



COURRIER

DES STATISTIQUES

N2 - Juin 2019

Rédaction en chef

Jean-Louis Lhéritier

Contribution

Insee : Dominique Bonnans,
Olivier Haag, Magali Demotes-Mainard,
Sylvie Lagarde, Pascal M. Rivière,
Philippe Scherrer

Office fédéral allemand de la
statistique (Destatis) :

Arno Bens, Stefan Schukraft
SSM Sécurité intérieure (SSMSI) :
François Clanché

Directeur de la publication

Jean-Luc Tavernier

Directeur de la collection

Pascal Rivière

Rédaction

Maryse Cadalanu, Pierre Glénat,
Jean-Louis Lhéritier, Odile Rascol,
Pascal M. Rivière

Composition

Agence **Latitude** Nantes

0274/19

02 51 25 06 06

www.agence-latitude.fr

Photo de couverture :

Fotolia®

Éditeur

Institut national de la statistique
et des études économiques
88, avenue Verdier
92541 MONTROUGE CEDEX

www.insee.fr

© Insee 2019 « Reproduction partielle
autorisée sous réserve de la mention
de la source et de l'auteur ».

Courrier des statistiques N2

SOMMAIRE

Présentation du numéro *Jean-Louis Lhéritier*4

• **Dossier : Allemagne**

- 30 ans de collaboration entre Destatis et l'Insee *Sylvie Lagarde*.....8
- Modernisation des registres administratifs en Allemagne : Développements actuels et enjeux pour la statistique publique *Arno Bens, Stefan Schukraft*.....10
- Le nouveau *Mikrozensus* en Allemagne : un dispositif intégré des enquêtes auprès des ménages à l'horizon 2020 *Pascal M. Rivière*21

La mesure de la sécurité et de la satisfaction vis-à-vis des institutions en France : l'impulsion donnée par les objectifs du développement durable des Nations Unies *François Clanché*33

RMÉS, le référentiel de métadonnées statistiques de l'Insee *Dominique Bonnans*.....46

Élire, un projet ambitieux au service du Répertoire électoral unique *Magali Demotes-Mainard*.....58

Statistiques conjoncturelles d'entreprises : des impulsions européennes, des avancées françaises *Philippe Scherrer*72

Le profilage à l'Insee : une identification plus pertinente des acteurs économiques *Olivier Haag*.....86

PRÉSENTATION DU NUMÉRO

Le numéro N2 du Courrier des statistiques illustre la diversité des thèmes que la revue a vocation à éclairer et la dimension internationale des enjeux portés par la statistique publique.

Le dossier consacré à l'Allemagne, introduit par Sylvie Lagarde, s'inscrit dans la tradition de coopération qui existe depuis près de trente ans entre l'Insee et Destatis, l'office fédéral de la statistique de ce pays. L'article d'Arno Bens et Stefan Schukraft, tous deux en poste à Destatis, décrit la démarche entreprise en Allemagne pour moderniser les registres administratifs et les enjeux qui en découlent pour la statistique publique allemande. Pascal M. Rivière présente l'une de ses opérations phare, le *Mikrozensus*, colonne vertébrale d'un dispositif intégré des enquêtes auprès des ménages à l'horizon de 2020.

Dans son article sur la mesure de la sécurité et de la satisfaction des citoyens vis-à-vis des institutions en France, François Clanché montre que les objectifs du développement durable des Nations Unies, et notamment celui intitulé « Paix, justice et institutions efficaces », ont donné une nouvelle impulsion à la statistique publique. Celle-ci a notamment fait évoluer son dispositif d'enquêtes pour appréhender la mesure des discriminations et de la corruption. Cependant, des indicateurs ne sont pas encore couverts, soit pour des raisons techniques, soit parce que les définitions internationales opératoires ne sont pas encore fixées, soit parce que la statistique publique hésite à investir des champs nouveaux, notamment celui de la gouvernance, pas toujours facile à appréhender conceptuellement et à mesurer de manière robuste.

Dominique Bonnans décrit RMÉS, le nouveau référentiel de métadonnées statistiques de l'Insee. Adoptant les standards internationaux, RMÉS assurera à la statistique publique française une grande cohérence de sa documentation, et surtout lui permettra de franchir un cap en gérant les métadonnées comme de véritables données. Il offrira des services tout au long du cycle de vie d'une opération statistique, tant pour alimenter le référentiel que pour accéder à son contenu. Plusieurs applications de l'Insee pourront également mobiliser les métadonnées, par exemple pour spécifier des questionnaires, décrire des contrôles ou des produits de diffusion. RMÉS aura également vocation à offrir un vecteur de coordination au sein du service statistique public.

Magali Demotes-Mainard rappelle le rôle de l'Insee dans le processus électoral depuis 1946, en particulier la tenue du Fichier général des électeurs à partir duquel étaient établies les listes électorales des communes. Les responsabilités de l'Insee ont été renforcées par la loi de 2016 qui a institué un Répertoire électoral unique et lui en a confié la gestion. L'Insee a lancé et mené à bien en trois ans le projet Élire de construction de ce nouveau répertoire. Il a ainsi livré un système intégré, automatisant les échanges d'information entre les différents acteurs, électeurs, communes, postes consulaires et Insee. Le système est fortement sécurisé et gère plus de deux millions de mouvements chaque année. L'initialisation du répertoire a constitué un projet en soi, qui a nécessité de valider une base de 48 millions d'enregistrements. Le Répertoire électoral unique aura été utilisé en vraie grandeur pour la première fois lors des élections européennes du 26 mai 2019.

Philippe Scherrer retrace ensuite la genèse du dispositif européen des statistiques conjoncturelles d'entreprises lors de la création de l'Euro et de l'Union économique et monétaire. Il rappelle que ce dispositif réglementaire a été modifié à plusieurs reprises pour améliorer la pertinence et la couverture des indicateurs et pour mieux répondre aux besoins de l'Union européenne et de la Banque centrale européenne. Le nouveau règlement Fribis

sur l'intégration de l'ensemble des statistiques des entreprises, voté en avril 2019, permettra une meilleure cohérence des statistiques d'entreprises, tant structurelles que conjoncturelles, ainsi qu'avec le Répertoire européen des groupes. Il offrira également à la statistique conjoncturelle une meilleure couverture du commerce et des services. La statistique conjoncturelle française, qui s'appuie sur un dispositif de production éprouvé, s'est adaptée ou a su anticiper ces évolutions européennes.

Enfin, Olivier Haag présente les travaux de profilage des entreprises réalisés par l'Insee pour mieux identifier les acteurs économiques pertinents. Ces opérations sont rendues nécessaires du fait de l'émergence des groupes de sociétés et des nouvelles pratiques comptables qui ont modifié parfois sensiblement l'autonomie des unités légales. L'Insee identifie ainsi les entreprises au sens économique, au sein des groupes les plus complexes et diversifiés. Il collecte ou calcule ensuite des statistiques sur ces nouveaux contours. Recourant à des monographies pour les plus grands groupes (ceux du CAC 40) et à des échanges avec leurs services, l'Insee s'attache à automatiser le processus de profilage pour couvrir à terme les groupes de taille intermédiaire, notamment ceux cotés en bourse.

Jean-Louis Lhéritier
Rédacteur en chef, Insee



Dossier

Allemagne

30 ANS DE COLLABORATION ENTRE DESTATIS ET L'INSEE

SYLVIE LAGARDE

8

MODERNISATION DES REGISTRES ADMINISTRATIFS EN ALLEMAGNE : DÉVELOPPEMENTS ACTUELS ET ENJEUX POUR LA STATISTIQUE PUBLIQUE

ARNO BENS, STEFAN SCHUKRAFT

10

LE NOUVEAU MIKROZENSUS EN ALLEMAGNE : UN DISPOSITIF INTÉGRÉ DES ENQUÊTES AUPRÈS DES MÉNAGES À L'HORIZON 2020

PASCAL M. RIVIÈRE

21

30 ANS DE COLLABORATION ENTRE DESTATIS ET L'INSEE

Sylvie LAGARDE*

L'Insee et son homologue allemand, Destatis, ont noué depuis le début des années quatre-vingt-dix une collaboration étroite, visant à renforcer leurs liens institutionnels dans un cadre marqué par une gouvernance et une organisation de leur système statistique fort différentes. Cette collaboration a été rythmée par des rencontres bilatérales, à l'origine d'équipes travaillant dans le domaine du marché du travail puis d'une partie du comité de direction des deux institutions, tous les deux à trois ans, organisées alternativement à Wiesbaden et Paris. Ces rencontres ont permis de lancer des travaux communs sur des problématiques statistiques bien spécifiques et pour lesquelles, sans cette coopération étroite, les résultats obtenus n'auraient pu être aussi approfondis. Les deux illustrations suivantes sont des exemples emblématiques de cette collaboration. Un infos-rapides annuel, doublé d'une publication simultanée de Destatis, comparant l'évolution des salaires de base en France et en Allemagne, à l'aide des données produites par les deux organismes, a été publié régulièrement jusqu'en 2015. Cette publication récurrente avait fait suite à une étude de la structure des salaires comparée en France et en Allemagne, publiée dans un article d'Économie et statistique de 1998 par deux co-auteurs de l'Insee et de Destatis. Plus récemment, un Insee Analyses de juin 2016 intitulé « la fragile comparabilité des durées du travail en France et en Allemagne » a résulté d'un partenariat entre les deux institutions sur la base d'une exploitation de l'enquête force de travail (enquête emploi) en France et en Allemagne. Ces exemples ont montré qu'un travail mené en partenariat par les responsables des enquêtes mobilisées (enquête structure des salaires et enquête emploi) dans chacun des deux instituts nationaux de statistique permettait de produire une information statistique de qualité sur les deux pays, intégrant toutes les précautions méthodologiques nécessaires à l'analyse de la comparaison des salaires et des durées du travail. Même dans le cas d'enquêtes fortement harmonisées au niveau européen, la comparaison des statistiques produites par les différents pays ne va en effet pas de soi.

PARTAGER NOS EXPÉRIENCES ET S'ENRICHIR DE NOS INNOVATIONS RESPECTIVES

Les enjeux de cette coopération bilatérale sont multiples : réaliser des comparaisons entre la France et l'Allemagne solides, fort attendues dans le débat public, ce qui nécessite de bien comprendre le mode de recueil et de traitement de l'information statistique de chacun ; partager nos expériences et s'enrichir de nos innovations respectives ; mais aussi faire alliance pour peser davantage dans le cadre de négociations européennes en cours ou à venir. La pleine réalisation de ces enjeux et tout particulièrement le dernier, requiert de bien appréhender comment s'organise et s'articule le travail de la statistique en Allemagne, dans son environnement administratif, culturel et politique. Pour nous, statisticiens, ce n'est souvent pas la première dimension à laquelle nous pensons dans notre activité, plutôt

* Directrice de la Méthodologie et de la Coordination Statistique et Internationale (DMCSI), Insee,
sylvie.lagarde@insee.fr

technique et pour autant, elle est fondamentale. La dernière rencontre bilatérale, qui s'est déroulée à Wiesbaden en septembre 2018 illustre d'ailleurs, dans la diversité des thèmes d'échanges, les principaux enjeux de notre coopération bilatérale. Figuraient notamment à l'ordre du jour : la mondialisation et son impact sur les comptes nationaux et les statistiques d'entreprises, le profilage des entreprises, le développement du big data et des méthodes de data science, la conduite de projets en mode agile, les projets d'évolution du recensement de la population ainsi que la gouvernance et l'organisation du système statistique des deux pays. Sur le point plus spécifique relatif à l'évolution du recensement de la population, les échanges entamés lors de la visite de l'Insee en septembre 2018 à Wiesbaden ont été poursuivis dans le cadre de la préparation du futur règlement cadre européen sur la population, au-delà du prochain recensement 2021. L'objectif est de finaliser la rédaction d'un document formalisant une position commune France-Allemagne sur les propositions d'Eurostat relatives au projet de règlement statistique cadre associé.

UN DOSSIER TÉMOIN D'UNE TRADITION ANCIENNE DE COOPÉRATION BILATÉRALE

In fine, le dossier sur l'Allemagne intégré dans ce numéro et dont l'origine remonte à la visite de septembre 2018 à Wiesbaden, est un output visible de la coopération bilatérale de longue date menée avec Destatis. Il renoue avec une tradition ancienne, puisque les numéros de décembre 1997 et décembre 1999 du Courrier des statistiques contenaient déjà des articles ou dossiers sur l'Allemagne. L'article d'Arno Bens et Stefan Schukraft (Destatis) explique la façon dont Destatis s'est lancé dans le recensement des registres et données administratives potentiellement intéressants pour la statistique publique, en profitant de la dynamique enclenchée par la demande du Conseil national de surveillance des normes pour accroître l'efficacité de l'administration grâce à la mise en place d'un ensemble de registres modernisés et interconnectés. Dans l'article suivant, Pascal M. Rivière (Insee) détaille comment, dans un contexte de négociation d'un règlement cadre européen sur les statistiques sociales, Destatis s'apprête à déployer un système intégré d'enquêtes auprès des ménages, s'appuyant sur le dispositif historique du *Mikrozensus* tout en introduisant la possibilité pour les enquêtés de répondre par internet.

MODERNISATION DES REGISTRES ADMINISTRATIFS EN ALLEMAGNE**

DÉVELOPPEMENTS ACTUELS ET ENJEUX POUR LA STATISTIQUE PUBLIQUE

Arno Bens*, Stefan Schukraft*

L'utilisation optimale des données disponibles est la condition d'administrations performantes et tournées vers l'avenir. Le Conseil national de surveillance des normes (Nationaler Normenkontrollrat) a ainsi mis à l'ordre du jour politique l'idée maîtresse de registres modernisés et interconnectés et a publié un rapport d'experts à ce sujet. En conséquence de cela et sous réserve de l'accord des personnes ou entreprises concernées, les autorités pourraient avoir accès par voie électronique à de grands volumes de données dont elles ne disposaient pas jusqu'à présent ou qu'elles devaient recueillir par ailleurs. Cela permettrait d'améliorer les services rendus par les administrations et d'alléger la charge des usagers. L'Office fédéral de la statistique (Destatis) s'est saisi du sujet et a identifié dans une procédure systématique les utilisations potentielles des données et registres administratifs existants. Une plateforme d'information collaborative relative aux sources de données administratives a été créée pour rassembler l'information existante sur les registres.

 *Optimum use of existing data is a key to effective and future-oriented public administration. Thus, Germany's National Regulatory Control Council (Nationaler Normenkontrollrat) has put the vision of modernised, digitally connected registers on the political agenda and has published an expert opinion. With the consent of persons and companies concerned, public authorities could have electronic access to many data that so far have not been available to them or have had to be collected specifically. This would lead to better and low-burden administrative services. The Federal Statistical Office of Germany (Destatis) now applies this vision, identifying through a systematic approach the potential of using existing administrative registers and data. A common platform on administrative data sources is created to gather knowledge on existing registers.*

La numérisation offre à l'administration d'énormes possibilités pour améliorer son efficacité et pour développer de nouvelles offres de services aux usagers. Sans être une fin en soi, elle peut soulager citoyennes et citoyens comme entreprises d'une bureaucratie inutile. La numérisation constitue aussi un aspect important de la stratégie future

*Arno Bens, chef de projet au bureau « Évaluation de la charge, estimation des coûts » de Destatis.

arno.bens@destatis.de

*Stefan Schukraft, Expert référent au bureau « Questions générales sur la mesure des coûts de la bureaucratie » de Destatis

stefan.schukraft@destatis.de

**Traduction d'une partie de l'article « Registermodernisierung und Verwaltungsdatennutzung in der amtlichen Statistik » paru dans WISTA - Wirtschaft und Statistik, Vol. 4/2018. Disponible dans sa version originale à l'adresse :

<https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2018/04/registermodernisierung-042018.html>

de l'administration allemande dans le contexte du changement démographique. Si ce sujet d'actualité suscite un certain enthousiasme, les objectifs liés à la protection des données, leur qualité et leur traitement ne doivent cependant pas être relégués au second plan. C'est dans ce contexte compliqué qu'ont émergé des propositions en vue de moderniser les registres, de les suivre et de les intégrer dans les processus de la statistique publique.

📍 LA VISION: DES REGISTRES MODERNES ET INTERCONNECTÉS POUR L'ADMINISTRATION NUMÉRIQUE

« Meilleure qualité de service pour les citoyennes et citoyens et les entreprises : numériser l'administration. Moderniser les registres », c'est le titre d'un rapport d'experts publié en octobre 2017 (McKinsey, 2017) à la demande du Conseil national de surveillance des normes (*Nationaler Normenkontrollrat*). Celui-ci, en tant que comité indépendant, conseille le Gouvernement fédéral sur les questions de réduction de la bureaucratie et d'amélioration de la réglementation. Le rapport présente des orientations permettant d'accroître l'efficacité et la performance de l'administration grâce à un ensemble de registres modernisés et interconnectés. Il s'intéresse au premier chef à la satisfaction des usagers, c'est-à-dire à celle des citoyens, mais aussi des entreprises ayant recours aux services de l'administration. Ce document s'intéresse particulièrement à la statistique publique, en tant que partie de l'administration qui recueille et traite une grande quantité de données.

Les idées centrales de ce rapport d'experts, auquel l'Office fédéral de la statistique, Destatis, a également apporté son concours, sont esquissées ci-après.

📍 LES DONNÉES MISES EN RÉSEAU, RESSOURCE MAJEURE POUR DE MEILLEURS SERVICES PUBLICS

Aujourd'hui, l'administration s'appuie dans une très large mesure sur les technologies de l'information. Pour cela, des données sont recueillies, traitées et vérifiées à partir de formulaires, de justificatifs ou d'interrogations de bases de données par exemple ; les résultats des procédures administratives sont en général produits et conservés de manière électronique. Les registres tenus par l'administration jouent un rôle particulier : ceux-ci contiennent en effet des données structurées relatives à des faits précis qui sont utilisées dans de nombreuses procédures, tant comme éléments de base pour des vérifications que pour la conservation à long terme de décisions et de certificats administratifs.

