourrait par ailleurs être développé à l'international, en relation avec d'autres Banques Centrales – le site internet de la Banque de France ayant déjà vocation à abriter prochainement un site d'information sur les accès aux données mis en place par les différentes Banques Centrales.

**Antoine BOZIO** constate qu'au cours des dernières années, des avancées très significatives ont été obtenues dans l'accès des chercheurs aux données, avec la mise à disposition, parfois de manière désordonnée, au fil des évolutions législatives catégorielles, d'un grand nombre de données de très grande qualité. Il y voit un facteur d'attractivité et de compétitivité pour la recherche et les travaux scientifiques en France, vis-à-vis des pays nordiques et des États-Unis notamment.

Il souligne qu'en parallèle, d'importants progrès ont été faits par l'Insee dans la construction de bases appariées accessibles *via* le CASD (avec l'EDP++, Filosofi, Fideli, etc.). Il note que la statistique publique s'affranchit ainsi de plus en plus de la frontière entre les données d'enquêtes et les données administratives. Il évoque également des progrès réalisés autour de l'ouverture, de la documentation et de l'utilisation des données fiscales, en lien avec la DGFIP. Il salue par ailleurs les efforts fournis par le Cnis et le Comité du secret statistique pour faciliter et fluidifier les procédures d'accès aux données.

Cela étant, il observe qu'hormis l'Acoss, peu d'administrations se sont inscrites dans le dispositif d'accès aux données *via* le Comité du secret statistique prévu par la loi pour une République numérique. Il constate que la CNAF, détentrice de nombreuses données utiles et documentées, continue ainsi de privilégier une procédure propre, reposant sur un conseil scientifique interne, sans que cette procédure permette de répondre favorablement aux demandes répétées des chercheurs. Il évoque des difficultés similaires pour accéder à certaines données administratives de la sécurité sociale. Il estime qu'une réflexion nécessiterait d'être menée sur les freins à lever et l'effort d'acculturation à produire pour encourager l'ouverture par les administrations de l'accès à leurs données *via* le Comité du secret statistique – ce dispositif, bien que non-obligatoire, ouvrant des perspectives extrêmement intéressantes pour la recherche.

Pour ce qui est des suites à donner au rapport sur l'accès des chercheurs aux données administratives, il confirme la nécessité d'éviter une multiplication des procédures d'accès, en incitant davantage les administrations à s'appuyer sur le Comité du secret statistique.

Il pointe également l'importance d'envisager les modalités d'accès communes à mettre en place pour les données relevant de cadres juridiques différents, s'agissant des données bancaires et fiscales d'entreprises notamment, ayant vocation à être utilisées à des fins de recherche et/ou d'évaluation des politiques publiques.

Il évoque par ailleurs un enjeu d'harmonisation des pratiques en matière d'archivage des données administratives, pour permettre la poursuite des travaux menés par les historiens notamment. A cet égard, il souligne que les recommandations de la Cnil en matière de non-conservation d'un certain nombre de données ne sauraient être assimilées à une obligation de destruction – la destruction de ces données demeurant soumise à l'autorisation des Archives.

Enfin, il insiste sur la nécessité de poursuivre les réflexions menées sur les appariements à des fins de recherche entre différentes sources administratives, avec en point de mire la question du tiers de confiance à mobiliser. Il espère notamment, dans le prolongement des réflexions engagées avec les chercheurs par la Dares et la Depp, des avancées autour de l'appariement des données relatives à l'emploi, aux salaires et à l'éducation, recouvrant des enjeux forts en termes de recherche et d'évaluation des politiques publiques.

**Nicolas DROMEL** observe que l'accès des chercheurs aux données administratives repousse aujourd'hui les frontières et les possibilités de production de la recherche. Le volume, la richesse et la précision de ces données permettent d'appréhender et d'analyser des caractéristiques individuelles et des phénomènes sociaux que nous ne pourrions pas observer autrement. Ces données sont aussi l'occasion de tester des théories de façon plus robuste et fiable que sur des échantillons plus restreints, collectés à des niveaux plus agrégés. De façon très intéressante, l'accès des chercheurs aux données administratives développe l'interface entre la science et la décision publique, et dynamise un cercle vertueux au bénéfice mutuel de la recherche, de l'action publique et de la société. Les données nourrissent la recherche, en y donnant parfois un avantage comparatif majeur. Les résultats de la recherche peuvent éclairer les débats démocratiques ainsi que l'action publique. Les enseignements de ces recherches sont pour l'État un moyen de gagner en efficience, en identifiant les politiques efficaces et celles qui le sont moins. La poursuite des débats

démocratiques et des choix publics appellent alors de nouvelles recherches, nécessitant un retour aux données, et ainsi de suite.

L'accès des chercheurs aux données administratives développe ainsi le potentiel d'impact multidimensionnel de la recherche, au-delà de l'aspect scientifique *stricto sensu*. Le Ministère de l'enseignement supérieur, de la recherche et de l'innovation (MESRI) y est donc particulièrement attentif.

Il souligne que l'accès des chercheurs aux données administratives rejoint deux dimensions stratégiques importantes de la politique nationale en matière de recherche : la stratégie nationale développée autour des infrastructures de recherche (en sciences humaines et sociales notamment), ainsi que les engagements pris par la France en matière de développement d'une science ouverte.

Vis-à-vis des infrastructures de recherche en sciences humaines et sociales, s'agissant à la fois de développer de nouvelles compétences disciplinaires et technologiques, de donner un accès optimisé aux données et aux ressources documentaires, d'offrir des services d'accompagnement aux chercheurs, d'assurer un développement territorial et de favoriser l'interdisciplinarité, une feuille de route nationale a été établie, s'articulant autour d'un certain nombre de projets ou de programmes : la très grande infrastructure de recherche pour les humanités numériques HUMA-NUM, la très grande infrastructure de production et de gestion de données PROGEDO, le Réseau National des Maisons des Sciences de l'Homme, la plateforme d'édition électronique et de communication scientifique ouverte en libre accès OpenEdition, ainsi que le projet d'infrastructure d'analyse, d'interprétation et de conservation des items du patrimoine naturel et culturel ERIHS-FR.

La très grande infrastructure de recherche PROGEDO a ainsi vocation, sous tutelle de l'EHESS et du CNRS, à développer une culture des données, à améliorer l'accès aux données (dans un cadre conforme à la législation), à permettre aux communautés de recherche d'utiliser au mieux les données disponibles et à renforcer la position de la France dans l'espace européen de la recherche (en soutenant notamment la participation de la France à de grandes enquêtes internationales telles l'European Social Survey ERIC, le SHARE ERIC, l'EVS, le GGP ou encore l'ISSP et en fournissant des services dans le cadre du consortium européen CESSDA ERIC). L'accès aux données françaises concerne le département Quetelet-PROGEDO Diffusion, regroupant des données issues d'enquêtes quantitatives et qualitatives de la statistique publique et de la recherche académique. L'accès à certaines données confidentielles est mis en œuvre par le CASD. Les plateformes universitaires de données (PUD) labellisées par PROGEDO contribuent à promouvoir l'utilisation des données dans les travaux en SHS à l'échelle locale des universités.

Nicolas Dromel souligne ensuite que les enjeux de la science ouverte ne sont pas restreints à la publication ouverte puisqu'ils concernent aussi les données, les processus et l'évaluation. Bien évidemment, une politique de science ouverte efficace se doit de maîtriser les enjeux de l'accès aux données confidentielles. Les autorités françaises ont réaffirmé, à travers une note publiée fin novembre 2017 concernant le prochain programme cadre de recherche et d'innovation de l'UE, le principe d'une approche « aussi ouverte que possible et aussi fermée que nécessaire ».

La Commission européenne a lancé un projet de cloud européen pour la science ouverte (EOSC), visant à structurer un réseau européen d'organisations et d'infrastructures soutenant la création ouverte et la diffusion des savoirs, à travers la mise en place d'une plateforme ouverte pour l'échange de données de recherche, en facilitant les collaborations de recherche en Europe, ainsi qu'en facilitant l'utilisation et la réutilisation des données et des résultats de la recherche à l'échelle européenne.

Pour tenir compte des enjeux relatifs à la structuration, à la conservation et à la « fairisation » des données de recherche, la France a par ailleurs amorcé, avec les Pays-Bas et l'Allemagne, l'initiative GO FAIR – la fairisation consistant à transformer les données pour les rendre faciles à trouver, accessibles au travers de règles explicites, interopérables et réutilisables.

Dans le cadre du G8, la France a également signé, en 2013, une charte relative à l'Open Data, prévoyant une ouverture par défaut des données publiques, dans le respect d'un principe de gratuité d'utilisation et de réutilisation et en privilégiant les formats ouverts et non-propriétaires.

La France a ensuite rejoint, en 2014, le partenariat pour un gouvernement ouvert (PGO), regroupant aujourd'hui 75 pays et plusieurs centaines d'organisations issues de la société civile agissant pour la transparence de l'action publique à travers le monde (*via* l'ouverture des données publiques notamment). Pour concrétiser un de ses engagements pris dans le cadre du PGO (« Construire un écosystème de la

Science Ouverte »), la France s'est notamment dotée d'un Comité pour la Science Ouverte. Ce comité d'experts – présidé par le DGRI et organisé autour de quatre collèges (publications ; compétences ; Europe et international ; données de la recherche), avec un comité technique et un comité de pilotage – aura en charge d'instruire des réflexions, de proposer des orientations et d'impulser et accompagner des actions.

**Nicolas DROMEL** ajoute que le Ministère de l'enseignement supérieur, de la recherche et de l'innovation est attentif aux recommandations du rapport Bozio-Geoffard et souhaite que leur implémentation soit régulièrement discutée en cohérence avec la politique nationale en matière de recherche. Il insiste sur la nécessité de développer ainsi une stratégie globale coordonnée, articulant l'accès des chercheurs aux données administratives, la stratégie nationale des infrastructures de recherche et la politique nationale de science ouverte, pour éviter une fragmentation des initiatives et, *in fine*, veiller à la fois à la simplicité des procédures d'accès aux données, à la qualité des données proposées, ainsi qu'au développement des compétences de la science ouverte (en matière de structuration et de diffusion des données de la recherche notamment).

**Xavier TIMBEAU, représentant l'OFCE**, salue les efforts fournis par la statistique publique et les administrations pour faciliter l'accès des chercheurs aux données. Cependant, il s'interroge sur les données susceptibles d'être utilisées pour évaluer la mise en place, dans le cadre du projet de loi de finance pour 2018, du prélèvement forfaitaire unique et de l'impôt sur la fortune immobilière. Il souligne que, pour conduire de telles évaluations, des données plus largement échantillonnées et plus détaillées que celles de l'enquête sur les revenus fiscaux et sociaux nécessiteraient de pouvoir être mobilisées, avec des informations complémentaires sur les stocks d'actifs détenus par les ménages. Au vu des enjeux budgétaires et démocratiques associés, il évoque la possibilité d'inviter les administrations compétentes à reconsidérer leur application du secret statistique, le cas échéant dans le cadre d'une procédure d'accès plus lourde. De manière générale, autour de tels sujets, il s'interroge sur la possibilité d'opposer au secret statistique la nécessité d'informer les citoyens.