Les données constituent donc une ressource en soi pour la gestion des administrations et ouvrent ainsi une voie importante pour améliorer l'offre de services, dans la mesure où l'on parviendrait à en optimiser l'accès et le traitement. Cette idée fondamentale est déjà perceptible dans la loi promulguée en 2013 qui visait à promouvoir l'administration électronique (*E-government-Gesetz – EGovG*) : celle-ci avait créé les conditions pour que les autorités fédérales puissent, avec l'accord des personnes concernées, demander des justificatifs par voie électronique directement auprès de l'administration les ayant établis elle-même de cette façon¹. En outre, dans le cadre de la numérisation, les autorités devaient documenter, analyser et optimiser leurs procédures administratives².

1. EGovG, § 5.

2. EGovG, § 9.

Mettre en œuvre cette idée directrice, centrée sur les usagers, génère de nouvelles exigences et des potentiels d'amélioration pour l'administration : dans un monde numérisé, les utilisateurs et utilisatrices sont habitués à ce qu'on puisse s'abstraire des limites habituelles, physiques ou autres, et que les applications informatiques offrent une approche intégrée. Les exemples sont notamment les sites de vente sur internet qui proposent à l'achat les articles les plus divers indépendamment de leur fournisseur respectif, ou bien les plateformes de voyages qui combinent vols, hôtels et services d'assurance. Ces types d'expérience diminuent la prédisposition à être en prise directe avec différents organismes compétents, à divers niveaux administratifs et dans divers cadres juridiques lors d'un recours aux services administratifs numérisés. La loi relative à l'accès en ligne (*Onlinezugangsgesetz*) du 14 août 2017 répond à ces nouvelles attentes des usagers par la création d'un regroupement de portails pour les services en ligne des administrations.

Mais comment y parvenir ? « Une grande remise en ordre s'impose dans la salle des machines de l'administration numérique », comme le formule le président du Conseil national de surveillance des normes, le Dr. Johannes Ludewig, dans son préambule au rapport

“ « Une grande remise en ordre s'impose dans la salle des machines de l'administration numérique », comme le formule le président du Conseil national de surveillance des normes. ”

d'experts³. L'organisation largement éclatée de l'administration a pour conséquence que les citoyens et les entreprises doivent « aller d'une administration à l'autre » dans de nombreuses situations de la vie. Il faut fournir des informations et des justificatifs à différentes autorités pour obtenir un service. Et l'on est souvent confronté à la situation rageante dans laquelle des informations identiques ou similaires doivent être fournies plusieurs

fois. De tels défauts systémiques sont récurrents dans les relations administrations-usagers (citoyens ou entreprises), mais aussi lors des interactions entre administrations.

Prenons par exemple la demande d'allocations parentales : sont exigés, entre autres justificatifs, des données sur la personne faisant la démarche (par exemple : adresse, numéro d'identification fiscale, nationalité, état civil) ainsi qu'un avis d'imposition et un certificat de naissance de l'enfant. Or toutes ces informations, qui sont en général déjà connues de l'administration, doivent être encore rassemblées et transmises par le demandeur, et souvent sur support papier. À la caisse d'allocations parentales, ces informations sont de nouveau saisies dans une application spécifique et les documents sont le cas échéant numérisés si une gestion des dossiers électronique est en place.

VERS UN ENVIRONNEMENT DE REGISTRES MODERNISÉ

L'idée de la modernisation des registres doit permettre de répondre à ces nouveaux enjeux sur deux plans :

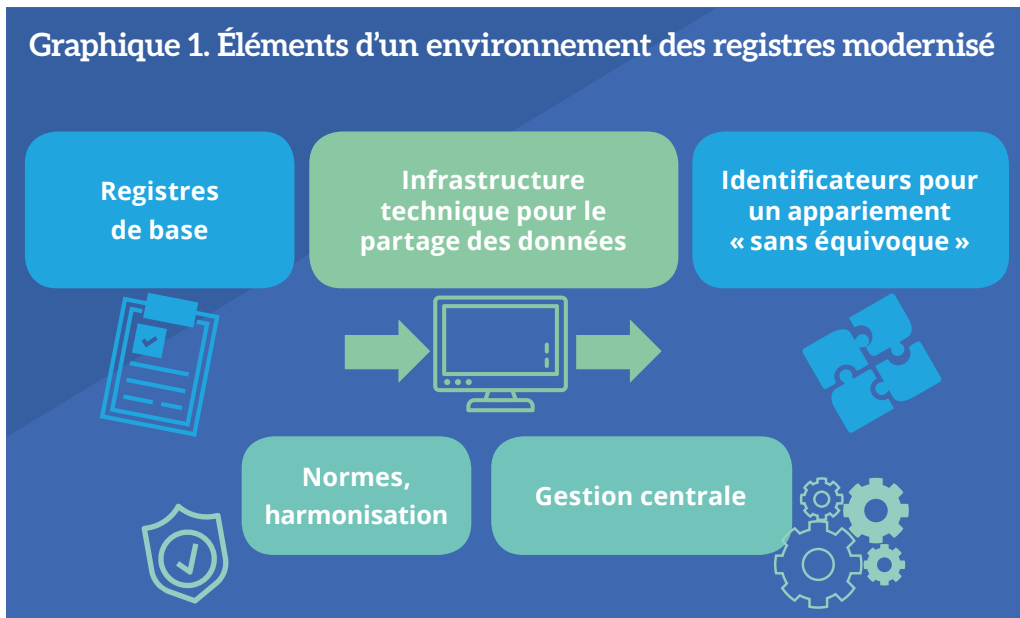
- ① **Fourniture des informations** : les informations déjà transmises à l'administration ne doivent pas être demandées une nouvelle fois (c'est ce que l'on appelle le principe « *Once-Only* »).

3. Rapport d'experts (McKinsey, 2017), cité plus haut.

❶ **Gestion des justificatifs** : la nécessité de présenter des justificatifs tombe lorsqu'ils peuvent être fournis en interrogeant par voie électronique des registres ou des bases de données administratives.

Dans un environnement de registres modernisé, des parties du formulaire de demande pourraient être déjà pré-remplies. De plus, la présentation des justificatifs deviendrait inutile lorsque, par exemple, la caisse d'allocations parentales qui examine le dossier peut obtenir les renseignements nécessaires par voie électronique, à partir des registres des déclarations domiciliaires et des registres d'État civil, et auprès du Trésor public. L'accord de la personne faisant la demande serait naturellement sollicité.

Même des statistiques comme le recensement de population pourraient être établies plus facilement, de manière plus économique et plus rapide en se fondant sur des registres. Les éléments clé d'un environnement de registres modernisé, selon le rapport d'experts cité précédemment, peuvent être décrits comme dans le **graphique 1** qui suit.



❶ DIFFÉRENTS REGISTRES DE BASE COMPORTANT LES RENSEIGNEMENTS LES PLUS UTILISÉS

Les informations importantes sur des unités pertinentes, comme les personnes et les entreprises, doivent être disponibles dans des registres dits « de base ». En particulier, des registres de base relatifs aux adresses, aux bâtiments, aux véhicules et aux installations ou institutions spécifiques sont envisageables et opportuns. Ils fournissent des informations essentielles, nécessaires dans de nombreuses procédures administratives et ce de manière actualisée, complète et avec un bon niveau de qualité. Il convient que le cadre juridique garantisse un accès adapté à ces données. En outre, toute la clarté et la transparence sont requises pour l'utilisation, la gestion et la gouvernance des données.

Il existe déjà de nombreux répertoires en Allemagne qui pourraient servir de base pour élaborer de tels registres. Les registres des déclarations de domicile et les registres d'État civil, le registre central des étrangers, le registre du commerce et le registre central

des véhicules, sont quelques exemples de quasi-registres de base qui permettent d'ores et déjà un partage des données.

Dans d'autres domaines, on trouve au contraire très peu de bases de données ou de fichiers disponibles grâce auxquels on pourrait développer des registres de base. C'est le cas pour des sujets tels que l'immobilier, l'éducation ou les travailleurs indépendants.

L'IMPORTANCE DE L'INFRASTRUCTURE TECHNIQUE

Pour un partage des données efficace entre les administrations et les registres de base, une infrastructure technique appropriée est nécessaire. Des normes homogènes doivent garantir une mise en réseau, modulable et indépendante des plateformes, des données des registres ; elles doivent correspondre au meilleur niveau de la technique pour ce qui concerne la protection et la sécurité des données. Une centralisation des fichiers n'est cependant pas obligatoire, mais peut avoir des avantages en termes de cohérence, de performance et de coûts. Outre les registres de base, des fichiers administratifs de domaines spécifiques doivent être mis en réseau par le biais d'interfaces *ad hoc*. Cela peut être préparé par des projets pilotes plus particulièrement appropriés pour fournir aux agents des services gestionnaires une approche vivante de l'utilisation des registres interconnectés. On pourrait donner comme exemple une demande d'allocations parentales électronique ou un recensement fondé sur les registres.

DES IDENTIFIANTS UNIQUES POUR FACILITER LES APPARIEMENTS

Pour travailler efficacement, une autorité doit pouvoir identifier avec précision dans un répertoire donné les enregistrements dont elle a besoin. Ils doivent en particulier être repérables sans ambiguïté. Dans les bases de données qui existent actuellement, cela se fait la plupart du temps par le biais de caractéristiques « permanentes ». Par exemple, pour les personnes : nom, adresse, date de naissance ; pour les entreprises : dénomination de la société ou de l'entreprise ainsi que de son siège. Cette procédure présente deux inconvénients : d'une part, des orthographes différentes ou des fautes de frappe peuvent nécessiter une correction ou une sélection manuelle ; d'autre part, de nombreuses informations doivent être fournies à des fins de vérification. Enfin, certaines caractéristiques permanentes peuvent changer, par exemple du fait d'un déménagement, d'un mariage ou d'un changement de raison sociale.

Les autorités qui gèrent de grandes quantités de données emploient ainsi des identifiants en vue d'un appariement efficace. C'est l'usage pour le numéro d'identification fiscale ou le numéro de sécurité sociale dans le cas des personnes, pour le numéro de registre du commerce⁴, le numéro d'identification⁵, le numéro EORI⁶ ou le numéro LEI⁷ pour les entreprises. Pour identifier les procédures, des numéros de dossier spécifiques aux administrations sont également donnés.

Comme le montre cette liste, tous ces identifiants n'ont qu'une portée restreinte, ce qui veut dire qu'ils ne peuvent être employés à des fins d'identification que dans un contexte thématique donné ou pour une administration précise. Si des données provenant d'une

4. Le numéro du registre du commerce est unique, en lien avec le tribunal d'immatriculation compétent.

5. Le numéro d'identification de l'entreprise est donné par l'Agence fédérale pour l'emploi.

6. EORI = *Economic Operators Registration and Identification* ; identifiant unique pour les acteurs économiques pour le dédouanement au sein de l'Union européenne.

7. LEI = *Legal Entity Identifier* ; identifiant unique global de l'entité légale, en particulier pour les entreprises présentes sur les marchés financiers.

autre sphère administrative doivent être traitées, il faut que l'identification de l'unité concernée se fasse soit par les données permanentes, soit *via* l'identifiant spécifique, demandé séparément dans la procédure administrative.

Un numéro d'identification unique pour les personnes, les entreprises et d'autres

“ *Un numéro d'identification unique pour les personnes, les entreprises et d'autres entités concernées procure d'énormes avantages sur le plan technique.* ”

entités concernées procure d'énormes avantages sur le plan technique : les interrogations dans les registres et les transmissions d'information aux administrations peuvent être toujours effectuées sans équivoque de manière automatisée. En outre, les

liens entre les unités peuvent être établis efficacement par voie électronique : si pour un véhicule, le numéro d'identification du propriétaire est enregistré, les changements d'adresse et de nom peuvent être transmis de manière automatisée aux autorités compétentes et être importés dans leurs fichiers.

🔍 NORMALISATION ET HARMONISATION, PRÉALABLES À LA QUALITÉ

Outre les importantes exigences techniques allant de pair avec la mise en œuvre de cette idée de modernisation et d'interconnexion des registres, des mesures supplémentaires devraient être également prises.

L'emploi d'informations venant de registres interconnectés dans la procédure administrative nécessite en particulier des normes de qualité uniformes. Pour de nombreux registres existants, la cohérence du contenu est assurée par la réglementation juridique. En cas de gestion des données décentralisée, il faut également garantir le respect de normes de qualité, concrètes et formelles, uniformes en matière d'exactitude, d'exhaustivité et d'actualité. Si par exemple, une administration, pour exécuter une procédure de planification ou d'autorisation, recourt à des données recueillies ou générées par une autre administration, la bonne conformité des données doit être garantie. Il faut déterminer pour cela les règles d'enregistrement dans la base de données, tant sur le plan pratique que sur celui du contenu, et produire une qualité des données uniforme.

Afin que la mise en œuvre du principe « *Once-Only* » puisse conduire à des allègements notables dans un environnement de registres modernisé, il faut également harmoniser les contenus des registres. Si une variable ou une unité est définie de manière différente d'une sphère administrative à l'autre, chaque procédure administrative doit nécessairement tenir compte des définitions propres à celle-ci. On peut citer comme exemples les grandeurs monétaires comme les revenus et les chiffres d'affaires. Dans de nombreuses procédures administratives, de tels renseignements sont nécessaires, mais il faut aussi souvent ajouter ou déduire des postes précis dans le calcul de la valeur à reporter. Dans ces cas, un recours, le cas échéant, aux données de registres existantes ne sera un progrès que si les données nécessaires aux calculs sont également bien différenciées dans les autres systèmes (par exemple, dans la procédure d'imposition).

La même remarque s'applique lorsque les unités de base d'un registre sont différentes. C'est souvent le cas pour les entreprises : on peut avoir besoin de renseignements relatifs à des unités définies différemment du point de vue juridique ou physique (par exemple groupes, entités juridiques, filiales, établissements et lieux de travail). Des registres performants et des données de qualité ne peuvent alors être obtenus que par l'action

Identifications des personnes spécifiques à chaque domaine : le modèle autrichien en matière de protection légale des données

Les risques liés à la protection des données augmentent à mesure que l'introduction de nouveaux identifiants élargit les utilisations possibles. La conservation et le transfert d'informations au sein d'un ensemble de registres interconnectés présentent toujours un risque résiduel. À elle seule, l'existence d'identifiants personnels uniques permettrait facilement à des pirates informatiques de combiner les différentes données relatives à une personne et de reconstituer ainsi des profils détaillés. De plus, le fait que les citoyens soient conscients que les autorités peuvent demander et combiner des données personnelles et sensibles pourrait conduire à une perte de confiance. Même une réglementation légale *ad hoc*, qui limiterait les possibilités des autorités, ne pourrait, lorsqu'un tel abus est techniquement possible, éliminer les doutes et craintes, partiellement fondés sur le plan historique (Martini *et al.*, 2017).

Une solution, alliant une identification sans équivoque des personnes à une forte garantie de la protection de leurs données, a été mise en œuvre en Autriche : le système y combine un identifiant personnel unique secret (« numéro permanent ») à différents identifiants personnels spécifiques à des domaines. 26 domaines d'activité publics ont été définis par le décret de délimitation des domaines du e-gouvernement autrichien (*E-government-Bereichsabgrenzungsverordnung*). Les utilisations de données d'un domaine d'activité relient des données personnelles uniquement à l'identifiant personnel spécifique de ce domaine. Cela empêche que des données d'une personne relevant d'un domaine administratif soient rapprochées de données de la même personne dans un autre domaine. C'est la situation actuelle en Allemagne : les données d'un citoyen sont enregistrées avec le numéro d'identification fiscal au sein du Trésor public, avec le numéro de sécurité sociale au sein de la sécurité sociale, etc. Un regroupement ou une comparaison des données personnelles entre domaines n'est alors possible que par le nom (et d'autres renseignements le cas échéant) ou lorsque les numéros d'identification respectifs des différents domaines sont connus.

L'avantage du modèle autrichien réside dans le fait que tous les identifiants personnels spécifiques à un domaine peuvent être déduits au moyen d'une opération mathématique irréversible à partir d'un « numéro de base » secret, permettant d'identifier un citoyen de manière unique tout au long de sa vie. Ce « numéro de base », produit uniquement par l'autorité autrichienne compétente, n'est enregistré que sur la carte de citoyen* de la personne concernée. Une utilisation des données entre domaines n'est ainsi possible que si le propriétaire de la carte de citoyen en a donné l'autorisation. En outre, lors de procédures administratives électroniques, le « numéro de base » peut être lu pour en inférer le nécessaire identifiant personnel, spécifique au domaine sécurisé. Dans une variante, l'autorité du registre des « numéros de base » peut fournir un jeu d'identifiants relatifs à une personne, chacun étant spécifique à un domaine. Pour cela, l'autorité demandeuse doit identifier la personne dans le registre des déclarations domiciliaires et être autorisée explicitement à consulter les identifiants personnels spécifiques à chacune des domaines. Les fonctions liées au registre des « numéros de base » sont assurées par l'autorité indépendante de protection des données autrichienne. Un système comparable pourrait être mis en place dans des conditions similaires en Allemagne, dans le respect de la protection des données et en conformité avec la Constitution.

* La carte de citoyen autrichienne est comparable à la nouvelle carte d'identité allemande avec fonction d'identification électronique. La carte de citoyen dispose aussi d'une fonction de signature.

conjointe des différents services et administrations concernés, en exploitant les moyens techniques appropriés.

📍 L'IMPORTANCE D'UN PILOTAGE CENTRAL

La mise en œuvre des dits éléments d'un environnement de registres modernisé nécessite une gestion coordonnée supra-fédérale. La création de normes techniques homogènes et les progrès dans l'harmonisation des registres exigent une collaboration englobant tous les départements ministériels. Ceci ne peut être garanti que par une entité centrale prenant en charge les travaux de coordination nécessaires. Celle-ci doit aussi pouvoir porter des recommandations et intervenir dans la sphère politique.