**Jean-Luc TAVERNIER, Directeur général de l'Insee**, souligne qu'autour de tels sujets, il reviendrait au législateur d'envisager la possibilité pour les assemblées parlementaires de se doter de capacités d'expertise propres. Dans l'état actuel des textes, il rappelle néanmoins la possibilité, pour conduire les évaluations évoquées, de mobiliser les données de l'enquête Patrimoine. Il mentionne également la possibilité d'accéder aux données relatives à l'impôt sur la fortune, sur décision du ministre en charge du budget, après instruction d'une demande par le Comité du secret statistique.

**Brice LEPETIT, représentant la DGFIP**, indique ne pas avoir connaissance d'une demande formulée pour accéder aux données relatives à l'impôt sur la fortune. Il rappelle qu'une évaluation de la réforme de la fiscalité du capital a néanmoins été prévue dans le cadre du projet de loi de finance pour 2018, avec un rapport d'évaluation à remettre avant l'adoption du projet de loi de finance pour 2020. Il souligne que, dans le cadre de l'évaluation du CICE, une démarche similaire a été mise en œuvre, donnant lieu à la mise à disposition de données auparavant inaccessibles. S'agissant d'évaluer la mise en place du prélèvement forfaitaire unique, il évoque un chantier devant aboutir dans les prochains mois, à partir d'une réropolation des données relatives à l'impôt sur le revenu.

**Antoine BOZIO** explique que, par le passé, la DGFIP avait indiqué que la décision ministérielle serait de ne pas donner suite à une demande d'accès aux données relatives à l'impôt sur la fortune. Le contexte ayant évolué, il évoque néanmoins la possibilité de réitérer cette demande.

**Roxane SILBERMAN, représentant les chercheurs**, souligne que l'ouverture à la recherche des données administratives, inscrite en particulier dans la Loi sur la République numérique laisse ouvert le choix des procédures d'accréditation des chercheurs et des modes d'accès sécurisés. On pourrait de ce fait assister à une multiplication des systèmes d'accès aux données confidentielles qui pourrait restreindre la possibilité d'utiliser conjointement celles-ci, à l'image de ce qui se passe déjà pour les données bancaires et les données de la statistique publique. Elle évoque un rapport remis récemment au Congrès publié aux États-Unis, préconisant la mise en place d'une procédure et d'un mode et d'un mode d'accès uniques pour l'accès aux données administratives, en vue de faciliter leur utilisation conjointe.

Elle souhaiterait par ailleurs que soit examinée la question de l'accès à distance aux données confidentielles *via* le CASD au-delà des équipes de recherche localisées dans l'Union européenne. Les demandes émanant des États-Unis et du Canada sont en effet de plus en plus nombreuses, qu'il s'agisse de chercheurs de ces pays ou de chercheurs français y résidant pour pouvoir travailler sur les données

françaises très riches à partir des universités ou institutions de recherche américains ou canadiens sans avoir besoin, comme c'est le cas actuellement de venir en France.

En ce qui concerne l'Union européenne, elle évoque la possibilité que le RGPD qui harmonise la protection des données au sein des États membres puisse faciliter, au-delà des accès à distance, des transferts physiques de données d'un centre d'accès à l'autre, s'agissant de permettre à des équipes de recherche d'utiliser conjointement des données de plusieurs pays à l'image de ce que font déjà plusieurs pays nordiques.

Elle s'interroge enfin sur les moyens qui seront alloués, tant par les administrations que par le ministère de l'enseignement supérieur, de la recherche et de l'innovation, pour accompagner la documentation et la mise à disposition effective des données administratives.

**Gilles POUZIN, représentant la CFTC**, s'étonne que, parmi les demandes formulées auprès de l'INDS, une minorité portent sur les données du SNDS et une majorité portent sur des dossiers médicaux.

**Dominique POLTON** explique que les équipes de recherche médicale, pour réutiliser des données de santé issues de dossiers médicaux, dès lors qu'elles ne peuvent satisfaire les conditions d'une simple déclaration de conformité vis-à-vis des méthodologies de référence, s'agissant notamment de répondre à l'obligation d'information des patients, doivent solliciter une autorisation de la Cnil. Elles sont pour cela amenées à formuler une demande auprès du guichet unique que constitue désormais l'INDS. Ces demandes demeurent majoritaires par rapport à celles portant sur les données du SNDS, malgré les attendus de la loi en matière d'ouverture des données administratives à la société civile (au-delà des chercheurs qui bénéficiaient déjà d'accès aux bases simples).

De manière générale, Dominique POLTON rappelle que les données de santé, compte tenu de leur caractère sensible, nécessitent une autorisation de la Cnil pour être utilisées. A ce titre, elle s'interroge sur la possibilité d'intégrer le domaine de la santé dans un mouvement d'unification des procédures d'accès.

**Benoît ROBIN, représentant la CGT-FO**, salue les progrès importants réalisés au cours des dernières années autour de l'accès aux données, en souhaitant que d'autres pays rejoignent ce mouvement – les pratiques en la matière demeurant hétérogènes. En parallèle, il confirme la nécessité de poursuivre les efforts en matière de formation et d'accompagnement, le cas échéant en s'inspirant de certaines pratiques développées à l'étranger.

**Françoise MAUREL** souligne que, compte tenu du droit existant, la question de l'accès aux données pour les chercheurs internationaux concerne avant tout les producteurs de données. Pour ce qui est du développement des filières d'accès, elle constate des évolutions positives. Elle gage que les réflexions en ce sens pourront être poursuivies dans le cadre du Comité pour la science ouverte, le cas échéant autour des possibilités de transferts de données entre centres d'accès. Elle assure que le Cnis s'attachera par ailleurs à suivre la mise en œuvre des recommandations du rapport Bozio-Geoffard.

**Patrice DURAN** constate que le rapport Bozio-Geoffard a soulevé de nombreuses interrogations et mis en évidence un certain nombre de pistes à explorer. Il confirme que les travaux autour de ces enjeux devront être poursuivis, avec l'appui du ministère de l'enseignement supérieur, de la recherche et de l'innovation.

## II. LA STATISTIQUE PUBLIQUE

**Jean-Luc TAVERNIER, Directeur général de l'Insee**, propose d'apporter un éclairage sur l'activité de la statistique publique en 2017. En préambule, il souligne que la statistique publique, regroupant l'Insee et 16 SSM, produit chaque année davantage de statistiques, en toute indépendance, en dépit d'une maîtrise sinon d'une baisse de ses moyens – les effectifs de l'Insee connaissant une diminution continue depuis plus de 15 ans.

Il ajoute que le principe de coordination des systèmes statistiques nationaux par les instituts nationaux de statistiques, inscrit dans le règlement européen relatif aux statistiques et repris dans le Code de bonnes pratiques de la statistique européenne, se traduit désormais, au sein du système statistique français, par une meilleure articulation des calendriers de publication, par une convergence des modalités de transmission préalable au Gouvernement des informations statistiques, ainsi que par le développement d'un cadre d'assurance qualité commun.

Enfin, il constate qu'en 2017, à l'occasion d'échéances électorales importantes, dans un contexte global marqué par le développement des *fake news*, la statistique publique n'a eu à intervenir qu'en de rares occasions dans le débat public pour rectifier une distance entre la présentation faite et la réalité statistique d'un phénomène. Il y voit un bon signe pour la démocratie française.

### **Le contexte européen et international**

**Jean-Luc TAVERNIER** rappelle qu'une liste de 244 indicateurs de développement durable a été adoptée, en mars 2017, par la Commission statistique de l'ONU puis, en septembre 2017, par l'Assemblée générale de l'ONU. Pour assurer la déclinaison de ces indicateurs au niveau français, un premier tableau de bord comportant une centaine d'indicateurs a été mis en ligne, fin mars 2017, sur le site insee.fr. Un groupe de travail du Cnis a par ailleurs été constitué pour définir les indicateurs les plus pertinents à retenir pour le suivi des politiques françaises contribuant aux objectifs de développement durable de l'ONU.

Au niveau européen, le règlement TERCET – complétant le règlement actuel sur la nomenclature des unités territoriales statistiques (NUTS) par une typologie des unités urbaines ou rurales, côtières ou intérieures, ainsi que par des éléments relatifs au carroyage – a été adopté et devrait être publié prochainement au Journal officiel de l'Union européenne.

Le Comité du système statistique européen (CSSE), dans un contexte marqué par la vacance persistante du poste de Directeur général d'Eurostat, a par ailleurs examiné trois projets de règlement, appelés à suivre leur processus d'adoption législative dans l'ordre chronologique suivant : un règlement relatif aux statistiques d'exploitations agricoles (IFS), devant constituer le fondement législatif du recensement agricole de 2020 ; un règlement relatif aux statistiques sociales (IESS), ayant vocation à couvrir la plupart des enquêtes réalisées par l'Insee auprès des ménages ; un règlement cadre relatif aux statistiques d'entreprises (FRIBS), devant organiser l'échange de micro-données d'entreprises entre les instituts nationaux de statistiques.

Le nouveau règlement européen relatif à la protection des données à caractère personnel devrait quant à lui entrer en vigueur le 25 mai 2018. Dans cette optique, une évolution de la loi informatique et libertés est en cours de discussion au sein de l'Assemblée nationale. Les dérogations dont bénéficie le SSP, pour réaliser des appariements notamment, devraient être reprises dans le projet de loi ainsi proposé. Un certain nombre de dispositions, concernant notamment le traitement des données sensibles ou du NIR en clair, devraient toutefois être renvoyées à un décret-cadre, potentiellement très impactant pour le SSP et vis-à-vis duquel une vigilance devra être exercée. Le passage prévu par le nouveau règlement européen d'un régime de formalités préalables à un régime de responsabilisation des acteurs, s'appuyant sur des *data protection officers*, devrait quant à lui pouvoir être abordé sereinement et être source de simplification – l'examen des projets du SSP par le Comité de label et les commissions du Cnis assurant déjà une transparence vis-à-vis des traitements et de la protection des données individuelles.

### **Le cadre institutionnel français**

**Jean-Luc TAVERNIER** indique que, dans le cadre de la mise en œuvre de la loi pour une République numérique, une réflexion a été engagée au sein de l'Insee sur l'utilisation du NIR haché, en vue d'aboutir à une méthode de cryptage satisfaisant aux exigences de la Cnil. Une première expérimentation devrait ainsi pouvoir être lancée en 2018, en lien avec la Dares.

En matière de diffusion des données en *open data*, le répertoire SIRENE a été mis à disposition dès le 5 janvier 2017, avec une mise à jour quotidienne du stock global d'établissements sur le site data.gouv.fr, ainsi qu'un service gratuit de constitution de listes d'établissements sur le site sirene.fr. Depuis lors, le site sirene.fr a enregistré un doublement de sa fréquentation, à hauteur de 50 000 visites par mois ; près de 4 600 téléchargements de la base de diffusion de SIRENE sont également enregistrés chaque mois.

En matière d'utilisation des données d'entreprises, le cadre juridique et les conditions de transmission des données de caisse de la grande distribution pour la détermination d'une partie de l'indice des prix ont été précisées par un décret et un arrêté, publiés respectivement le 31 mars 2017 et le 13 avril 2017. Un conventionnement avec les grandes enseignes a ensuite été engagé pour assurer la pérennité de cette transmission. La plupart des grandes enseignes se sont ainsi engagées dans une démarche d'alimentation début 2018, en vue de permettre un passage en production d'une partie de l'indice des prix avec un recul suffisant à partir de 2020.