Sa tâche la plus importante est de développer une infrastructure constituée de registres de base, de procédures de partage des données et d'identifiants. Pour cela, il faut impliquer étroitement les acteurs – « métier » et « technique » – chargés de la

mise en œuvre opérationnelle, et réaliser les normalisations nécessaires.

« Il convient que le service centralisateur fasse sentir le bénéfice de la modernisation des registres par la mise en place de projets pilotes concrets. »

De surcroît, il convient que le service centralisateur fasse sentir le bénéfice de la modernisation des registres par la mise en place de projets pilotes concrets. Pour faire connaître la thématique au sein de l'administration et du public, une communication large et ciblée est nécessaire. Il faut également expliciter les solutions pratiques pour améliorer la protection des données et la transparence. Outre son rôle

de force motrice, le service centralisateur doit aussi jouer un rôle d'information et de conseil. Il doit informer les autorités intéressées sur les potentiels et les possibilités de la modernisation des registres, et soutenir les projets correspondants.

📍 LA MODERNISATION DES REGISTRES, UN SUJET À L'AGENDA POLITIQUE...

Dans leur accord de coalition, les partis au gouvernement de l'État fédéral soulignent l'importance de la numérisation. Le rapport d'experts est clair à cet égard : « Pour cela, nous autorisons les administrations à appairer des données de plusieurs registres grâce à des identifiants communs sans équivoque (principe « Once-Only »). Nous allons moderniser les registres publics et examiner pour cela les propositions du Conseil national de surveillance des normes ». Le nouveau gouvernement veut mettre en place une agence d'e-gouvernement pour développer des normes et des solutions pilotes supra-fédérales.

Le Conseil national de surveillance des normes continue de faire avancer cette question dans la sphère politique et demande sa mise en œuvre dans le cadre d'une loi sur la modernisation des registres. Le Conseil statistique de Destatis souligne également l'importance de cette modernisation, en vue d'utiliser les données administratives dans la statistique publique : c'est en effet important en termes de qualité des données, d'allègement de la charge des unités interrogées (surtout les entreprises) et de réduction des coûts⁸.

8. Lors de sa 64^e séance, le Conseil de la statistique a émis des recommandations en ce sens. La modernisation des registres reprend en outre des recommandations du Conseil pour le développement de la statistique officielle.

❶ ... DONT SE SAISIT LA STATISTIQUE PUBLIQUE

Avec le rapport d'experts du Conseil national de surveillance des normes, une nouvelle dynamique s'est développée, que l'Office fédéral de la statistique met à profit : celui-ci a mis en place une procédure systématique pour identifier des sources de données administratives appropriées et fait progresser leur utilisation. La numérisation offre à cet effet de nouvelles opportunités et possibilités d'applications pour la statistique. Elle apporte aussi un allègement de la charge des répondants, notamment par la mise en application du principe « *Once-Only* ».

Actuellement en Allemagne, près de la moitié des statistiques sont recueillies d'une manière indirecte, c'est-à-dire que les citoyens et entreprises ne sont pas interrogés directement⁹. Au lieu de cela, les informations nécessaires sont recueillies auprès des services compétents, principalement dans l'administration.

« Actuellement en Allemagne, près de la moitié des statistiques sont recueillies d'une manière indirecte. »

Afin que les registres et les données administratives puissent être utilisés à des fins statistiques, ils doivent être conformes aux exigences correspondantes en termes de qualité, d'actualité et de champ.

Pour la statistique publique, il est donc fondamental de classer les sources de données envisageables selon leur degré d'adéquation à leur utilisation en statistique.

❷ UNE REVUE GÉNÉRALE DES SOURCES ADMINISTRATIVES EXISTANTES

À cet effet, l'Office fédéral de la statistique a expérimenté la mise en place d'une procédure systématique d'identification et d'examen des registres et données administratives. Pour cela, il s'est appuyé sur les nouveaux instruments juridiques créés par l'amendement de la loi fédérale sur la statistique et l'introduction du § 5a (Engelter et Sommer, 2016), qui autorise notamment l'Office fédéral de la statistique à demander aux administrations concernées des métadonnées et des données de test. Cette procédure comporterait cinq étapes :

- ❶ **Étape 1** : définition des domaines et des conditions d'utilisation des sources administratives ;
- ❷ **Étape 2** : cartographie des registres existants ;
- ❸ **Étape 3** : identification des sources de données potentielles ;
- ❹ **Étapes 4a et 4b** : demande de métadonnées et de données de test ;
- ❺ **Étape 5** : utilisation effective des données administratives.

L'Office fédéral de la statistique a ainsi rassemblé de l'information sur les 214 registres administratifs cités dans le rapport d'experts précité. Ces données alimentent la plateforme d'informations de données administratives qu'il a créée.

Dans les domaines de l'industrie, de la construction et de l'énergie, il a mené à bien les étapes 1 à 4a (métadonnées) et va étendre progressivement ces travaux à d'autres secteurs.

9. Voir (Lorenz et Opfermann, 2017), page 50.

DÉJÀ QUELQUES ENSEIGNEMENTS

Les premières expériences montrent clairement que des difficultés restent encore à surmonter :

- ❶ Sans une harmonisation des définitions et des concepts entre les variables traitées au sein de l'administration et celles nécessaires à la statistique, la mise en œuvre d'un « *Once-Only* » est fortement limitée.
- ❶ La qualité dans le temps des données administratives doit être garantie si celles-ci doivent avoir une utilisation statistique.
- ❶ L'infrastructure technique hétérogène au sein de l'administration nécessite des interfaces spécifiques pour la statistique.
- ❶ Lors de l'examen de l'adéquation et de l'exploitation de nouvelles sources de données administratives, les questions de droit doivent encore et toujours être clarifiées.
- ❶ La transparence nécessaire sur l'emploi des données, la protection des données et le droit de regard des usagers sur leurs données personnelles doivent être garantis.
- ❶ L'enrichissement des données individuelles déjà disponibles au travers d'autres sources de données administratives est très lourd si aucun identificateur unique adapté n'est disponible dans les deux sources.
- ❶ L'éclatement administratif et le manque fréquent de coordination entre les services métiers et les services informatiques, entre les différentes strates fédérales ou entre les entités compétentes alourdissent fortement la communication.

Le rapport d'experts propose des solutions pour l'administration dans son ensemble. Leur mise en œuvre offrirait également à la statistique publique de nouvelles marges de manœuvre et créerait une nouvelle dynamique pour l'utilisation des données administratives.

BIBLIOGRAPHIE

ENGELTER, Marion et SOMMER, Kay, 2016. Die Novellierung des Bundesstatistikgesetzes 2016. In : *WISTA – Wirtschaft und Statistik*. Wiesbaden, Destatis, 15 décembre 2016. Vol. 6/2016, pp. 11-19. ISBN 978-3-8246-1048-8.

KOALITIONSVERTRAG ZWISCHEN CDU, CSU UND SPD 19. LEGISLATURPERIODE, 2018. *Ein neuer Aufbruch für Europa. Eine neue Dynamik für Deutschland. Ein neuer Zusammenhalt für unser Land* [en ligne]. [Consulté le 22 mai 2019]. Disponible à l'adresse : <https://www.bundesregierung.de/resource/blob/975226/847984/5b8bc23590d4cb2892b-31c987ad672b7/2018-03-14-koalitionsvertrag-data.pdf?download=1>

LORENZ, Robin et OPFERMANN, Rainer, 2017. Verwaltungsdaten in der Unternehmensstatistik. In : *WISTA – Wirtschaft und Statistik*. Wiesbaden, Destatis, 15 février 2017. Vol. 1/2017, pp. 49-66. ISBN 978-3-8246-1060-0.

MARTINI, Mario, WAGNER, David et WENZEL, Michael, 2017. *Rechtliche Grenzen einer Personen – bzw. Unternehmenskennziffer in staatlichen Registern* [en ligne]. Speyer, Université allemande des Sciences Administratives de Speyer, 17 septembre 2017. [Consulté le 22 mai 2019]. Disponible à l'adresse : <http://www.normenkontrollrat.bund.de/resource/blob/72494/476034/eebab686008cfe-c0a7919ca03e51abe3f/2017-10-06-download-nkr-gutachten-2017-anlage-untersuchung-datenschutz-data.pdf>

MCKINSEY & COMPANY, 2017. *Mehr Leistung für Bürger und Unternehmen : Verwaltung digitalisieren. Register modernisieren* [en ligne]. [Consulté le 22 mai 2019]. Disponible à l'adresse : <https://www.normenkontrollrat.bund.de/resource/blob/72494/476004/12c91ffb877685f-4771f34b9a5e08fd/2017-10-06-download-nkr-gutachten-2017-data.pdf>

STATISTISCHER BEIRAT, 2012. *Fortentwicklung der amtlichen Statistik – Empfehlung des Statistischen Beirats für die Jahre 2018 bis 2022* [en ligne]. Novembre 2012, p. 5 [Consulté le 22 mai 2019]. Disponible à l'adresse : https://www.destatis.de/DE/Ueber-uns/Leitung-Organisation/Statistischer-Beirat/fortentwicklung-nov-2018-2022-teil3.pdf?__blob=publicationFile&v=4

STATISTISCHES BUNDESAMT, 2017. *Ein Blick in die Registerlandschaft in Deutschland* [en ligne]. Wiesbaden [Consulté le 22 mai 2019]. Disponible à l'adresse : <https://www.normenkontrollrat.bund.de/resource/blob/300864/476024/04a6019c-945895d3587136ff2ce46b73/2017-10-06-download-nkr-gutachten-2017-anlage-untersuchung-staba-register-data.pdf?download=1>

FONDEMENTS JURIDIQUES

Loi relative à la statistique pour l'État fédéral (*Bundesstatistikgesetz – BStatG*) dans la version publiée le 20 octobre 2016 (BGBl. I p. 2394), modifiée par l'Article 10 Paragraphe 5 de la loi du 30 octobre 2017 (BGBl. I p. 3618).

Loi relative à la promotion de l'administration électronique (*E-government-Gesetz – EGovG*) du 25 juillet 2013 (BGBl. I p. 2749), modifiée par l'article 1 de la loi du 5 juillet 2017 (BGBl. I p. 2206).

Loi relative à l'amélioration de l'accès en ligne aux services administratifs (*Online – Zugangsgesetz – OZG*) du 14 août 2017 (BGBl. I pp. 3122, 3138).

LE NOUVEAU MIKROZENSUS EN ALLEMAGNE

UN DISPOSITIF INTÉGRÉ DES ENQUÊTES AUPRÈS DES MÉNAGES À L'HORIZON 2020

Pascal M. Rivière*

Mené depuis plus de 60 ans, le Mikrozensus allemand est une enquête sur l'ensemble de la population, au taux de sondage de 1 %. Il s'agit de la plus grande enquête auprès des ménages de l'Union européenne. Conformément à l'organisation de son système statistique, l'office fédéral de statistique allemand définit sa méthodologie, qui n'est pas sans rappeler celle de l'enquête emploi française. Et ce sont les offices statistiques des Länder qui assurent la collecte des données. Privilégiant depuis le début les questions de composition du ménage, de logement, d'activités, de revenus et de formation/qualification, les thématiques du questionnaire ont changé au fil du temps. La diversité des domaines traités permet d'effectuer des analyses variées et fines des évolutions structurelles de la société allemande. À l'horizon de 2020-21, le nouveau Mikrozensus va permettre une plus grande rationalisation du dispositif allemand des enquêtes statistiques auprès des ménages, pour plusieurs raisons : intégration d'enquêtes européennes (TIC, SRCV) dans le questionnaire, modalités de collecte modifiées pour l'enquête emploi (intégrée depuis 2005 dans le Mikrozensus), et introduction de la collecte par Internet. Ceci dotera l'Allemagne d'un dispositif de grande qualité pour l'observation de la situation démographique et sociale du pays, et suffisamment souple pour répondre aux besoins nouveaux de connaissance.

 *Conducted for more than 60 years, the German Mikrozensus is a survey of the whole population, with a sampling rate of 1 %. It is the largest household survey in the European Union. In accordance with the organisation of its statistical system, the German Federal Statistical Office defines its methodology, which looks like that of the French employment survey. And the data collection is carried out by the statistical offices of the Länder. From the beginning, focusing on issues of household composition, housing, activities, income and training/qualification, the themes of the questionnaire have changed over time. The diversity of the fields covered makes it possible to carry out varied and detailed analyses of the structural developments in the German society. By 2020-21, the new Mikrozensus will allow a greater rationalisation of the German household statistical survey system, for several reasons : integration of European surveys (ICT, SILC) into the questionnaire, new collection methods for the Labour Force Survey (integrated since 2005 into the Mikrozensus), and introduction of Internet collection. This will provide Germany with a high-quality framework for monitoring the country's demographic and social situation, and flexible enough to meet new data needs.*

* Chef-adjoint de l'unité Ressources documentaires et Archiving, Insee,
pascal-m.riviere@insee.fr

UNE TRÈS GRANDE ENQUÊTE, QUI A RÉGULIÈREMENT ÉVOLUÉ

Le *Mikrozensus* allemand a été institué *via* la loi du 16 mars 1957 pour une statistique représentative de la population, du travail et des revenus. Appliquant un taux de sondage de 1 % à l'ensemble de la fédération, elle concerne chaque année de l'ordre de 800 000 personnes dans 370 000 ménages, ce qui en fait la plus grande enquête auprès des ménages¹ du système statistique européen. Ce taux de sondage important est motivé par deux types de raison : interroger une population de plus grande taille permet des analyses portant sur des sous-populations plus fines (selon tout type de critère) et une meilleure précision des résultats.

Contrairement à ce que le terme laisse imaginer, le *Mikrozensus* n'a rien d'un micro-recensement : cette enquête ne sert pas particulièrement à estimer des grandeurs démographiques fédérales (comme la population totale), et s'appuie d'ailleurs sur le recensement pour redresser ses résultats². En outre, elle permet de traiter une vaste gamme de sujets, qui s'est élargie avec le temps : le questionnaire est passé de 52 questions en 1957 à 190 en 2005, pour atteindre 220 questions en 2018.

Au cours d'une période aussi longue, les modalités de mise en œuvre ont naturellement évolué. Ainsi, l'échantillonnage qui était fondé sur les logements a changé pour s'appuyer sur les aires dans les années 60. Pour les enquêteurs, la collecte par micro-ordinateur portable, testée en 1991, s'est généralisée en 2005. Toujours en 2005, pour répondre à une demande européenne croissante, la collecte a été étalée sur l'ensemble de l'année avec l'introduction de la notion de semaine de référence glissante (au lieu de prendre une semaine fixe au mois d'avril). Et de nouvelles améliorations sont attendues pour 2020 : intégration d'enquêtes européennes, en lien avec l'adoption du règlement cadre européen sur les statistiques sociales (IESS), collecte multi-modes.

Malgré son indéniable utilité et sa riche histoire, le *Mikrozensus* n'est paradoxalement pas très connu du grand public. En effet, si les résultats de celui-ci interviennent dans un grand nombre de publications, le nom de *Mikrozensus* n'y est pas reporté de façon visible. De plus, il est largement éclipsé par la notoriété du recensement exhaustif de population. Il en résulte un besoin régulier d'information des personnes interrogées³.

UN VÉHICULE POUR DE MULTIPLES THÉMATIQUES

Au sein de la statistique publique allemande, le *Mikrozensus* joue un rôle tout à fait particulier. On peut le considérer comme le pilier central du dispositif annuel d'observation de la situation économique et sociale de la population. Son rôle principal est de fournir au Parlement, au gouvernement et aux administrations de la fédération et des *Länder* des informations récentes, exhaustives et fiables. Ces informations concernent la structure démographique, la situation économique et sociale des personnes, des familles et des ménages, le marché du travail, la formation et les conditions de logement.

Le *Mikrozensus* est notamment le véhicule de l'enquête emploi (*Labour Force Survey - LFS*) de l'Union européenne en Allemagne qui produit les données de référence relatives au chômage et au marché du travail. Le questionnaire comporte des thématiques variées,

1. Au sens où il s'agit d'une enquête dont le mode de collecte privilégié est l'interrogation en face-à-face. En tant qu'enquête multi-modes (internet et papier), le recensement français (pour ne prendre que cet exemple) est plus volumineux que le *Mikrozensus* allemand.

2. Ajusté des naissances, des décès et des flux migratoires annuels ultérieurs.

3. Sur l'histoire du *Mikrozensus*, voir (Anders, 2018).

telles que : données démographiques détaillées, sources de revenus, couverture retraite, établissement scolaire ou universitaire fréquenté, diplôme, données d'activité détaillées, utilisation d'internet, logement, santé, handicap, privations matérielles, caisse d'assurance maladie, trajet quotidien domicile/travail. Les questions sont toutes obligatoires, à de rares exceptions près⁴.

Le caractère annuel de ces informations structurelles permet en outre de répondre à une demande institutionnelle croissante de données aussi récentes que possible.

Son rapport coût-bénéfice est d'autant plus favorable que le *Mikrozensus* fournit un échantillon des ménages réutilisable pour d'autres besoins : il constitue ainsi un véritable « véhicule » auquel peuvent se raccrocher des enquêtes. En particulier, son questionnaire annuel peut être modulé pour répondre à des demandes récemment exprimées. Par ailleurs, les résultats du *Mikrozensus* peuvent servir de données de redressement - ou de comparaison - à certaines autres enquêtes ménages.

Cette souplesse apparente du *Mikrozensus* doit néanmoins être modérée. En effet, en Allemagne, la loi⁵ inscrit explicitement les statistiques publiques comme un instrument de décision, et elle encadre fortement le fonctionnement du système statistique. Ainsi, l'établissement de nouvelles statistiques, ou la modification de statistiques existantes (comme celles issues du *Mikrozensus*) doit se faire au travers de l'adoption d'une loi fédérale *ad hoc* (**encadré 1**).