Au niveau du Comité du secret statistique, une procédure a par ailleurs été définie pour sécuriser la diffusion des fichiers de production et de recherche, sans en allonger les délais d'obtention.

Suite à la promulgation de la loi de modernisation de la Justice du 21<sup>ème</sup> siècle, l'enregistrement des conclusions, modifications et dissolutions de PACS par les officiers d'état civil ne devrait pas induire de perte d'informations pour la statistique publique, avec la possibilité de mobiliser le circuit de remontée des actes d'état civil des communes. En revanche, l'enregistrement des divorces par consentement mutuel par les notaires pourrait soulever un risque de rupture de séries.

Dans le cadre de la mise en œuvre de la loi relative à l'égalité réelle outre-mer, un document a été produit pour alimenter le rapport devant être remis par le Gouvernement au Parlement sur le périmètre de calcul du PIB et le calcul des taux de pauvreté en métropole et dans les Outre-mer. En conformité avec les textes européens, seuls la métropole, les 5 DOM et la région ultrapériphérique de Saint-Martin ont été retenus dans le périmètre de calcul du PIB. Pour ce qui est de la mesure de la pauvreté, des précisions ont été apportées sur la pertinence et l'utilisation faite des taux se rapportant au revenu médian de l'ensemble de la France ou de chacun des territoires d'outre-mer. Un rapport complet accessible sur le site Insee.fr a également permis de souligner les efforts particuliers fournis par la statistique publique dans les collectivités d'outre-mer – les carences encore observées à cet endroit renvoyant souvent à la qualité des données administratives disponibles.

Un dernier recensement quinquennal a par ailleurs été opéré à Mayotte, avant l'application des nouvelles modalités de recensement dans ce territoire. Les populations légales de Mayotte ont pu ainsi être diffusées en décembre 2017, peu après celles de la Polynésie française. A Saint-Martin, pour tenir compte des conséquences du passage de l'ouragan Irma, l'ouverture de la période de collecte du recensement a été repoussée à mars 2018.

### **Les principaux chantiers d'innovation du SSP en 2017**

**Jean-Luc TAVERNIER** indique que, dans le champ de l'emploi, les évolutions d'emplois salariés ont donné lieu, en juin 2017, à une première publication commune à la Dares, à l'Insee et à l'Acoss, portant sur l'ensemble du champ salarié – l'Acoss ayant assuré le suivi du secteur marchand non agricole et les compléments ayant été apportés par l'Insee et la Dares.

L'intégration de la DSN a par ailleurs été poursuivie, pour permettre notamment la publication par la Dares et Pôle Emploi de séries sur l'intérim et, à partir de 2019, la publication de données structurelles sur l'emploi et les salaires dans le secteur privé. A terme, la DSN de la fonction publique a également vocation à être intégrée.

Le 9 novembre 2017, dans le prolongement des réflexions menées par un groupe de travail du Cnis présidé par Jacques Freyssinet, la directrice de la Dares et le responsable des statistiques de Pôle Emploi ont annoncé la refonte de la publication Dares-Pôle Emploi, marquant le passage de celle-ci à un rythme trimestriel, avec un maintien de la diffusion en ligne des chiffres mensuels.

Un groupe de travail du Cnis a par ailleurs été constitué autour de la rénovation de la nomenclature des catégories socio-professionnelles.

Dans une optique d'évaluation des politiques publiques, la Dares a également initié le projet TRAJAM, visant le suivi, à travers un panel, de l'accès à l'emploi des jeunes bénéficiant de mesures actives du marché du travail. Ce projet devrait permettre d'expérimenter l'utilisation du NIR haché.

**Jean-Luc TAVERNIER** souligne que, dans le champ de la santé, la mise en place du SNDS a déjà été évoquée lors de cette assemblée.

Dans le champ des statistiques relatives au secteur public, il indique que le SSM de la DGAFP a publié, sur la base des données du portail des offres d'emploi de la fonction publique, de premières statistiques sur les vacances d'emploi dans la fonction publique d'État. La DGAFP a par ailleurs enrichi ses données relatives à l'absentéisme dans la fonction publique et engagé des travaux sur l'évolution de son système d'information.

En vue d'améliorer la connaissance des dépenses publiques au niveau territorial, un travail a été mené par l'Insee autour des dépenses de l'État. Des difficultés ont toutefois été identifiées, en lien avec l'absence de localisation fine des dépenses de l'État dans l'outil Chorus – cet outil étant principalement utilisé à des fins

gestionnaires et non statistiques. A cet égard, un travail de sensibilisation nécessiterait d'être mené par la DGFIIP.

Autour de la connaissance des dépenses des collectivités locales, d'importants investissements méthodologiques ont été consentis par le SSM de la DGCL, pour favoriser la consolidation des comptes entre communes et, à terme, permettre une consolidation entre communes et régions (dans l'optique de produire un compte consolidé des collectivités locales sur le champ régional), voire entre l'État et les collectivités locales.

Autour de la dimension territoriale, l'enquête sur l'occupation et l'utilisation du territoire (TERUTI), mise en œuvre par le SSM du ministère de l'agriculture, a par ailleurs fait l'objet d'une première collecte dans son format rénové – ce nouveau format ayant vocation à améliorer les taux de collecte et à intensifier l'observation dans les zones à forte probabilité de changement.

Dans le champ de l'éducation nationale et de l'enseignement supérieur, un identifiant national élève (INE) unique a été mis en place dans l'ensemble des établissements scolaires sous tutelle du ministère de l'éducation nationale, en vue de permettre un meilleur suivi des élèves tout au long de leur scolarité et, à terme, une meilleure appréhension de leur parcours dans l'enseignement supérieur.

En parallèle, l'accent a été mis sur les indicateurs de suivi des parcours et de la réussite dans l'enseignement supérieur. Depuis l'été 2017, des données sont ainsi diffusées pour les sections de technicien supérieur ; depuis la fin d'année 2017, des données sont diffusées à un niveau fin pour les niveaux licence, 1<sup>ère</sup> année des études de santé et DUT.

Une enquête biannuelle sur l'insertion professionnelle des docteurs a également été mise en place, ayant vocation à impliquer la quasi-totalité des établissements d'enseignement supérieur délivrant des diplômes de doctorat et dont les premiers résultats devrait être disponibles fin 2018.

En réponse au règlement européen sur les indicateurs de court terme, l'Insee a mis en place de nouveaux indices mensuels de volume des ventes dans le commerce et de production dans les services.

En matière de diffusion, l'ergonomie du site internet du SSM des Douanes a été revue. Un *data lake* a été mis en place par le SSM des ministères de la transition écologique et solidaire et de la cohésion des territoires. Une nouvelle collection a été créée par l'INJEP, baptisé « INJEP analyses et synthèses ».

Sur le site insee.fr, les séries chronologiques de la BDM ont été intégrées, de même que les statistiques structurelles d'entreprises. Un nouveau site a également été mis en ligne pour les statistiques locales, avec un nouvel outil de cartographie.

Ces évolutions du site insee.fr semblent aujourd'hui répondre aux besoins des utilisateurs, malgré quelques interrogations, de la part des utilisateurs de séries macroéconomiques longues notamment – le service Insee Contact n'ayant par ailleurs pas enregistré d'augmentation particulière du nombre de demandes d'internautes ne parvenant pas à s'orienter sur le site.

### **La communication du SSP en direction des jeunes**

**Jean-Luc TAVERNIER** indique que, dans le cadre des partenariats développés avec le ministère de l'éducation nationale, une réflexion a été engagée sur la rénovation du site « statapprendre ».

L'Insee s'est par ailleurs investi, avec le concours de l'Inspection générale de l'Éducation nationale et de la Société française de statistiques, dans l'organisation de la première compétition européenne de statistiques ouverte aux classes de lycées. Près de 400 équipes ont candidaté dans le cadre des sélections nationales, programmées le 20 octobre 2017. Le concours à l'échelle européenne se déroulera ensuite au printemps 2018.

Les réflexions seront également poursuivies autour des moyens de communiquer dans un but pédagogique en direction des collégiens.

## **Les publications du SSP en 2017**

**Jean-Luc TAVERNIER** souligne la publication d'un premier numéro d'Insee Références consacré au thème de l'environnement, principalement élaboré par le SDES, avec le concours du SSM du ministère de l'agriculture. Dans le champ de l'environnement, il salue également la publication de chiffres clés du climat et de statistiques relatives à l'utilisation des énergies renouvelables.

Autour des évolutions du tissu productif et de la société, il mentionne : un article publié dans Insee Références « L'économie française » sur l'apport des données massives à la mesure de l'économie (pour l'établissement des diagnostics conjoncturels, la mesure des prix et la quantification de l'économie numérique) ; un document d'étude de la Dares sur l'économie des plateformes et ses impacts potentiels sur la croissance, le travail et l'emploi ; la publication par l'Insee de premières statistiques sur les plateformes d'hébergement des touristes par des particuliers ; une note de conjoncture présentant les résultats d'une enquête sur les freins à l'embauche ressentis par les entreprises (mettant en évidence la perception de freins liés, par ordre décroissant d'importance, aux incertitudes quant à la situation économique, aux difficultés rencontrées pour trouver une main d'œuvre compétente, aux coûts du travail et au droit du travail) ; une étude réalisée par la Dares à partir des résultats de l'enquête OFER sur les caractéristiques des recrutements en CDI ou en CDD.

Dans le domaine de la santé, il signale la publication par la Drees (ayant été dotée d'un nouveau directeur en la personne de Jean-Marc Aubert), des résultats de l'enquête nationale périnatale 2016, mettant en évidence, au cours des 20 dernières années, une forte réduction du nombre de maternités, ainsi que l'augmentation de certains facteurs de risques chez les mères, avec une progression du taux de nouveau-nés prématurés ou de faible poids. Il mentionne également la publication des premiers résultats de l'enquête CARE Domicile.

Dans le domaine de l'éducation, il souligne qu'un numéro de la revue de la Depp a proposé un panorama de l'ensemble des panels de la Depp. Il mentionne également la publication des résultats du dispositif CEDRE, visant la mesure des compétences en langues étrangères en fin d'école et de collège, et du dispositif TIMSS, visant la mesure des compétences en mathématiques.

Dans le champ du secteur public, il évoque la diffusion par la DGAFP d'une étude sur les nouveaux apprentis dans la fonction publique.

Autour des fonctions régaliennes de l'État (Intérieur, Justice et Immigration), il mentionne : une étude sur le parcours des étudiants étrangers arrivés en France entre 2002 et 2016 ; une étude sur l'immigration en France au cours des 45 dernières années (mettant en évidence une augmentation de la part des femmes dans la population immigrée) ; une étude sur les délits commis pour des motifs racistes, xénophobes ou antireligieux ; une étude exploratoire sur les infractions commises en raison de l'origine ou de la religion ; une étude sur la part des femmes parmi les élus locaux (excédant désormais 40 %) ; une publication conjointe Insee/Eurostat sur la vie des femmes et des hommes en Europe (diffusée à l'occasion de la Journée européenne de la statistique).