Encadré 1. Les incontournables pour une loi sur une nouvelle statistique —

La loi statistique à des fins fédérales (*Bundesstatistikgesetz*) est une loi chapeau. Elle prévoit notamment que chaque nouvelle statistique doit faire l'objet d'une loi. Cette dernière doit *a minima* définir les variables à collecter, les variables d'identification des unités observées, le mode de collecte, la période et la périodicité, le champ.

Des statistiques peuvent être exemptées d'une loi dans des circonstances bien déterminées, notamment le recours à des données déjà existantes et d'accès public, ou le besoin de répondre à un besoin spontané émanant des « hautes instances de la fédération ».

À titre d'illustration, la loi du 17 janvier 1996 définissait les caractéristiques du *Mikrozensus* jusqu'à l'année de collecte 2004 incluse. Les thématiques étaient liées aux caractéristiques du ménage, du logement, aux revenus, à l'activité, au niveau de qualification, aux prestations sociales, à la situation de santé ; il n'y avait qu'une seule semaine de référence dans l'année pour la collecte ; les ménages ne pouvaient pas être interrogés plus de quatre années de suite*.

La loi du 24 juin 2004 instaurait le *Mikrozensus* jusqu'à l'année de collecte 2016 incluse. On y trouvait un nombre plus important de questions relatives à l'éventuel statut de migrant de la personne. La collecte est depuis lors étalée sur l'ensemble de l'année.

La loi du 7 décembre 2016, qui contient des changements importants (dont l'intégration de questionnaires européens et le multi-modes pour la collecte) ne prévoit pas de date de fin pour la collecte du *Mikrozensus*.

* Symptomatique de la sensibilité autour des données individuelles en Allemagne, on relèvera le paragraphe 6 de la loi du 17 janvier 1996 qui précise que « les enquêteurs ne devront pas collecter de données dans leur voisinage direct ».

4. Par exemple, pour les femmes de 15 ans à 75 ans, le nombre d'enfants nés vivants, les modalités de recherche de travail et le handicap ou certaines questions de SRCV (statistiques sur les ressources et conditions de vie) et de TIC (technologies de l'information et de la communication).

5. La « loi statistique à des fins fédérales » (*Bundesstatistikgesetz*), telle que modifiée au 20 octobre 2016.

UN ÉCHANTILLONNAGE ASSURÉ PAR DESTATIS...

L'organisation du processus de production du *Mikrozensus* fait intervenir le niveau fédéral et les *Länder*. Les directives de la collecte émanent de Destatis, l'office fédéral de statistique allemand. Elles font l'objet de concertations avec les offices statistiques des *Länder* en vue de leur mise en œuvre effective sur le terrain, notamment au sein du Conseil statistique (*Statistischer Beirat*). S'appuyant sur des comités spécialisés, le secrétariat de ce conseil est assuré par Destatis. Il réunit les directeurs des offices statistiques des *Länder* et les représentants de nombreux utilisateurs, dont les partenaires sociaux. Son rôle est de conseiller Destatis, tant sur le programme de travail que sur certaines questions de méthodologie.

Le *Mikrozensus* couvre une très large part de la population résidant en Allemagne, y compris celle vivant en communautés⁶ (notamment les personnes en cités étudiantes, en maisons

de retraite, à l'hôpital...). Il n'enquête toutefois pas les personnes sans logement déclaré, et celles vivant en logement provisoire, comme les demandeurs d'asile notamment.

“ Le *Mikrozensus* couvre une très large part de la population résidant en Allemagne, y compris celle vivant en communautés [...] En cas de nécessité, les échantillons peuvent être augmentés dans certaines zones de niveau NUTS 2 de plus petite taille. ”

Techniquement, le *Mikrozensus* recourt à un échantillonnage de groupes de logements contigus, c'est-à-dire par grappes (Bihler et Zimmermann, 2016). Cette méthode a été choisie pour des raisons pratiques : le travail d'un enquêteur est facilité puisque les personnes qu'il devra interroger seront regroupées dans une même zone géographique. Destatis assume les inconvé-

nients de cette méthode qui induit une certaine perte de précision. En effet, ce procédé revient à interroger des populations qui peuvent avoir tendance à se « ressembler » (comme des habitants d'une même rangée de maisons mitoyennes) ; mais la taille globale de l'échantillon (1 % de la population totale) est suffisamment importante pour assurer la qualité statistique nécessaire aux résultats publiés.

L'objectif est d'avoir des résultats représentatifs au niveau géographique dit NUTS 2⁷. En cas de nécessité, les échantillons peuvent être augmentés dans certaines zones de niveau NUTS 2 de plus petite taille ; ceci se fait en diminuant les tailles d'échantillons d'autres zones, puisque le taux global de sondage reste de 1 %. Le tirage de l'échantillon de 1 % de ces unités est décrit dans l'**encadré 2**.

... UNE COLLECTE MENÉE PAR LES OFFICES STATISTIQUES DES LÄNDER...

Si l'échantillonnage du *Mikrozensus* est effectué par Destatis, ce sont les offices statistiques des *Länder* qui pilotent et mettent en œuvre son processus de collecte. Ils recrutent les

6. Contrairement à l'enquête emploi en France qui n'enquête que les personnes vivant en logement ordinaire.

7. Nomenclature (européenne) des unités territoriales statistiques. Le niveau 2 correspond en France aux 22 anciennes régions, à des regroupements de cantons (*Kreise*) en Allemagne.

enquêteurs et leur font suivre, au moins une fois par an, une formation spécifique. Celle-ci porte sur le contenu du questionnaire, la manière de mener un entretien et le travail de saisie associé (*via* CAPI/Blaise).

En amont de la collecte se déroulent deux opérations préparatoires :

- ① en début d'année ou de trimestre, les adresses de plus de 10 logements font l'objet d'un repérage de façon à déterminer les logements précis à visiter ;
- ① une semaine avant le recensement proprement dit, les enquêteurs font une tournée pour annoncer leur visite et identifier les personnes qui seront effectivement concernées. Une brochure explicative est laissée à l'intention de chaque foyer, ainsi qu'une proposition concrète de rendez-vous.

De façon à assurer une meilleure qualité des réponses, l'entretien en face-à-face est privilégié par l'enquêteur. À défaut, un questionnaire papier est laissé au ménage⁸. En 2016, le taux de réponse était de 97,4 %. Les réponses obtenues se répartissaient en 67 % en face-à-face, 30 % *via* le questionnaire papier, 3 % par téléphone. Un entretien avec un ménage d'une personne dure en moyenne 15 à 20 minutes, et 10 minutes de plus, s'il y a deux personnes. Bien entendu, si le ménage est plus grand ou si ses membres ont une mauvaise connaissance de l'allemand, le temps d'entretien est plus long. En cas de nécessité, un interprète utilisant le langage des sourds peut être mobilisé. Les ménages ne répondant pas font l'objet d'un avertissement par l'office statistique du *Land* ; celui-ci mentionnant une amende de

Encadré 2. Le tirage de l'échantillon de 1%

Source : *Die neue Mikrozensus-Stichprobe 2016* (W. Bihler, D. Zimmermann)

L'échantillonnage est réalisé par strates. Celles-ci sont définies à partir de deux critères : les zones géographiques et les groupes d'adresses. On compte 243 zones géographiques « d'un seul tenant » (d'au moins 200 000 habitants), qui découpent le territoire fédéral, et quatre groupes d'adresses définis selon le nombre de logements qu'ils comprennent. Les adresses sont actualisées annuellement au vu des nouvelles constructions enregistrées.

Le croisement de ces deux critères détermine $4 \times 243 = 972$ strates. Au sein de chaque strate, des unités de tirage homogènes sont définies. La cible est d'avoir des unités correspondant, en moyenne globale, à 9 logements et 15 personnes. Au cours de ce processus, près de 19 millions d'adresses définissent ainsi environ 4,6 millions d'unités de tirage.

De façon à assurer une base de sondage aussi représentative que possible au niveau local, les unités (i.e. les paquets de logements) sont ordonnées géographiquement** et regroupées par zones de 100 unités. Puis, au sein de chaque zone, une unité sur cent est sélectionnée pour faire partie de l'échantillon courant. De façon à respecter les termes de la loi selon laquelle « une unité ne doit pas être interrogée plus de quatre fois au cours d'une période de cinq années », un système de rotation d'année en année est organisé parmi toutes les unités de chaque zone.

** Schématiquement : cantons, communes, quartier, rue, numéro et numéro-lettre dans la rue.

8. En 2020, se rajoutera le questionnaire sous forme électronique (internet).

300 euros (la consigne étant de l'appliquer 9-10 semaines après la première visite), il s'ensuit le plus souvent le renvoi par le ménage du questionnaire rempli⁹.

L'ensemble des questionnaires sont traités par les offices des *Länder*. Ceux-ci calculent et diffusent les agrégats au niveau local, sur la base d'une méthode élaborée par Destatis. Les données individuelles anonymisées sont envoyées par voie électronique chaque mois à Destatis, qui produit les agrégats au niveau fédéral.

... DES PUBLICATIONS ET ANALYSES AU NIVEAU FÉDÉRAL ET À CELUI DES LÄNDER

Les résultats du *Mikrozensus* sont publiés très rapidement, environ trois mois après la fin de la collecte de l'année antérieure. Seules des données portant sur l'ensemble de l'année sont publiées par le système statistique allemand. En effet, les résultats mensuels ou trimestriels qu'il est possible de calculer du fait de l'étalement annuel de la collecte sont considérés comme étant trop volatils.

Le *Mikrozensus* donne chaque année une information structurelle de la situation démographique et sociale en Allemagne, sous forme de

« Des documentations motivant la conduite du *Mikrozensus* accompagnent systématiquement les publications sur les sites internet, insistant notamment sur son utilité collective, sa base légale, et le soin apporté à la gestion de la confidentialité. »

tableaux publiés par Destatis au niveau fédéral et par les offices statistiques des *Länder* pour leur champ géographique. Ils sont organisés par grands domaines tels que « Niveau de formation de la population », « Questions sur la santé, le poids, la consommation de tabac... », « Trajets domicile/travail ou lieu de formation », etc. Des documentations motivant la conduite du *Mikrozensus* accompagnent systématiquement les publications sur les sites internet, insistant notamment sur son utilité collective, sa base légale, et le soin apporté à la gestion de la confidentialité.

L'Office fédéral édite d'autre part des publications plus détaillées, notamment sur les thèmes suivants : « Participation de la population au marché du travail », « Population issue de la migration » (**encadré 2**), « Ménages et familles », « Caractéristiques des logements et conditions de logement des ménages », « Données sur l'assurance-maladie », « Les parents isolés ».

Grâce à leur richesse, ces publications éclairent le débat social en Allemagne. Ainsi, celle publiée en août 2018 sur les parents isolés décrit en 50 pages la situation démographique de ces personnes (âge, sexe, nombre d'enfants à charge), leur niveau de qualification, leur situation sur le marché du travail, leurs conditions de vie, leur situation vis-à-vis du surendettement. De même, la publication « Population issue de la migration¹⁰ » dresse depuis 2005 un état très complet de la composition de la population allemande. Au sein de ce document de pas moins de 500 pages figurent ainsi des données très détaillées sur le pays d'origine des personnes, la composition de leur famille, la langue parlée à la maison, les raisons de la migration, le niveau de formation, l'activité en emploi, les revenus, la santé, etc.

9. Il est à noter que l'obligation de répondre demeure si un contrôle de plausibilité détecte une erreur « de saisie ». Le ménage est alors contacté par écrit en vue de vérifier/corriger ses réponses.

10. *Bevölkerung mit Migrationshintergrund* (Destatis, 2018a).

Divers publics, universitaires notamment, font usage des résultats du *Mikrozensus* pour leurs propres analyses et études. Une conférence annuelle organisée par le laboratoire de microdata de l'institut Leibniz de sciences sociales (situé à Mannheim) est l'occasion de présenter certains de ces travaux. Par exemple, en dehors de sujets liés à la méthodologie même, la conférence de fin novembre 2018 a abordé les diverses thématiques suivantes : famille, inégalités sociales, marché du travail, migration et intégration, indicateurs régionaux. Sur le sujet sensible de la migration, cette conférence a traité des « Discriminations locales des populations à passé d'immigration », des « Processus intergénérationnels d'intégration » et du thème « Fumer et surpoids parmi les personnes à passé d'immigration ».

UNE GRANDE TRANSPARENCE SUR LA QUALITÉ DES RÉSULTATS

Comme pour un certain nombre de ses statistiques, Destatis publie sur son site internet un rapport qualité annuel portant sur le *Mikrozensus* de l'année précédente.

Ce rapport de 16 pages expose le contenu et l'utilité du *Mikrozensus*, la méthodologie appliquée, il analyse la fiabilité et la cohérence des résultats (avec ceux d'autres statistiques), et donne les différents canaux de diffusion. Il fournit de nombreuses précautions de lecture et d'interprétation : elles concernent notamment les écarts dans les statistiques d'emploi

Encadré 3. Le rapport qualité de juillet 2018

Le rapport qualité met en lumière certaines spécificités de la collecte du *Mikrozensus* pouvant avoir un impact sur la lecture et l'interprétation des résultats. Sa structure reste similaire d'une année sur l'autre.

Il insiste tout particulièrement sur les statistiques de l'emploi. En 2017, le nombre d'actifs en Allemagne selon le *Mikrozensus* est inférieur de 2,2 millions (5,7 %) à celui donné par les comptes nationaux. Ces écarts s'expliquent principalement par des définitions et des méthodes de collecte différentes*. Il faut ainsi souligner, dans la collecte du *Mikrozensus*, la propension des personnes à oublier de mentionner leur éventuel statut de travailleur occasionnel. Des contre-interviews ont confirmé cette tendance à négliger la déclaration de petits emplois rémunérés ; c'est le cas en particulier pour les étudiants, les retraités : les personnes concernées se perçoivent d'abord comme étudiantes, ou retraitées, avant de se considérer comme travailleurs occasionnels.

Le rapport qualité rappelle également que, depuis 2005, la semaine de référence du questionnaire n'est plus située au mois d'avril mais devient « la semaine précédant celle de l'interview » (on est passé à une collecte tout au long de l'année). L'existence de fluctuations infra-annuelles des phénomènes observés limite par conséquent la portée des comparaisons avec des données antérieures à 2005. Par ailleurs, le choix d'une semaine de référence « juste antérieure à celle de l'interview effective du ménage » conduit à une répartition non uniforme au cours de l'année des semaines effectivement collectées. En effet, durant les périodes de congés et en toute fin d'année, les ménages sont moins facilement joignables et ne répondent ainsi au questionnaire qu'avec retard par rapport à la date prévue a priori dans l'échantillonnage.

Enfin, en termes de précision statistique, le taux de sondage à 1 % conduit à ne pas publier des informations concernant une sous-population totale inférieure à 5 000 individus.

* En 2016, Destatis estimait les écarts dus aux définitions différentes à 401 000 personnes et ceux dus à des questions de méthodologie ou de sources à 1 802 000 personnes (NB : les comptes nationaux s'appuient principalement sur des données administratives).

entre le *Mikrozensus* et les comptes nationaux, les conséquences du changement de semaine de référence de 2005 et celles des modifications dans le temps de certaines questions relatives au statut d'emploi et à la formation.

DES GAINS D'EFFICACITÉ À VENIR GRÂCE À L'INTÉGRATION D'ENQUÊTES EUROPÉENNES...

Confrontés à une demande croissante de statistiques démographiques et sociales, notamment en provenance de l'Union européenne, à une baisse de leurs budgets et à des demandes d'allègement de la charge de réponse des enquêtes, Destatis et les offices statistiques des *Länder* ont décidé d'aller vers un système intégré d'enquêtes auprès des ménages, en lien avec le règlement cadre européen sur les statistiques sociales (IESS)¹¹.

La loi fédérale *ad hoc* du 7 décembre 2016 traduit ces évolutions dans le droit. Elle pérennise le *Mikrozensus* allemand dans la mesure où elle ne définit plus de date de fin pour cette opération¹². Elle acte par ailleurs l'intégration de plusieurs enquêtes européennes existantes et des modifications des modalités de collecte.

« L'intégration de ces enquêtes va permettre une rationalisation du dispositif des enquêtes auprès des ménages. »

Le nouveau *Mikrozensus* va ainsi prendre en compte des besoins européens et notamment les règlements SRCV (statistiques sur les ressources et conditions de vie) et TIC (technologies de l'information et de la communication), dont il

intégrera les questionnaires. Ceci permettra à Destatis de supprimer les deux enquêtes spécifiques SRCV et TIC.

L'intégration de ces enquêtes va permettre une rationalisation du dispositif des enquêtes auprès des ménages, qui partageront dès lors un même schéma d'échantillonnage, un même dispositif de collecte, les mêmes outils et infrastructures informatiques. Leur mise en commun dégagera des synergies évidentes, mais accroîtra le degré de complexité de la gestion et du pilotage du *Mikrozensus*.

En particulier, si le taux de sondage de 1 % est conservé pour les questions du « tronc commun » des enquêtes auprès des ménages, pour les enquêtes européennes, ce seront des sous-échantillons qui sont interrogés. Ainsi 45 % de l'échantillon complet (correspondant au taux de sondage à 1 %) seront également questionnés au titre de l'enquête emploi, 12 % au titre de SRCV et 3,5 % au titre de TIC (*figure 1*). Pour prolonger la longueur du panel de l'enquête SRCV, il sera demandé aux ménages s'ils veulent bien continuer à répondre à l'enquête au-delà du maximum prévu par la loi de quatre fois sur cinq années.

Cette intégration d'enquêtes sera également l'occasion d'un travail d'harmonisation de questionnaires, notamment ceux de SRCV et de l'enquête emploi, qui actuellement présentent plusieurs différences. C'est le cas, par exemple, pour le choix du travail à temps partiel :

11. *Integrated European Social Statistics*.