Dans le domaine de l'agriculture, il évoque la publication d'études sur la mondialisation des systèmes d'approvisionnement et les conduites alimentaires comme reflet de la mondialisation.

Dans le champ de la démographie et des conditions de vie, il mentionne la publication d'un nouveau numéro de « France, portrait social », portant un regard sur les ménages à niveau de vie médian.

Enfin, il souligne que le numéro d'Insee Références consacré aux entreprises en France a apporté un éclairage sur la localisation des emplois auprès des petites ou grandes entreprises. Il ajoute qu'un nouvel outil cartographique a été mis à disposition sur le site de l'Insee, permettant de visualiser la répartition de la population française selon la proximité avec de grands établissements.

### **III. DISCUSSION ET VOTE DES ORIENTATIONS ET DE L'AVIS SUR LE PROGRAMME STATISTIQUE 2018**

**Patrice DURAN** indique qu'en 2018, au-delà du travail des commissions thématiques, le Cnis s'attachera à suivre la mise en œuvre des recommandations du rapport sur l'accès des chercheurs aux données

administratives, en liaison avec le cabinet de la ministre de l'enseignement supérieur, de la recherche et de l'innovation.

Il annonce qu'après l'organisation, en 2017, d'une rencontre sur la mesure des discriminations (ayant donné lieu à la rédaction d'un document de travail essentiel), le Cnis organisera, le 7 mars 2018, une rencontre autour du thème de l'économie numérique.

Il ajoute que le groupe de travail constitué pour examiner la déclinaison française des indicateurs relatifs aux ODD, présidé par Jean-René Brunetière, devrait remettre son rapport au second semestre 2018. Le groupe de travail consacré à la rénovation de la nomenclature PCS, présidé par Thomas Amossé, devrait quant à lui remettre son rapport en avril 2019. Les réflexions sur l'élaboration du prochain moyen terme du Cnis, dont la coordination a été confiée à Cédric Afsa, seront également poursuivies, en lien avec les présidents des commissions thématiques.

Il souligne que le Cnis peut également s'appuyer sur un site internet rénové, plus convivial et plus attractif.

### **Adoption des orientations générales du Cnis pour 2018**

**Patrice DURAN** observe qu'aucune proposition d'amendement n'a été adressée préalablement. Il y voit le signe d'une bonne gestion du moyen terme s'achevant, au cours duquel des ajustements ont pu être opérés dans les orientations et la programmation des travaux du Cnis, pour tenir compte de l'évolution des enjeux associés.

**Françoise MAUREL** indique que les orientations générales du Cnis pour 2018 ont été structurées autour de six avis, en cohérence avec les orientations du moyen terme 2014-2018.

#### **Avis n°1 : Contribuer à faciliter l'accès aux données**

**Françoise MAUREL** précise que cet avis formule un certain nombre de recommandations, en rappelant aussi les avancées obtenues ces dernières années.

**Nicolas PROKOVAS, représentant la CGT**, suggère de faire référence à l'importance de faciliter la compréhension et l'appropriation des données « lors de la publication » de celles-ci, afin de ne pas laisser entendre que certaines données pourraient ne pas être publiées.

**Patrice DURAN** propose de retenir cet amendement.

*Compte tenu des modifications apportées en séance, l'avis n°1 est adopté à l'unanimité.*

#### **Avis n°2 : Améliorer la comparabilité des statistiques au plan international**

**Françoise MAUREL** souligne que cet avis s'inscrit dans le prolongement des travaux et réflexions menés par le Cnis sur ce thème dans le cadre du moyen terme, à travers l'organisation d'un colloque en 2016 notamment.

**Jean-Luc TAVERNIER** observe qu'il demeure complexe de sensibiliser les partenaires de la statistique publique, et notamment Eurostat, aux enjeux de la comparabilité des statistiques. Il gage que l'ESGAB, dont la présidence a été confiée à Enrico Giovannini et au sein duquel Mireille Elbaum a été nommée, devrait néanmoins être sensible à cette préoccupation.

*L'avis n°2 est adopté à l'unanimité.*

#### **Avis n°3 : Mettre le citoyen en capacité de comprendre et d'utiliser les données de la statistique publique**

**Françoise MAUREL** souligne que l'enjeu de cet avis serait de renouveler les recommandations du Cnis à la statistique publique, afin que celles-ci poursuivent les efforts et avancées évoqués ce jour.

**Brigitte GOTTI, représentant les Chambres de commerce et d'industrie**, évoque le risque que la formulation retenue dans le titre de cet avis infantilise le citoyen. Elle estime qu'il conviendrait davantage de

faire référence à la volonté « d'aider le citoyen à comprendre et utiliser les données de la statistique publique ».

**Patrice DURAN** propose de retenir cet amendement.

**Nicolas PROKOVAS** suggère d'insister dans cet avis sur la nécessité de veiller à ce que le SSP dispose des moyens lui permettant d'exercer ses missions et d'alimenter le débat public.

**Jean-Luc TAVERNIER** accueille favorablement cette proposition – le SSP ayant besoin d'un consensus le plus large pour défendre les moyens lui permettant d'exercer ses missions.

**Nicolas PROKOVAS** propose la formulation suivante : « *Le Conseil insiste sur la nécessité pour le SSP de disposer des moyens lui permettant de mener à bien ses missions et renouvelle ses recommandations au service statistique public : [...].* »

**Patrice DURAN** propose de retenir cet amendement.

*Compte tenu des modifications apportées en séance, l'avis n°3 est adopté à l'unanimité.*

**Avis n°4 : Suivre la diffusion des indicateurs de développement durable et de ceux du tableau de bord complétant la mesure du produit intérieur brut (PIB) et des informations utiles à leur compréhension**

**Françoise MAUREL** indique que cet avis rend compte des travaux menés et des avancées autour de cette thématique.

*L'avis n°4 est adopté à l'unanimité.*

**Avis n°5 : Améliorer la connaissance de la dépense territoriale**

**Françoise MAUREL** souligne que l'enjeu serait d'encourager les progrès dans ce domaine. Elle gage que cette orientation devrait également être reprise dans le cadre du prochain moyen terme du Cnis.

**Christine CHOIGNOT, représentant l'UNIOPSS**, fait état d'une véritable difficulté pour les utilisateurs à se saisir de sources multiples et à dialoguer avec les autorités ou collectivités locales compétentes en la matière (ARS, conseils départementaux, etc.). Elle évoque également une problématique liée à l'émergence de données médico-économiques ou socio-économiques alimentant les décisions publiques sans pour autant relever de la statistique publique, avec des enjeux de fiabilité, de qualité et de comparabilité associés.

**Françoise MAUREL** gage que ces préoccupations pourront être prises en compte dans le cadre du prochain moyen terme.

**Xavier NIEL, représentant la DGCL**, suggère de supprimer la référence à une connaissance de la dépense des collectivités locales demeurant insuffisante, compte tenu des améliorations obtenues et des perspectives ouvertes en la matière par la mise à disposition par la DGFIP de bases exhaustives.

**Patrice DURAN** propose de retenir cet amendement.

*Compte tenu des modifications apportées en séance, l'avis n°5 est adopté à l'unanimité.*

**Avis n°6 : Progresser dans la mesure de l'économie numérique**

**Françoise MAUREL** rappelle qu'un colloque sera organisé par le Cnis sur ce thème en mars 2018.

**Nicolas DÉMOULIN, député de l'Hérault**, estime qu'au-delà de la mesure des impacts du numérique sur les activités économiques, il conviendrait d'insister dans cet avis sur la mesure des impacts sociaux et des nouveaux risques induits par le développement de l'économie numérique, vis-à-vis des phénomènes de harcèlement, de la sécurité des données, de la protection des personnes vulnérables, etc.

**Françoise MAUREL** souligne que ces enjeux ont été abordés dans le cadre de la commission « Démographie et questions sociales ». Elle ajoute que ces préoccupations pourront être discutées dans le cadre du colloque programmé en mars 2018.

**Christine CHOIGNOT** évoque la possibilité de reprendre dans les orientations générales la formulation retenue dans l'avis proposé par la commission « Démographie et questions sociales ».

**Gilles POUZIN** suggère de faire référence aux enjeux pour la protection des citoyens des transformations liées au développement de l'économie numérique.

**Nicolas PROKOVAS** confirme qu'il conviendrait de faire référence aux enjeux économiques et sociaux du développement du numérique.

**Pierre-Yves GEOFFARD, Président de la commission « Services publics et services aux publics »**, estime qu'il conviendrait d'évoquer également les opportunités liées au développement du numérique, afin de ne pas véhiculer une vision biaisée du phénomène, uniquement focalisée sur les risques.

**Patrice DURAN** abonde en ce sens.

**Françoise MAUREL** suggère de faire simplement référence aux enjeux économiques et sociaux du développement de l'économie numérique, afin de ne pas modifier en profondeur cette orientation à ce stade du moyen terme – les réflexions sur ce thème ayant vocation à être poursuivies dans le cadre du colloque organisé en mars 2018.

**Jean-Luc TAVERNIER** propose de faire référence aux impacts majeurs du développement actuel du numérique sur les activités économiques et les relations sociales.

**Nicolas DÉMOULIN** évoque un équilibre à trouver entre les opportunités liées au développement du numérique et les enjeux associés en termes de protection des citoyens.

**Jean-Christophe SCIBERRAS, Président de la commission « Emploi, qualification et revenus du travail »** pointe le risque que le terme « relations sociales » renvoie davantage à une dimension institutionnelle qu'aux enjeux en matière de protection des personnes.

**Pierre-Yves GEOFFARD** suggère de faire référence aux impacts du développement actuel du numérique sur l'organisation sociale.

**Benoît ROBIN** estime que le terme « organisation sociale » pourrait soulever un problème d'interprétation.

**Patrice DURAN** propose de retenir la formulation générale suivante : *« Le Conseil considère que le développement actuel du numérique a un impact majeur sur les activités économiques et sociales, lié à la réduction drastique des coûts de mise en relation des agents économiques et qui a permis l'émergence de nouveaux modes d'intermédiation (plateformes marchandes ou non marchandes mais aussi de nouveaux services tels que le financement participatif des projets). Ces évolutions constituent des enjeux économiques et sociaux majeurs ».*

Il souligne que cette thématique pourra ensuite être approfondie dans le cadre de la préparation du prochain moyen terme.

*Compte tenu des modifications apportées en séance, l'avis n°6 est adopté à l'unanimité.*

### **Adoption du projet d'avis des commissions thématiques**

**Patrice DURAN** propose d'adopter par un vote unique l'ensemble des avis proposés par les commissions thématiques.

*Le projet d'avis des commissions thématiques est adopté à l'unanimité.*

## CONCLUSION

**Patrice DURAN** remercie l'ensemble des participants à cette assemblée plénière. En conclusion, il souligne que l'activité du Cnis est appelée à prendre une importance croissante, compte tenu des questions méthodologiques soulevées par le développement du *Big Data*, du rôle décisif du Conseil dans la constitution d'un espace public de communication autour des données statistiques et des enjeux liés à la maîtrise de celles-ci pour l'évaluation des politiques publiques. Il indique avoir insisté sur ces points auprès de l'Autorité de la statistique publique.