12. La présentation du projet de loi au Parlement note qu'il n'y a plus d'opposition de principe à la collecte *via* le *Mikrozensus* (dans les années 80, c'est plus le caractère obligatoire du *Mikrozensus* qui a été remis en cause que son existence même). En revanche, le besoin d'ajuster régulièrement ce dispositif aux demandes nouvelles est bien identifié.

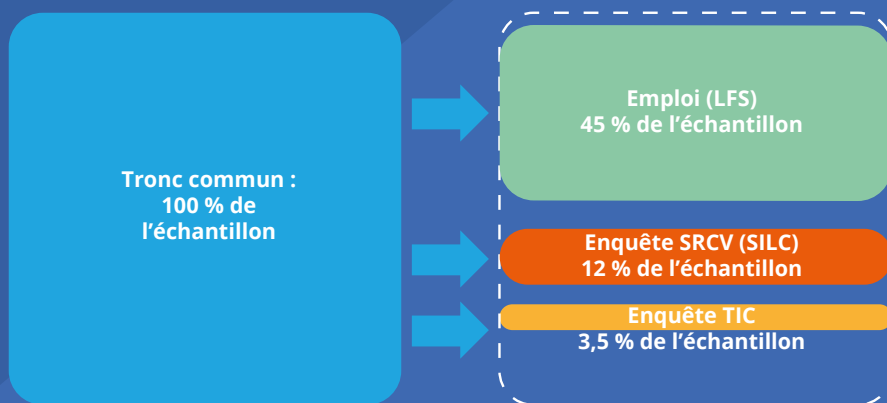
« Cette intégration d'enquêtes sera également l'occasion d'un travail d'harmonisation de questionnaires. »

actuellement, l'enquête SRCV demande « la raison d'un travail de moins de 30 heures hebdomadaires » alors que l'enquête emploi demande « la raison d'un travail à temps partiel ». Les différences qui existent pour d'autres questions, notamment sur la recherche d'emploi, le nombre d'heures travaillées, et les raisons du changement d'emploi, seront également supprimées.

La qualité des résultats devrait s'en trouver améliorée. En effet, le caractère obligatoire du *Mikrozensus*, contrairement à la plupart des enquêtes ménages actuelles, entraînera automatiquement un meilleur taux de réponse pour les enquêtes SRCV et TIC dès lors que leurs questions y seront intégrées. En outre, l'échantillon concerné par les questions SRCV va passer de 14 000 à 40 000 ménages, ce qui permettra une représentativité des résultats au niveau de chaque *Land* comme le prévoit le nouveau règlement cadre européen sur les statistiques sociales (IESS). Le délai de production des résultats sera raccourci ; par exemple, le délai de transmission des résultats de SRCV devrait être ramené de un an à six mois.

Enfin, à travers la présentation du projet de loi au Parlement, il apparaît que la rationalisation induite du dispositif d'enquêtes devrait contenir les coûts, tout en garantissant le degré de précision imposé.

Figure 1. Intégration des enquêtes européennes dans le *Mikrozensus*



Note de lecture : 100 % de l'échantillon interrogé répond aux questions du tronc commun du Mikrozensus. Parmi eux, 45 % maximum de l'échantillon répond à l'enquête « emploi », 12 % à l'enquête « ressources et conditions de vie » et 3,5 % répond à l'enquête sur les « technologies de l'information et de la communication ».

... ET À L'ÉVOLUTION DES MODALITÉS DE COLLECTE, NOTAMMENT LE MULTI-MODES

Les évolutions porteront aussi sur la collecte elle-même. Tout d'abord, le nombre d'interviews de l'enquête emploi (règlement LFS), déjà intégrée au *Mikrozensus* depuis 2005, va nettement augmenter : il passera de 340 000 à 500 000 par an. Ses modalités de collecte seront aussi modifiées, afin de faire baisser la volatilité des résultats mensuels et trimestriels. Cela permettra à l'Allemagne de répondre aux futures demandes européennes de publication de statistiques infra-annuelles harmonisées entre pays issues de l'enquête emploi (Marder-Puch, 2018). L'objectif de ces changements est aussi de pouvoir disposer de données plus représentatives de l'ensemble de l'année.

En particulier, une rotation infra-annuelle de l'échantillon du nouveau *Mikrozensus* sera réali-

« Pour les répondants, l'allègement recherché du questionnaire et la suppression de certaines enquêtes auprès des ménages devraient conduire à terme à une moindre charge globale de réponse. »

sée, comme c'est la pratique dans les autres pays de l'Union européenne. D'une part, les ménages interrogés au titre de l'enquête emploi le seront à plusieurs reprises durant 18 mois¹³. D'autre part, la semaine de référence sera modifiée. Actuellement, le fait d'avoir une « semaine de référence juste antérieure à celle de l'interview effective du ménage » conduit à une répartition non uniforme de la collecte au cours de l'année et donc de la représentation des phénomènes étudiés ; dans le nouveau *Mikrozensus*, les ménages se verront attribuer un questionnaire avec une semaine de référence parfaitement déterminée (quitte

à répondre avec retard en cas de congés) en sorte que la collecte porte sur chacune des 52 semaines de l'année. L'interrogation répétée des mêmes ménages durant 18 mois peut éventuellement générer des biais¹⁴ : les statisticiens allemands conduisent des tests sur le terrain pour déterminer la meilleure façon de les diminuer.

Par ailleurs, à partir de 2020, les enquêtés auront désormais la possibilité de répondre par Internet (en mode CAWI¹⁵). L'interview en face-à-face restera toutefois privilégiée (en mode CAPI¹⁶), notamment à chacune des quatre vagues d'interrogations annuelles pour l'échantillon SRCV, et lors de la première interrogation pour les questions du tronc commun ainsi que pour celles de l'enquête emploi ou de TIC. Hormis pour les ménages interrogés au titre de SRCV, les vagues suivantes d'interrogations se font par téléphone (en mode CATI¹⁷).

Au total, la rationalisation du dispositif d'enquêtes auprès des ménages et le développement de la collecte multi-modes devraient permettre de contenir les coûts globaux du *Mikrozensus* tout en améliorant la qualité des statistiques produites. Pour les répondants, l'allègement recherché du questionnaire et la suppression de certaines enquêtes auprès des ménages devraient conduire à terme à une moindre charge globale de réponse.

13. Un même ménage sera désormais interrogé deux trimestres de suite (T1 et T2) et ne sera pas sollicité les deux trimestres suivants (T3 et T4), avant de l'être à nouveau les deux trimestres ultérieurs (T5 et T6).

14. Pour abrégé l'interview, les ménages peuvent être en particulier tentés de ne pas déclarer de changement de leur situation par rapport à l'interview précédente.

15. CAWI = Computer-Assisted Web Interview.

16. CAPI = Computer-Assisted Personal Interview.

17. CATI = Computer-Assisted Telephone Interview.

Le calendrier de mise en application de ce nouveau *Mikrozensus* est le suivant :

- 🟢 **2017-2019 :** Simplification et mise en cohérence interne du questionnaire ;
- 🟢 **2019 :** Tests de la mise en place des nouvelles modalités évoquées ci-dessus (sur un échantillon de 0,025 %, soit environ 20 000 personnes) ;
- 🟢 **2020-2021 :** Mise en œuvre du nouveau *Mikrozensus*.

À cet horizon, le système statistique allemand devrait bénéficier d'un dispositif intégré d'enquêtes auprès des ménages, efficient et fournissant des résultats d'une grande qualité. Sa relative souplesse lui permettra de s'adapter plus facilement qu'actuellement aux besoins nouveaux de connaissance de la situation démographique et sociale de l'Allemagne et de ses territoires.

BIBLIOGRAPHIE

ANDERS, Carsten, 2018. 60 Jahre Mikrozensus – Ein Rück- und Ausblick. In : *Statistisches Monatsheft Baden-Württemberg* [en ligne]. Vol. 2/2018, pp. 3-9. [Consulté le 23 mai 2019]. Disponible à l'adresse :

https://www.statistik-bw.de/Service/Veroeff/Monatshefte/PDF/Beitrag18_02_01.pdf

BIHLER, Wolf et ZIMMERMANN, Daniel, 2016. Die neue Mikrozensusstichprobe ab 2016. In : *WISTA – Wirtschaft und Statistik* [en ligne]. 15 décembre 2016. Vol. 6/2016, pp. 20-29. [Consulté le 23 mai 2019]. Disponible à l'adresse :

https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2016/06/die-neue-mikrozensusstichprobe-062016.pdf?__blob=publicationFile&v=3

DESTATIS, 2018a. *Bevölkerung mit Migrationshintergrund – Ergebnisse des Mikrozensus 2017* [en ligne]. Fachserie 1 Reihe 2.2. [Consulté le 23 mai 2019]. Disponible à l'adresse : https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Publikationen/Downloads-Migration/migrationshintergrund-2010220177004.pdf?__blob=publicationFile&v=4

DESTATIS, 2018b. *Mikrozensus – Qualitätsbericht – 2017* [en ligne]. 7 décembre 2018. [Consulté le 23 mai 2019]. Disponible à l'adresse :

https://www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Bevoelkerung/mikrozensus-2017.pdf?__blob=publicationFile&v=4

LEFÈVRE, Cécile, 1999a. La statistique en Allemagne – Un homologue proche, mais en partie méconnu. In : *Courrier des statistiques* [en ligne]. Décembre 1999. N° 91-92, pp. 5-14. [Consulté le 23 mai 2019]. Disponible à l'adresse :

<https://www.epsilon.insee.fr/jspui/bitstream/1/8525/1/cs91-92b.pdf>

LEFÈVRE, Cécile, 1999b. Le Mikrozensus, pierre angulaire des enquêtes statistiques auprès des ménages allemands. In : *Courrier des statistiques* [en ligne]. Décembre 1999. N° 91-92, pp. 15-22. [Consulté le 23 mai 2019]. Disponible à l'adresse :

<https://www.epsilon.insee.fr/jspui/bitstream/1/8523/1/cs91-92c.pdf>

MARDER-PUCH, Katharina, 2018. The new German Mikrozensus as an integrated survey for household statistics. In : STATISTICS ICELAND. *Labour Force Survey Workshop 2018, Reykjavík, 17 et 18 mai 2018*.

FONDEMENTS JURIDIQUES

Gesetz zur Neuregelung des Mikrozensus und zur Änderung weiterer Statistikgesetze (Bundestag, 7 décembre 2016). *Bundesgesetzblatt.de* [en ligne]. 13 décembre 2016. [Consulté le 23 mai 2019]. Disponible à l'adresse :

https://www.bgbl.de/xaver/bgbl/start.xav?start=%2F%2F*%5B%40attr_id%3D%27bgbl116s2826.pdf%27%5D#_bgbl_%2F%2F*%5B%40attr_id%3D%27bgbl116s2826.pdf%27%5D__1558620398332

LA MESURE DE LA SÉCURITÉ ET DE LA SATISFACTION VIS-À-VIS DES INSTITUTIONS EN FRANCE

L'IMPULSION DONNÉE PAR LES OBJECTIFS DU DÉVELOPPEMENT DURABLE DES NATIONS UNIES

François Clanché*

La mesure de l'efficacité des institutions et à un moindre niveau celle de la sécurité ne font pas partie des domaines couverts de longue date par la statistique publique française. Pour autant, la France est d'ores et déjà en mesure de renseigner plusieurs des indicateurs du 16^e objectif de développement durable des Nations Unies (ODD 16) « Paix, justice et institutions efficaces ». Parfois, c'est en mobilisant des sources administratives, mais le plus souvent grâce aux enquêtes de victimation. Afin d'améliorer sa couverture du champ des ODD, la statistique publique française fait évoluer en 2018 son dispositif d'enquêtes auprès des ménages, pour appréhender notamment la mesure des discriminations et de la corruption. Certains indicateurs restent cependant non couverts pour le moment, soit pour des raisons techniques, soit parce que les définitions internationales opératoires ne sont pas encore fixées, soit parce que la statistique publique hésite à investir des champs nouveaux, notamment celui de la gouvernance, pas toujours facile à appréhender conceptuellement et à mesurer de manière robuste.

 *French official statistics do not consider subjects related to governance and security as part of their core business. Nonetheless, the French official statistics system is already able to produce several indicators of the Sustainable Development Goal 16, entitled « Peace, justice and strong institutions ». Some of them derive from administrative data, but most of them from victimisation surveys. In order to improve its monitoring of SDG 16, especially on the measurement of discriminations and corruption, the French official statistics system is adapting its household surveys on these subjects. However, some indicators cannot still be produced, either for technical reasons, or because international operational definitions are not yet fixed, or because official statistics hesitate to invest in new fields, notably governance, which is not always easy to understand conceptually and to measure in a robust way.*

* Ancien chef du service statistique ministériel de la sécurité intérieure, Ministère de l'Intérieur,
francois.clanche@pm.gouv.fr

Les Nations Unies ont adopté en septembre 2015 dix-sept objectifs de développement durable pour leur agenda 2030 (**encadré 1**). Le 16^e objectif « Paix, justice et institutions efficaces » met l'accent sur la promotion de sociétés pacifiques et inclusives pour le développement durable, l'accès à la justice pour tous et le renforcement des institutions responsables et efficaces à tous les niveaux. Les cibles et indicateurs qu'il promeut (**encadré 2**), et qu'il demande aux statisticiens publics de renseigner, portent sur la sécurité des personnes et des biens dans leurs sociétés (y compris au sein des familles), ainsi que sur la confiance dans les institutions et dans l'efficacité et l'honnêteté perçue de ces dernières. Le présent article fait le point sur la disponibilité des indicateurs relatifs à la paix, la justice et les institutions efficaces en France.

Encadré 1. Trois buts essentiels sont poursuivis : lutter contre les inégalités, l'exclusion et les injustices ; faire face au défi climatique ; mettre fin à l'extrême pauvreté.

- | | |
|---|---|
| 1 : Éradication de la pauvreté | 10 : Réduction des inégalités |
| 2 : Sécurité alimentaire et agriculture durable | 11 : Villes et communautés durables |
| 3 : Santé et bien-être | 12 : Consommation et production responsables |
| 4 : Éducation de qualité | 13 : Lutte contre les changements climatiques |
| 5 : Égalité entre les femmes et les hommes | 14 : Vie aquatique marine |
| 6 : Gestion durable de l'eau pour tous | 15 : Vie terrestre |
| 7 : Énergies propres et d'un coût abordable | 16 : Paix, justice et institutions efficaces |
| 8 : Travail décent et croissance durable | 17 : Partenariats pour la réalisation des objectifs |
| 9 : Infrastructures résilientes et innovation | |

Pour plus d'information consulter le site des Nations Unies <https://www.un.org/sustainabledevelopment/fr> [Consulté le 6 juin 2019].

MESURER LA VIOLENCE OU LA QUALITÉ DES INSTITUTIONS N'EST PAS DANS LA TRADITION STATISTIQUE FRANÇAISE

Cet objectif a donné une impulsion au développement des statistiques relatives à la sécurité et à la qualité des institutions. En effet, la mesure de la violence (qu'elle soit physique, sexuelle ou psychologique), de la corruption, ou encore de l'appréciation par les individus de la qualité des institutions ne font pas partie de la tradition statistique française, telle qu'elle s'est construite et développée, essentiellement dans la seconde moitié du XX^e siècle. Celle-ci s'était alors essentiellement centrée sur le développement économique, sur l'emploi et les conditions de travail ainsi que sur les ressources et les conditions de vie des ménages, dans une société où la sécurité était considérée comme posant peu de problèmes, d'où la corruption était jugée quasi absente dans un pays dit « développé » et où la légitimité des institutions démocratiques, au lendemain de la deuxième guerre mondiale, faisait l'objet d'un vaste consensus. Dans ces conditions, le statisticien public devait essentiellement mesurer les progrès de l'économie, la répartition des emplois et des ressources entre groupes sociaux et leur accès progressif aux différents biens et services produits par une société en croissance. Les difficultés économiques apparues dans les dernières décennies ont accru les besoins autour des questions du chômage, et des inégalités de revenus et de la précarité économique et sociale, mais sans pour autant faire remonter au premier plan

les sujets de sécurité et de bonne gouvernance publique. Pour autant, au début des années 2000, dans un contexte de montée en puissance des préoccupations liées à l'insécurité encore renforcée très récemment après les vagues d'attentats, un observatoire national de la délinquance a vu le jour en 2003, puis en 2014 un service statistique ministériel en charge des questions de sécurité intérieure et des statistiques de la délinquance. Dans ces conditions, que peut fournir le statisticien public pour répondre aux enjeux de l'ODD 16, et comment s'organise-t-il pour y répondre dans les années à venir ? On verra que, tout en continuant à s'appuyer sur les données administratives produites par les services de sécurité publique et le ministère de la Justice, la statistique publique va poursuivre le développement de ses enquêtes auprès de la population, notamment son enquête de victimation.

Encadré 2. Le 16^e objectif de développement durable

Parmi les dix-sept objectifs de développement durable des Nations Unies à l'horizon de 2030, le 16^e est consacré à « Promouvoir l'avènement de sociétés pacifiques et ouvertes à tous aux fins du développement durable, assurer l'accès de tous à la justice et mettre en place, à tous les niveaux, des institutions efficaces, responsables et ouvertes à tous ».

Ce 16^e objectif est décliné en douze cibles :

- 16.1 Réduire nettement, partout dans le monde, toutes les formes de violence et les taux de mortalité qui y sont associés
- 16.2 Mettre un terme à la maltraitance, à l'exploitation et à la traite, et à toutes les formes de violence et de torture dont sont victimes les enfants
- 16.3 Promouvoir l'état de droit aux niveaux national et international et donner à tous accès à la justice dans des conditions d'égalité
- 16.4 D'ici à 2030, réduire nettement les flux financiers illicites et le trafic d'armes, renforcer les activités de récupération et de restitution des biens volés et lutter contre toutes les formes de criminalité organisée
- 16.5 Réduire nettement la corruption et la pratique des pots-de-vin sous toutes leurs formes
- 16.6 Mettre en place des institutions efficaces, responsables et transparentes à tous les niveaux
- 16.7 Faire en sorte que le dynamisme, l'ouverture, la participation et la représentation à tous les niveaux caractérisent la prise de décisions
- 16.8 Élargir et renforcer la participation des pays en développement aux institutions chargées de la gouvernance au niveau mondial
- 16.9 D'ici à 2030, garantir à tous une identité juridique, notamment grâce à l'enregistrement des naissances
- 16.10 Garantir l'accès public à l'information et protéger les libertés fondamentales, conformément à la législation nationale et aux accords internationaux
- 16.a Appuyer, notamment dans le cadre de la coopération internationale, les institutions nationales chargées de renforcer, à tous les niveaux, les moyens de prévenir la violence et de lutter contre le terrorisme et la criminalité, en particulier dans les pays en développement
- 16.b Promouvoir et appliquer des lois et politiques non discriminatoires pour le développement durable

UN CINQUIÈME DES INDICATEURS PRODUITS À PARTIR DE DONNÉES ADMINISTRATIVES

Certains des indicateurs font partie, depuis très longtemps, de la production statistique issue de l'exploitation de sources administratives. C'est le cas par exemple du nombre de victimes d'homicides volontaires (ODD 16.1.1) : elles étaient 872 en 2015 en France. Il est possible de décliner les sexes et les âges, conformément à la recommandation de l'ONU. Il est également possible de chiffrer le nombre de décès liés à des conflits (ODD 16.1.2), qui est égal à zéro en France en 2015, car les attentats terroristes en sont explicitement exclus.

« Ces indicateurs, tirés des données administratives, ont l'avantage d'être peu coûteux à produire. »

Ces données sont issues des statistiques établies par les forces de sécurité (police et gendarmerie), grâce à l'analyse de leurs procédures judiciaires. De même, la statistique publique du ministère de la Justice permet de connaître la proportion de la population carcérale en instance de jugement : 27,2 % en 2015. C'est un indicateur statistique également produit de longue date et collecté par l'ONU dans sa base de données sur la criminalité et la justice pénale¹.

Enfin, grâce au système d'information sur l'emploi dans la fonction publique construit par agrégation de fichiers administratifs par le service statistique du ministère en charge de la fonction publique, il est possible de renseigner l'ODD 16.7.1 « Répartition des postes dans les institutions publiques par sexe et âge » : en 2015, les hommes occupaient 37,2 % de ces postes et les femmes 62,8 %, les jeunes de moins de 24 ans 5,6 % et les personnes de 50 ans et plus 32,5 %.

Ces indicateurs, tirés des données administratives, ont l'avantage d'être peu coûteux à produire. On peut aussi les construire chaque année. Mais ils ne représentent que quatre des vingt indicateurs statistiques du domaine. Pour aller plus loin, il faut recourir à des enquêtes auprès de la population, plus riches, mais aussi plus coûteuses à produire.

LENQUÊTE DE VICTIMATION AUPRÈS DE LA POPULATION PERMET DE MESURER PLUSIEURS INDICATEURS

Ainsi trois indicateurs sont issus des enquêtes de victimation, qui ont commencé à se développer en France à la fin du XX^e siècle, d'abord à l'initiative de chercheurs, puis en réponse à la demande européenne d'indicateurs sociaux. Depuis 2007, en réponse à une demande sociale et à un intérêt politique croissants sur les thèmes de l'insécurité et de la délinquance, l'Insee a mis en place une importante enquête annuelle de victimation, « Cadre de vie et sécurité »², qui interroge en face-à-face un échantillon représentatif de la population sur les vols et sur les agressions dont elle a été récemment victime.

Grâce à cette enquête, on connaît la proportion de la population victime de violences physiques, ou sexuelles au cours des 12 mois précédents, qui s'élève à 2,5 % des personnes âgées de 18 à 75 ans. Notons que ce chiffre ne correspond pas parfaitement à l'indicateur ODD 16.1.3, qui inclut également les violences psychologiques, pour des raisons que nous exami-

1. Voir sur le site internet de l'UNODC : <http://www.unodc.org/unodc/en/data-and-analysis/crime-and-criminal-justice.html>

2. Pour une présentation synthétique de l'enquête voir (Insee, 2019a) ; pour des résultats détaillés voir (Guedj, 2017).

« Grâce à cette enquête, on connaît la proportion de la population victime de violences physiques, ou sexuelles au cours des 12 mois précédents. [...] Cette même enquête permet de remplir un indicateur central du bien-être. »

nerons plus tard. En revanche, il prend en compte les violences subies au sein de la famille, y compris au sein du couple, grâce à un protocole de collecte spécifique qui limite la gêne éventuelle de l'enquête vis-à-vis de l'enquêteur et évite de mettre la victime de violence en danger par ses déclarations.

Cette même enquête permet de remplir un indicateur central du bien-être, qui porte non pas sur le fait d'avoir été victime, mais sur l'impression de vivre en sécurité autour de chez soi. Le nombre de personnes considérant qu'il n'y a pas de danger à se déplacer seules à pied dans leur zone de résidence, en proportion de la population totale s'élève à 89,4 % des 14 ans et plus en 2015 en France (indicateur ODD 16.1.4). Notons qu'avec ce type de questions, qui fait consensus dans les travaux sur la mesure de l'insécurité, on quitte le confort des mesures « factuelles » habituelles dans la statistique publique pour entrer dans celui, plus inhabituel, des questions subjectives.

Enfin l'enquête « Cadre de vie et sécurité » permet de renseigner un indicateur de confiance dans l'efficacité ou l'utilité des institutions, à savoir la « proportion des personnes victimes de violences au cours des 12 mois précédents ayant signalé les faits aux autorités compétentes ou recouru à d'autres mécanismes de règlement des différends officiellement reconnus ». En France, seules 19 % des personnes de 18 ans à 75 ans qui se sont déclarées victimes de violences physiques ou sexuelles en ont informé les forces de sécurité.

Telle qu'elle existe, l'enquête française de victimation permet donc de répondre à une partie des indicateurs de l'objectif 16, mais au total seuls sept indicateurs statistiques sont aujourd'hui renseignés au niveau national par la France dans ce domaine. Ils sont rassemblés sur un espace dédié du site internet de l'Insee³. À noter que, compte tenu des tailles d'échantillons et des taux de victimation, les variations annuelles que l'on peut tirer de ces enquêtes ne sont pas significatives.

DEUX NOUVEAUX THÈMES : COMPORTEMENTS DISCRIMINATOIRES...

En partie pour améliorer sa réponse aux ODD, mais aussi pour mieux satisfaire une demande sociale nationale croissante, les responsables de l'enquête nationale de victimation ont décidé d'en développer le questionnaire dans deux directions à partir de 2018 : les comportements discriminatoires et la corruption.

Concernant les comportements discriminatoires, l'indicateur ODD 16.b.1 est rédigé ainsi « Proportion de la population ayant déclaré avoir personnellement fait l'objet de discrimination ou de harcèlement au cours des 12 mois précédents pour des motifs interdits par le droit international des droits de l'homme ». Il s'agit donc d'une conception assez large des comportements de nature discriminatoire, qui va au-delà des strictes discriminations entendues comme le refus d'un droit pour des raisons illégales.

Jusqu'à présent, l'enquête permettait de repérer les motifs discriminatoires liés notamment à la race, à l'origine, au genre ou à l'orientation sexuelle, pour les seules personnes qui

3. Voir (Insee, 2019b).

s'étaient déclarées victimes de violences, de menaces ou d'injures. Pour aller au-delà, en s'inspirant des enquêtes déjà réalisées sur le thème des discriminations et après discussion avec les spécialistes du domaine⁴, les concepteurs de l'enquête ont ajouté un nouveau module sur les discriminations et les attitudes blessantes ou vexantes (**encadré 3**).

La collecte de cette enquête rénovée s'est déroulée en 2018. Des résultats ont été publiés dans une étude du SSMSI sur le sexisme⁵, d'autres seront publiés dans le rapport annuel sur la lutte contre le racisme, l'antisémitisme et la xénophobie de la Commission nationale consultative des droits de l'homme (CNCDH)⁶. Du fait de son caractère novateur, des ajustements seront peut-être nécessaires avant de parvenir à une mesure satisfaisante du phénomène.

Encadré 3. La mesure des discriminations et des attitudes blessantes ou vexantes dans l'enquête de victimation à partir de 2018

Le nouveau module du questionnaire commence ainsi :

« Nous allons maintenant aborder les comportements dont vous avez pu être victime en raison, par exemple, de votre couleur de peau, de votre religion, de votre orientation sexuelle, de votre âge ou du fait d'être un homme ou une femme. Il peut s'agir :

1. De violences, de menaces ou d'injures.

2. De discriminations, c'est-à-dire de traitements défavorables comme le refus de vous accorder un emploi, un logement, un prêt ou tout autre bien ou service.

Par exemple, une personne ayant un nom d'origine étrangère qui ne reçoit pas de réponse lorsqu'elle postule pour un emploi ou une personne qui a la peau noire à laquelle on dit qu'un appartement est déjà loué alors que c'est faux ou encore une personne à laquelle on refuse un crédit parce qu'elle vient d'un quartier défavorisé.

3. Il peut aussi s'agir d'attitudes ou de propos vexants ou blessants.

Par exemple, des propos comme « Les femmes ne comprennent rien à la mécanique » ou « Tous les noirs ont le sens du rythme » ou des questions gênantes : « De quel pays viens-tu ? » ou « C'est vrai que tu es gay/lesbienne ? », il peut aussi s'agir d'attitude de rejet ou d'évitement, comme un refus de serrer la main, de monter dans l'ascenseur avec vous ou le fait de changer de trottoir.

Ces comportements peuvent se produire sur le lieu de travail ou d'études, à l'entrée d'une boîte de nuit, lors d'un contrôle de police ou autres ».

Suivent des questions permettant aux personnes de signaler si elles ont été victimes au cours des deux années précédant l'enquête, puis pour celles qui répondent positivement, quelques questions sur la nature du comportement discriminatoire, la cause de ce comportement ainsi que le déroulement et les circonstances des faits.

4. Notamment les enquêtes réalisées par l'Insee et l'Ined « Trajectoires et origines » de 2008 (Insee, 2019d) et par l'Insee « Histoires de vie » en 2003 (Insee, 2019c).

5. Voir (Guedj, 2019)

6. Parallèlement à ce rapport qui a été rendu public en mars 2019, le SSMSI a publié un Interstats Analyse sur les atteintes à caractère raciste, xénophobe ou antireligieux en 2018 (SSMSI, 2019).

S'agissant de corruption, celle-ci constitue un sujet entièrement nouveau pour la statistique publique française. L'indicateur ODD 16.5.1 est la « Proportion de personnes ayant eu, au moins une fois, affaire à un agent public auquel elles ont versé un pot-de-vin ou qui leur a demandé un pot-de-vin au cours des 12 mois précédents ». Après avoir consulté des spécialistes du domaine, et observé les expériences étrangères ou menées par des ONG spécialisées, il a été décidé de le mesurer grâce à un nouveau module spécifique de l'enquête « Cadre de vie et sécurité », qui commence ainsi :

« Parfois, certaines personnes se voient demander des dessous-de-table, des pots-de-vin, des bakchichs, des cadeaux, pour obtenir un service, une place en crèche, un logement social, un document administratif, une autorisation, un permis, un rendez-vous médical, etc. Au cours des dernières années, dans le cadre de vos affaires personnelles et privées, ou dans le cadre de votre travail, est-ce que quelqu'un a demandé ou attendu de votre part le versement d'un pot-de-vin ou un cadeau pour fournir ses produits ou ses services ? »

Comme dans le sujet précédent, la difficulté consiste à faire comprendre à la personne interrogée ce dont nous souhaitons parler. C'est d'autant plus essentiel ici qu'une grande majorité des répondants n'ont jamais connu ce phénomène, ou du moins ne l'ont jamais interprété comme tel. Nous avons ainsi évité d'utiliser le mot « corruption » dans l'intitulé ou même dans le texte de la question, car il risquait d'être mal compris ou de susciter rejet ou incompréhension.

Après cette question « filtre » suit une petite série de questions visant à cerner le domaine administratif ou économique concerné.

Comme pour les comportements de nature discriminatoire, les données ont commencé à être expertisées mais aucune étude n'a pour le moment été publiée. Du fait de la relative rareté du phénomène⁷, il faudra probablement disposer d'une année d'enquête supplémentaire pour produire des statistiques suffisamment robustes.

DES INDICATEURS NON PRODUITS EN FRANCE, FAUTE DE CONSENSUS INTERNATIONAL SUR LES DÉFINITIONS...

Malgré ces évolutions programmées, certains indicateurs de développement durable de l'objectif 16 ne seront toujours pas renseignés par la France dans les années qui viennent, et ceci pour des raisons variées. Cette situation n'est pas anormale ni propre à la France. D'ailleurs, parmi les indicateurs pour lesquels la statistique publique française n'a à ce jour pas de solution, plusieurs d'entre eux font encore l'objet de discussions conceptuelles au niveau international⁸, notamment dans le groupe dit de Praia sur la mesure de la gouvernance, et pour la majorité des autres, le mode de collecte des informations pose problème à un grand nombre de pays⁹.

7. Notons que l'institut de statistique italien (Istat) a réalisé une enquête selon une méthode très proche en 2015-2016. Selon cette enquête, la part des ménages italiens confrontés à des « demandes d'argent, de faveurs, de cadeaux ou autres en échange de services ou de facilitations de transactions » au cours de la dernière année est de 1,2 %. Cette part monte à 2,7 % lorsqu'elle est mesurée sur trois années et à 7,9 % au cours de la vie. Les secteurs les plus concernés par ces comportements de corruption sont la recherche d'emploi, puis la justice, les services sociaux et la santé (Istat, 2017).

8. Indicateurs dits en « Tier III ».

9. Indicateurs dits en « Tier II ». Le nombre d'indicateurs statistiques relatifs à l'ODD 16 est particulièrement élevé (23), ce qui traduit la complexité des discussions internationales sur le domaine. Voir notamment sur le sujet (Cling, Razafindrakoto et Roubaud, 2016).

« Parfois, la raison en est qu'il n'existe pas encore de définition opératoire reconnue internationalement. C'est le cas des violences psychologiques. »

Parfois, la raison en est qu'il n'existe pas encore de définition opératoire reconnue internationalement. C'est le cas des violences psychologiques, mentionnées dans les indicateurs ODD 16.1.3 et ODD 16.2.1¹⁰. Quelques enquêtes, notamment l'enquête « Violence et rapports de genres » de l'INED et, pendant quelques années, un module

spécifique de « Cadre de vie et sécurité »¹¹) ont approché à travers des questions sur des « situations typiques », les violences psychologiques au sein du couple (par exemple : « Votre conjoint a-t-il exigé de savoir avec qui et où vous étiez ? », « Votre conjoint a-t-il dévalorisé ce que vous faisiez ? »), ou dans les relations de travail («... avez-vous eu des critiques répétées et injustifiées concernant votre travail, vous a-t-on rabaissé ou humilié... », « dans votre travail, avez-vous été tenu à l'écart, empêché de communiquer avec les autres ? »)¹². Des mesures sont donc possibles, et ont fait l'objet d'études et d'interprétations, mais à défaut d'un cadre international uniformisé, la construction d'indicateurs fiables et comparables dans le temps et entre pays est pour le moment impossible.

... OU SUR LES TECHNIQUES STATISTIQUES ADÉQUATES

Le plus souvent, ce sont des questions de technique statistique qui se heurtent au bon renseignement des indicateurs. Ces problèmes sont de nature très variée selon les indicateurs. L'indicateur relatif à la violence envers les enfants 13.2.1 « Proportion d'enfants âgés de 1 à 17 ans ayant subi des châtiments corporels ou des agressions psychologiques infligés par une personne s'occupant d'eux au cours du mois précédent » pose un problème de technique de collecte : pour des raisons de droit, on peut difficilement interroger un échantillon d'enfants sans l'accord de leurs parents, surtout sur un tel sujet. L'interrogation des parents introduit bien entendu un biais de réponse important, puisqu'ils sont le plus souvent les auteurs de ces violences, et les données administratives sous-estimeront systématiquement l'ampleur du phénomène. Seule une enquête rétrospective auprès des très jeunes adultes pourrait permettre d'approcher la cible, à l'image de ce qui est proposé dans l'indicateur ODD 16.2.3 « Proportion de jeunes femmes et hommes de 18 à 29 ans ayant été victimes de violences sexuelles avant l'âge de 18 ans¹³ ».

L'indicateur ODD 16.5.2 portant sur la corruption dont sont victimes les entreprises « Proportion d'entreprises ayant eu au moins une fois affaire à un agent public auquel elles ont versé un pot-de-vin ou qui leur a demandé un pot-de-vin au cours des 12 mois précédents » pose également un problème de base de sondage pour une éventuelle enquête : qui enquêter au sein de l'entreprise ? Tout d'abord, il y a un risque que des responsables économiques souhaitent cacher ou minimiser un événement qui a pu être jugé nécessaire à la vie de leur établissement. Ensuite, les actes de corruption passive ont pu s'adresser à plusieurs

10. Bien que ces deux indicateurs soient classés en Tier II, il est indiqué explicitement dans les métadonnées présentes sur le site <https://unstats.un.org/sdgs/metadata/> : « Il n'y a pas encore de consensus au niveau international sur la définition précise de la notion de violence psychologique et il n'existe pas encore de méthodologie généralement bien établie pour mesurer la violence psychologique. »

11. Voir les résultats publiés dans (Burricand, Jamet, 2016).

12. Voir (Hamel, 2014).

13. Sous réserve d'expertise, le module « Violences avant l'année N-2 », qui a figuré plusieurs années dans le questionnaire de l'enquête CVS, ainsi que l'enquête Virage de l'INED pourraient fournir des ordres de grandeur dans ce domaine.

personnes dans l'entreprise, et l'information a pu ne pas circuler jusqu'à celui qui, le plus souvent, répond aux questionnaires statistiques adressés à l'entreprise.

① UNE EXPÉRIMENTATION EN FRANCE SUR LA MESURE DE LA CORRUPTION DANS LE MONDE PROFESSIONNEL

Sur la suggestion de spécialistes de la délinquance économique, la France expérimente à partir de 2018 le dispositif suivant, dans le cadre de son enquête auprès des ménages : en plus de demander aux personnes si elles ont réclamé un dessous-de-table dans

le cadre de leur vie privée ou familiale, on les interroge aussi sur le fait de savoir si elles ont été victimes de cette forme de corruption dans leur vie professionnelle. Ainsi, la statistique publique va chercher à mesurer une pratique relative à la vie des entreprises à travers une enquête auprès des ménages. Suivant la même logique, les personnes sont également interrogées sur les situations où, dans le cadre de leur

“ La statistique publique va chercher à mesurer une pratique relative à la vie des entreprises à travers une enquête auprès des ménages. ”

travail, on leur a proposé un pot-de-vin, par exemple pour obtenir un marché ou une commande. Les témoignages de corruption passive sont donc complétés par des témoignages de corruption active.

À défaut de répondre exactement à la question de la proportion d'entreprises concernées, il sera donc possible de disposer d'une indication de la proportion des personnes en activité qui ont fait face à des pratiques de corruption, sous réserve que le dispositif d'enquête ait bien fonctionné.

① TRAFICS ET FRAUDE : DES INDICATEURS ENCORE PLUS DIFFICILES À MESURER

Nous passerons rapidement sur les sujets pour lesquels, pour le moment, on n'entrevoit pas de solutions techniques, que ce soit par des enquêtes spécifiques, par l'usage de sources administratives ou *via* des synthèses de sources : le chiffrage des flux financiers internationaux illicites, la traçabilité des armes à feu ou le nombre de victimes de traite des êtres humains. Dans ces trois domaines, on est loin, au niveau national mais aussi international, de disposer d'outils fiables permettant des chiffrages comparables entre pays et dont on pourrait tracer des dynamiques temporelles¹⁴.

① SATISFACTION VIS-À-VIS DES INSTITUTIONS : PAS DE CONSENSUS DANS LES PAYS DÉVELOPPÉS...

Avant de conclure, arrêtons-nous sur un autre type d'indicateurs pour le moment non produits par la statistique publique française que sont les indicateurs de satisfaction vis-à-vis des institutions : « Proportion de la population dont la dernière expérience avec les services publics a été satisfaisante » (ODD 16.6.2) et « Proportion de la population qui estime que la prise de décisions est ouverte et réactive » (ODD 16.7.2).

14. Deux de ces indicateurs sont officiellement encore classés en « Tier III », et le troisième en « Tier II ».

La nature subjective de ces deux indicateurs n'est pas ici en cause, car la statistique publique française s'est ouverte depuis une dizaine d'années à la mesure des questions subjectives dites « d'opinion ». Il est vrai que celles-ci, jugées peu fiables, en avaient été longtemps bannies. Certes les soldes d'opinion sur des grandeurs économiques telles que les perspectives économiques, les baromètres conjoncturels, la satisfaction des utilisateurs de certains services publics sur la qualité de l'accueil ou sur celle du traitement de leurs demandes étaient depuis longtemps régulièrement utilisés. Mais désormais, à la suite du rapport Stiglitz-Sen-Fitoussi sur « la mesure de la performance économique et du progrès social » de 2009, plusieurs opérations statistiques se sont enrichies d'outils de mesure sur des aspects plus subjectifs comme le bien-être global des individus ou la satisfaction associée à certaines activités.

On pourrait croire qu'il ne s'agit pas non plus, à première vue, d'une difficulté technique à les construire à partir d'une enquête auprès de la population. Ces indicateurs posent même moins de problèmes de technique d'enquête que les questions, par exemple, sur les violences sexuelles, et ne nécessitent pas forcément des tailles d'échantillon élevées.

Cependant, les autorités statistiques nationales expriment de nombreuses réserves au sein des pays développés sur le périmètre des indicateurs et leur méthodologie de construction, alors même que certaines institutions d'étude ou de recherche, privées ou publiques, sont plus allantes.

Ainsi l'OCDE, qui a en particulier largement exploré ce domaine dans son étude « How's Life ? 2017 Measuring Well-being »¹⁵, souligne toutefois dans ce même rapport la difficulté à mesurer la gouvernance, du fait de son caractère multidimensionnel, d'une absence pour l'heure de consensus au niveau international sur le périmètre conceptuel des statistiques de la gouvernance mais aussi d'une méthodologie encore insuffisamment robuste et éprouvée. C'est notamment le cas pour ce qui concerne la mesure de la confiance dans les institutions, alors que la mesure de la confiance entre les individus, à partir des enquêtes ménages, apparaît plus consensuelle et éprouvée.

De la même façon, des réserves ont été exprimées au niveau européen. Eurostat, l'Office statistique de la Commission européenne, avait en effet expérimenté en 2013 l'ajout d'un module de questions sur ce thème dans l'enquête européenne SILC sur les ressources et conditions de vie des ménages. Ce module intégrait quatre questions sur la confiance : une sur la confiance dans les autres et trois sur celle dans les institutions (« sur une échelle allant de 0 à 10, quel degré de confiance avez-vous dans les institutions politiques et les responsables politiques et publics, quel degré de confiance avez-vous dans la justice et quel degré de confiance avez-vous dans la police et la gendarmerie »). Les résultats de ce module expérimental n'avaient pas convaincu les instituts nationaux de statistique, notamment pour la mesure de la confiance dans les institutions et plus spécifiquement encore pour celle de la confiance dans le système politique. En particulier, la confiance dans les hommes politiques est apparue très dépendante du cycle politique qui n'est évidemment pas le même selon les pays, d'où une difficulté à interpréter les résultats. Au final, quand il s'est agi de décider s'il fallait reprendre ces questions sur la confiance dans le module SILC en 2018, le groupe de travail en charge de la conception de ce module a préféré ne pas

15. On y apprend notamment, dans une optique très proche de l'ODD 16.7.2, qu'en France, seuls 10 % de la population a le sentiment d'avoir son mot à dire dans ce que fait le gouvernement, ce qui est bien inférieur à la moyenne de l'OCDE (33 %) (<http://www.oecd.org/statistics/how-s-life-23089679.htm>). C'est dans l'enquête internationale sur l'usage de l'information par les adultes (PIAAC) que ces informations ont été recueillies (Insee, 2019e).

le faire. Un point important à approfondir sur ces sujets, à partir d'une analyse empirique solide étayée par des résultats d'enquêtes sur le terrain, est la comparabilité internationale des statistiques produites ainsi que la comparabilité dans le temps des mesures pour un même pays.

... ALORS QUE LA MESURE DE LA GOUVERNANCE SE DÉVELOPPE DANS LES AUTRES PAYS

Ces réticences des autorités nationales dans les pays développés ne sont pas partagées dans le monde entier puisque, par exemple, plusieurs pays d'Afrique ont confié à leurs instituts nationaux de statistique la responsabilité de mener des enquêtes sans complaisance sur leur gouvernance, dont le questionnaire a été conçu par des chercheurs de l'Institut de recherche pour le développement (IRD) et de l'UMR DIAL (Développement, Institutions et Mondialisation) de l'Université de Paris-Dauphine.

“ Les travaux menés jusqu'à présent ont prouvé non seulement leur faisabilité technique, mais aussi leur pertinence et leurs capacités à mettre en lumière les qualités de gouvernance des pays concernés. ”

Ces modules d'enquête rassemblent des informations de victimation, de sentiment d'insécurité, mais aussi de témoignages sur la corruption, sur la qualité des relations avec les administrations et la confiance dans les autorités. Les travaux menés jusqu'à présent ont prouvé non seulement leur faisabilité technique, mais aussi leur pertinence et leurs capacités à mettre en lumière les qualités de gouvernance des pays concernés (Razafindrakoto et Roubaud, 2015).

DES INDICATEURS QUI S'INSCRIVENT MAL DANS LES TRAVAUX D'ÉVALUATION PLUS SECTORIELS DE LA FRANCE

En France, outre les difficultés techniques qui viennent d'être citées à construire ces indicateurs, ceux-ci constituent, à eux seuls, des indicateurs globaux d'évaluation de l'action publique, politiquement sensibles, qui ne s'inscrivent pas dans la tradition des travaux d'évaluation de la statistique publique et qui portent généralement sur une politique spécifique ou sur un champ particulier de l'action publique. Jusqu'à présent, cela n'a pas permis la mise en place d'un dispositif d'évaluation global des institutions. Cet état de fait pourra évoluer dans les années qui viennent. Notamment le groupe de travail du Conseil national de l'information statistique (Cnis), instance de concertation entre producteurs et utilisateurs de la statistique publique, chargé depuis l'été 2017 de sélectionner les indicateurs pertinents reflétant les efforts de la France pour atteindre les cibles associées aux ODD, pourra émettre des recommandations pour faire évoluer le dispositif statistique actuel associé à l'objectif 16.

En conclusion, si elle veut répondre aux exigences des ODD dans ce domaine, la statistique publique française va devoir investir le champ nouveau de la statistique sur la Paix, la justice et les institutions efficaces. Elle devra continuer à enrichir ses outils, et éventuellement faire évoluer certains de ses axes stratégiques, pour augmenter la production d'indicateurs. Ces limites ne sont pas propres à la France, et l'enrichissement des données françaises dépendra aussi des discussions multilatérales encore en cours sur les concepts et les méthodes qui restent à préciser pour certains indicateurs.

BEAUCHEMIN, Cris, HAMEL, Christelle et SIMON, Patrick, 2016. *Trajectoires et origines – Enquête sur la diversité des populations en France*. Ined, Collection Grandes Enquêtes. ISBN 978-2-7332-8004-1.

BURRICAND, Carine et JAMET, Lucile, 2016. Atteintes psychologiques et agressions verbales entre conjoints. In : *Insee Première* [en ligne]. 4 juillet 2016. N°1607. [Consulté le 6 juin 2019]. Disponible à l'adresse : <https://www.insee.fr/fr/statistiques/fichier/version-html/2019028/ip1607.pdf>

CLING, Jean-Pierre, RAZAFINDRAKOTO, Mireille et ROUBAUD, François, 2016. L'ODD 16 sur la gouvernance et sa mesure : l'Afrique en tête. In : *Afrique contemporaine* [en ligne]. Décembre 2016. N°258, pp. 73-93. [Consulté le 6 juin 2019]. Disponible à l'adresse : <https://www.cairn.info/revue-afrique-contemporaine-2016-2-page-73.htm>

CLING, Jean-Pierre, RAZAFINDRAKOTO, Mireille et ROUBAUD, François, 2018. SDG16 on Governance and its measurement: Africa in the Lead. In : *Document de travail DIAL* [en ligne]. Janvier 2018. N°2018-02. [Consulté le 6 juin 2019]. Disponible à l'adresse : http://www.dial.ird.fr/media/ird-sites-d-unites-de-recherche/dial/documents/publications/doc_travail/2018/2018-02

GUEDJ, Hélène, 2017. Rapport d'enquête Cadre de vie et sécurité 2017. In : *site du SSMSI Interstats* [en ligne]. 7 décembre 2017. [Consulté le 6 juin 2019]. Disponible à l'adresse : <https://www.interieur.gouv.fr/content/download/99046/778274/fle/Rapport%20d'enqu%C3%AAt%20CVS%202017.pdf>

GUEDJ, Hélène, 2019. Les victimes du sexisme en France – Approche croisée sur 2017 à partir des procédures enregistrées par les forces de sécurité et l'enquête « Cadre de vie et sécurité ». In : *Interstats Analyses* [en ligne]. 8 mars 2019. N° 19. [Consulté le 6 juin 2019]. Disponible à l'adresse : <https://www.interieur.gouv.fr/content/download/115264/923400/fle/IA19.pdf>

HAMEL, Christelle, 2014. *Enquête Violence et rapports de genre : contextes et conséquences des violences subies par les femmes et par les hommes* [en ligne]. Juin 2014. Ined, Collection documents de travail, N°212. [Consulté le 6 juin 2019]. Disponible à l'adresse : https://virage.site.ined.fr/fichier/s_rubrique/20838/212_genre_violence.fr.pdf

INSEE, 2019a. Enquête de victimation – cadre de vie et sécurité. In : *site de l'Insee* [en ligne]. 5 juin 2019. [Consulté le 6 juin 2019]. Disponible à l'adresse : <https://www.insee.fr/fr/metadonnees/source/serie/s1278/>

INSEE, 2019b. Indicateurs pour le suivi national des objectifs de développement durable – 17 objectifs de développement durable. In : *site de l'Insee* [en ligne]. 29 avril 2019. [Consulté le 6 juin 2019]. Disponible à l'adresse : <https://www.insee.fr/fr/statistiques/2654964>

INSEE, 2019c. Enquête Histoire de vie (HDV). In : *site de l'Insee* [en ligne]. 5 juin 2019. [Consulté le 6 juin 2019]. Disponible à l'adresse : <https://www.insee.fr/fr/metadonnees/source/serie/s1246>

INSEE, 2019d. Enquête Trajectoires et Origines en 2008 (TeO 2008). In : *site de l'Insee* [en ligne]. 5 juin 2019. [Consulté le 6 juin 2019]. Disponible à l'adresse : <https://www.insee.fr/fr/metadonnees/source/s1276>

INSEE, 2019e. Programme pour l'évaluation internationale des compétences des adultes (PIAAC). In : *site de l'Insee* [en ligne]. 5 juin 2019. [Consulté le 6 juin 2019]. Disponible à l'adresse : <https://www.insee.fr/fr/metadonnees/source/s1257>

ISTAT, 2017. Corruption in Italy: the citizen's point of view. In : *site d'Istat* [en ligne]. 12 octobre 2017. [Consulté le 6 juin 2019]. Disponible à l'adresse : https://www.istat.it/it/files//2017/10/EN_Corruption_in_Italy.pdf

RAZAFINDRAKOTO, Mireille et ROUBAUD, François, 2015. Les modules Gouvernance, Paix et Sécurité dans un cadre harmonisé au niveau de l'Afrique (GPS-SHaSA) : développement d'une méthodologie d'enquête statistique innovante. In : *Statéco* [en ligne]. N°109, pp. 103-133. [Consulté le 6 juin 2019]. Disponible à l'adresse : <http://www.epsilon.insee.fr/jspui/bitstream/1/39255/4/Stateco109.pdf>

RAZAFINDRAKOTO, Mireille et ROUBAUD, François, 2018. Responding to the SDG16 Measurement Challenge : The Governance, Peace and Security Survey Modules in Africa. In : *Global Policy*. 16 avril 2018. University of Durham and John Wiley & Sons, Ltd.

SSMSI, 2019. Les atteintes à caractère raciste, xénophobe ou antireligieux en 2018. In : *Interstats Analyses* [en ligne]. 23 avril 2019. N° 20. [Consulté le 6 juin 2019]. Disponible à l'adresse : <https://www.interieur.gouv.fr/content/download/115882/928737/fle/IA20.pdf>

RMÉS, LE RÉFÉRENTIEL DE MÉTADONNÉES STATISTIQUES DE L'INSEE

Dominique Bonnans*

Permettre de comprendre ce que recouvre un résultat statistique afin de faciliter son interprétation ou sa réutilisation, voilà tout l'enjeu des métadonnées statistiques. Elles sont de multiples natures : définitions, nomenclatures, caractéristiques de la source mobilisée pour produire le résultat, etc. Pour rendre ces informations accessibles, le référentiel de métadonnées statistiques de l'Insee, RMÉS, les structure, les centralise, assurant la cohérence de l'information grâce à des règles d'identification. L'adoption de standards internationaux pour décrire les métadonnées facilite leur usage, automatique ou manuel, et les échanges avec d'autres organismes. L'environnement permettant de communiquer avec le référentiel a également été mis en place : des services sont offerts tout au long du cycle de vie d'une opération statistique, aussi bien pour alimenter le référentiel que pour accéder à son contenu. L'usage d'RMÉS va plus loin qu'une simple réutilisation : certaines applications clientes peuvent mobiliser les métadonnées pour produire automatiquement des composants du processus statistique, par exemple pour spécifier des questionnaires, voire à terme décrire des contrôles ou des produits de diffusion. Transversal et structurant pour la production statistique, RMÉS apparaît comme un potentiel vecteur de la coordination au sein du Service Statistique Public.

 *Understanding what a statistical result covers in order to facilitate its interpretation or reuse is what statistical metadata is all about. They are of many kinds: definitions, classifications, characteristics of the source used to produce the result, etc. To make this information accessible, Insee's statistical metadata repository, RMÉS, structures it and centralizes it, ensuring the consistency of the information through identification rules. The implementation of international standards for describing metadata facilitates their use, automatically or manually. This also allows sharing with other entities. The environment for communicating with the repository has also been set up: services are offered throughout the life cycle of a statistical operation, both to supply the repository and to access its content. The use of RMÉS goes beyond simple reuse: some client applications can rely on metadata to automatically produce components of the statistical process, for example to specify questionnaires, or even eventually describe controls or dissemination products. Cross-functional and structuring for statistical production, RMÉS appears as a potential vector for coordination within the Public Statistical Service.*

* Cheffe de l'Unité Qualité, Insee,
dominique.bonnans@insee.fr

En mettant en place dès les années 80 un Dispositif de Documentation Structurée (DDS), l'Insee soulignait l'importance accordée à toutes les informations permettant de mieux comprendre les statistiques produites. Sans information sur les concepts, nomenclatures, variables, listes de codes, les résultats statistiques ne pourraient être correctement utilisés et interprétés, ni appariés à d'autres données : on ne saurait tout simplement pas ce qu'ils mesurent. Ces informations sont des données... sur les données : des métadonnées. Encore faut-il les élaborer et les mettre à disposition de façon efficace : cet article présente comment y parvient RMÉS, le Référentiel de Métadonnées Statistiques de l'Insee, qui s'est substitué à DDS en 2018.

DOCUMENTER POUR FAIRE COMPRENDRE LES CHIFFRES

La devise qui accompagne son logo « *mesurer pour comprendre* » traduit l'engagement de l'Insee à produire des données utiles pour éclairer le débat public. Dans le cadre de cet engagement, l'Insee met en place des enquêtes, mobilise des sources administratives, combine différents dispositifs d'information. L'exploitation de ces sources permet de produire des chiffres qui sont diffusés dans des publications, mis en ligne sur le site insee.fr, repris dans la presse ou réutilisés par des chercheurs ou d'autres utilisateurs.

Mais ces chiffres, pris isolément, ne sont porteurs d'aucune information. Pour devenir « données », ils doivent être accompagnés d'un ensemble d'informations qui permettent de les comprendre, de les interpréter correctement. Très souvent, ces informations supplémentaires passent inaperçues, car elles viennent émailler les commentaires qui accom-

“ *Le commentaire livre les éléments indispensables pour décoder.[...] L'utilisateur dispose ainsi d'éléments lui permettant de savoir ce à quoi les chiffres se réfèrent.* ”

pagnent les données analysées dans une publication ou reprises dans un article de journal. Prenons l'exemple du bilan de la saison touristique d'été 2018, présenté dans la publication *Insee Focus* n°134 de novembre 2018 : « *Durant les six mois de la saison d'été 2018, les hébergements collectifs touristiques de France métropolitaine ont enregistré 311 millions de nuitées, soit une hausse de 1,3 % par rapport à la saison estivale de 2017.* » Le commentaire livre les éléments indispensables pour décoder en première lecture ce que recouvre la hausse de 1,3 %. Il s'agit là d'une comparaison par rapport à la même

période de l'année précédente ; cette période est la « saison d'été », concept dont la définition est précisée dans l'article (les mois d'avril à septembre). Cette comparaison porte sur le nombre de « nuitées », concept également précisé par l'article. Le territoire concerné est la France métropolitaine. Le périmètre est celui des hébergements collectifs touristiques, concept explicité dans la rubrique « Définitions » sur insee.fr, et qui correspond à plusieurs groupes d'activités de la division 55 (Hébergement) de la nomenclature d'activités française, également consultable sur insee.fr.

L'utilisateur dispose ainsi d'éléments lui permettant de savoir ce à quoi les chiffres (311 millions, 1,3 %) se réfèrent, les catégories qu'ils utilisent, et peut ainsi comprendre ce que traduit la hausse de 1,3 %. S'il s'agit d'un utilisateur averti, il peut en outre trouver sur insee.fr plus d'informations sur les enquêtes de fréquentation touristique qui ont permis de produire ce résultat et qui sont référencées dans les sources de l'article. Il peut par exemple connaître le nombre et le type d'unités enquêtées, le nombre et le type de relances, le taux de

Encadré 1. Un exemple de métadonnées associées à l'enquête TIC* 2016

L'enquête sur l'usage de l'informatique, des technologies de la communication et le commerce électronique...

... permet de mesurer des concepts...

Effectif annuel moyen (en nombre de personnes occupées)

Concerner les personnels salariés et non-salariés de l'entreprise, y compris le dirigeant, le gérant, les associés non salariés et les apprentis, et n'inclut pas les intérimaires, ni les stagiaires, ni les saisonniers, ni le personnel rattaché à d'autres entreprises d'un même groupe. Il s'agit du nombre de personnes physiques.

... en utilisant des variables...

Effectif annuel moyen
durant le dernier exercice
comptable (en nombre
de personnes occupées) (EFF)

... dans un questionnaire, ... s'appuie sur des nomenclatures...



Couverture sectorielle

Les entreprises appartiennent aux secteurs suivants de la NACE rév. 2 :

- l'industrie manufacturière (section C)
- la production et la distribution d'électricité, de gaz, de vapeur et d'air conditionné (section D)
- la production et la distribution d'eau, l'assainissement, la gestion des déchets et la dépollution (section E)
- la construction (section F)
- le commerce, la réparation d'automobiles et de motocycles (section G)
- les transports et l'entreposage (section H)
- l'hébergement et la restauration (section I)
- l'information et la communication (section J)
- les activités immobilières (section L)
- les activités spécialisées, scientifiques et techniques hors activités vétérinaires (divisions 69 à 74)
- les activités de services administratifs et de soutien (section N)
- la réparation d'ordinateurs et d'équipements de communication (groupe 951).

... et alimente un rapport qualité.

Plan de sondage

L'échantillon de l'enquête TIC-TPE est tiré dans une base de sondage construite à partir du répertoire Sirius.

La méthode d'échantillonnage est un sondage aléatoire simple stratifié. La stratification correspond au croisement entre 31 secteurs d'activité, 3 tranches d'effectif en nombre de personnes occupées et 4 tranches de chiffres d'affaires (soit 372 strates).

Le nombre d'entreprises à échantillonner par strate provient d'une allocation proportionnelle au nombre d'unités (avec contraintes de précisions locales), modifiée pour prendre en compte la dispersion des montants de vente web selon les secteurs. En effet, l'allocation proportionnelle au nombre d'unités par strate est a priori plus adaptée aux paramètres d'intérêt de type proportion. Pour améliorer l'estimation des paramètres de type montant, la dispersion de la précédente enquête TIC-TPE a été prise en compte en appliquant un coefficient multipli-

cateur à l'allocation proportionnelle sur certains secteurs, afin que celle-ci soit plus importante sur les regroupements d'activités ayant une forte dispersion de montants de ventes web.

La part de la population de la strate à tirage exhaustif est :

- dans l'échantillon : 3,62 % (= 402 / 11 100)
- dans la population totale : 0,02 % (= 402 / 2 292 643)

La distribution des taux de sondage par strate est la suivante :

- 1^{er} quartile = 1 %
- médiane = 5 %
- 3^e quartile = 30 %

Taille de l'échantillon : 11 100 unités.

*TIC : technologie de l'information et de la communication.

réponse, les méthodes de redressement, les produits de diffusion, etc. Dans notre exemple, les éléments permettant de comprendre les chiffres se présentent sous la forme de commentaires explicatifs au sein d'une publication. Mais on peut aller plus loin : les informations qui définissent et décrivent des données, comme les concepts (ex : nuitée) ou les nomenclatures (ex : nomenclature d'activités française) sont en elles-mêmes des données qu'il faut pouvoir isoler et gérer en tant que telles. On les nomme « métadonnées », et elles interviennent dans de nombreux aspects du processus de production statistique.

❶ DES MÉTADONNÉES STATISTIQUES TRÈS DIVERSES...

Les exemples que l'on vient de citer sont des métadonnées importantes pour la statistique, mais il en existe beaucoup d'autres, de natures variées. En raison de cette hétérogénéité il n'est pas toujours facile de bien appréhender la notion. Pour y voir plus clair on peut néanmoins esquisser quelques façons de les caractériser.

Les métadonnées se distinguent par le type d'information qu'elles apportent (**encadré 1**) : certaines définissent rigoureusement ce à quoi se réfère un chiffre (nomenclatures, concepts, liste de codes...), d'autres visent à informer sur la qualité des résultats statistiques (méthodologie employée, précision d'une série de données, taux de réponse, technique de désaisonnalisation), d'autres enfin explicitent le processus de production (dates de collecte, calendrier et fréquence de diffusion). Ensemble, selon des angles différents, elles permettent à l'utilisateur de savoir précisément de quoi on parle, conformément au principe de clarté du code de bonnes pratiques de la statistique européenne. En revanche, la description des procédures ou des consignes de gestion ne fait pas partie des métadonnées statistiques. En effet, elle n'est pas indispensable à la compréhension des résultats, même si cette documentation « métier » est très utile aux gestionnaires ou aux chargés d'études.

La portée de ces métadonnées est variable. Certaines sont transverses, susceptibles d'être utilisées par plusieurs producteurs de données et ont ainsi une existence propre, indépendante d'une opération statistique donnée : c'est le cas de la plupart des nomenclatures (Guibert *et alii*, 1971), partagées par nature, et de plusieurs concepts qui garantissent la cohérence de la production statistique (par exemple l'activité principale exercée, l'âge révolu, l'unité urbaine, etc.). D'autres sont au contraire spécifiques à une opération : sa description, le numéro de visa ou le questionnaire s'il s'agit d'une enquête, le cas échéant son plan de sondage, ainsi que ses variables et listes de codes. À mi-chemin entre transverses et spécifiques, certaines peuvent néanmoins être mutualisées par plusieurs opérations statistiques : c'est le cas justement de quelques variables ou listes de codes mis en commun entre plusieurs enquêtes.

Les métadonnées d'une opération statistique se distinguent aussi par le moment où elles sont recueillies. Car leur « collecte » ne s'effectue pas en une seule étape : au contraire, elles sont renseignées par le producteur tout au long du cycle de vie d'une opération.

- ❶ Au moment de la définition des besoins, le producteur peut décrire les objectifs de l'opération statistique envisagée, sa place dans le système d'information et ses principales caractéristiques, comme la couverture sectorielle, la date et le mode de collecte ou encore la zone géographique de référence. Il peut, le cas échéant, demander la création de concepts lors de cette étape d'identification des besoins.
- ❶ Dans la phase de conception puis d'élaboration des outils de collecte, le producteur définit les variables qu'il souhaite collecter (le chiffre d'affaires, l'effectif salarié...) et les modalités qu'elles peuvent prendre (listes de codes). Lors de cette étape, le concepteur peut réutiliser des concepts ou des nomenclatures existants.

- ❶ Lors des phases de collecte, de traitement et d'analyse, le producteur peut disposer de métadonnées utiles pour qualifier les statistiques qui seront issues du processus : le nombre d'unités enquêtées, le taux de réponse, les procédures de contrôle, d'imputation, de redressement de la non-réponse.
- ❷ Lors de la phase de diffusion, le producteur peut enrichir les rubriques relatives aux publications prévues.

À l'issue de ces étapes, le producteur a ainsi progressivement construit toutes les métadonnées utiles à la compréhension des résultats. Il reste alors à les utiliser, pour alimenter les rapports qualité pour Eurostat, ou le site internet de l'Insee.

❸... STANDARDISÉES, ACCESSIBLES ET RÉUTILISABLES...

Pour que les métadonnées aient un rôle de médiation entre le chiffre et son utilisateur, il ne suffit pas qu'elles existent, elles doivent être facilement accessibles, aussi facilement que les chiffres eux-mêmes. Comme on l'a vu précédemment, les informations associées aux statistiques sont parfois directement intégrées aux commentaires : dans ce cas il ne s'agit pas réellement de données, isolables ; y accéder requiert une interprétation humaine. Cependant, la majorité des résultats statistiques ne font pas l'objet d'un commentaire spécifique. Ils sont diffusés dans de vastes fichiers de données mis en ligne sur le site insee.fr ou encore transmis directement de machine à machine *via* des interfaces. L'accès aux métadonnées doit alors être simple, visible, normalisé, cohérent dans le temps et adapté à ces différentes voies de transmission. Cette accessibilité revêt deux dimensions :

- ❶ **Un modèle de description des métadonnées partagé** : comme il y a une profusion et une diversité de métadonnées, il est nécessaire de modéliser leur description pour faciliter le dialogue entre les producteurs et les utilisateurs. L'existence d'une modélisation commune permet à l'utilisateur de savoir à quel endroit retrouver telle ou telle information. Ainsi, pour la centaine de métadonnées susceptibles de décrire les opérations statistiques, une norme a été adoptée en 2015 au niveau européen¹, les organisant en une vingtaine de groupes, comme « unité de mesure », « période de référence », « fréquence de diffusion », « coûts et charges », « traitements statistiques », etc.
- ❷ **Des formats et langages standardisés** : l'adoption de normes pour décrire les métadonnées facilite leur réutilisation, que ce soit pour rechercher et reprendre des variables, des nomenclatures ou pour alimenter diverses applications clientes. Elle est indispensable pour réaliser des appariements entre différentes sources. Par exemple, pour les enquêtes, un standard de documentation technique a été défini par un consortium international (projet DDI, *Data Documentation Initiative* initié en 1995). La standardisation de cette documentation consiste en une modélisation des différents objets statistiques (questions, questionnaires, variables, listes de codes...) et de leurs liens, le tout sous la forme de documents XML. Cette normalisation est d'abord conceptuelle, dans la mesure où chaque type de métadonnée a une place bien définie dans le modèle conceptuel de description des métadonnées. Mais elle doit aussi comporter un format technique d'échanges entre machines. La normalisation conceptuelle et technique favorise une réutilisation par des machines, *via* des API² (*Application Programming Interface*). DDI n'est pas le seul exemple : le standard SDMX (*Statistical Data and Metadata eXchange*) est porté par des organismes internationaux, en particulier Eurostat, pour promouvoir au niveau international

1. La norme SIMS (*Single Integrated Metadata Structure*) a été adoptée en novembre 2015 par le comité du système statistique européen (Comité SSE). Elle va progressivement être reprise dans les différents règlements et s'imposer pour la structuration des rapports qualité. Elle facilitera les comparaisons des données produites dans le cadre de règlements européens.

2. Interfaces de programmation applicative qui permettent à un logiciel d'offrir des services à un autre logiciel.

la diffusion et les échanges de données et de métadonnées statistiques. Ces standards sont très largement utilisés par les instituts nationaux de statistique.

1... CENTRALISÉES DANS UN RÉFÉRENTIEL : RMÉS

Comme on l'a souligné plus haut, les métadonnées statistiques sont essentielles à la compréhension des résultats statistiques qu'elles caractérisent. Leur normalisation permet d'y accéder facilement, voire de les réutiliser. Le producteur d'une seule opération statistique pourrait s'en tenir là. Mais pour un organisme qui gère un grand nombre d'opérations statistiques, comme l'Insee, la centralisation est nécessaire. Elle permet de garantir l'unicité de la représentation des métadonnées transverses et la cohérence de l'information mise à disposition, grâce à des règles d'identification.

Une gestion des « droits de propriété » permet de désigner pour chaque métadonnée le responsable de son actualisation. Cette caractéristique est particulièrement intéressante pour les métadonnées transverses, comme les nomenclatures et les concepts. Ainsi même si plusieurs services utilisent une même métadonnée, seul l'un d'entre eux (le « propriétaire ») est autorisé à la modifier. Par exemple à l'Insee, c'est le pôle Tourisme de la direction régionale d'Occitanie qui est « propriétaire » du concept de « nuitée »³. L'actualisation des métadonnées transverses peut être effectuée par le « propriétaire », une seule fois pour le bénéfice de toutes les opérations qui les utilisent. La centralisation et la normalisation des métadonnées rendent possibles d'autres mutualisations : par exemple l'usage des mêmes outils de gestion, pour créer les métadonnées non spécifiques, les modifier, les renseigner, quelle que soit l'opération statistique.

Cette fonction de centralisation (« guichet unique ») des métadonnées statistiques, auparavant garantie à l'Insee par le Dispositif de Documentation Structurée (DDS ; **encadré 2**), est désormais assurée par le Référentiel de Métadonnées Statistiques (RMÉS)⁴. Ce référentiel est en effet avant tout un lieu unique d'enregistrement d'objets, doté de règles communes de gestion. L'Insee a en outre délibérément adopté dans ce référentiel les normes et standards internationaux évoqués précédemment pour décrire les métadonnées statistiques, afin de faciliter leur réutilisation (cf. point précédent) et ouvrir la voie à des échanges avec d'autres instituts de statistiques, des services statistiques ministériels, des partenaires nationaux ou internationaux.

1 UN CONTENU DÉJÀ RICHE

Mis en production en 2018, RMÉS a repris les métadonnées transverses précédemment contenues dans DDS. Il contient ainsi environ 1 200 concepts (ex : chiffre d'affaires, taux de marge, taux de pauvreté, revenu disponible des ménages, bassin de vie...). Chacun de ces concepts a un « propriétaire » bien identifié au sein du service statistique public, responsable de la définition associée. RMÉS contient aussi six « séries » de nomenclatures. Structurantes pour la production statistique, les nomenclatures constituent des références,

3. Par délégation de la maîtrise d'ouvrage, le département des synthèses sectorielles au sein de la direction des statistiques d'entreprises.

4. La référence internationale sur le sujet des référentiels de métadonnées (*metadata repository*) est probablement Bo Sundgren, mathématicien suédois et professeur en traitement de l'information à la *Stockholm School of Economics* de 1987 à 2005 (<https://sites.google.com/site/bosundgren/my-life>).

validées par des groupes d'experts, permettant de classer l'information économique et sociale. L'Insee joue par exemple un rôle central pour la définition et l'actualisation de la nomenclature des Professions et Catégories Socioprofessionnelles (PCS). L'Insee assure en outre la gestion des nomenclatures d'activités et de produits françaises (NAF et CPF), en cohérence avec les nomenclatures équivalentes aux niveaux européen et international.

RMÉS s'est aussi progressivement enrichi de métadonnées plus spécifiques à certains dispositifs statistiques mais qui pourraient cependant être partagées. Il s'agit des variables, des listes de codes et du questionnaire utilisé par une enquête (voire *des* questionnaires dans le cadre d'une collecte en multimode). À titre d'exemple, l'enquête emploi en continu 2018 comprend environ 1 200 variables qui sont décrites dans RMÉS.

Toutes ces métadonnées sont associées à des opérations statistiques, qu'il s'agisse d'enquêtes, de dispositifs d'intégration de données administratives (ex : l'enquête de fréquentation dans l'hôtellerie 2018, les Déclarations Annuelles de Données Sociales 2015) ou d'autres types d'opérations qui mobilisent de nouvelles sources de données, comme les données de caisse.

RMÉS contient environ 130 séries d'opérations statistiques. Ces opérations peuvent être millésimées. Elles sont classées selon une arborescence en familles/séries/opérations qui structure le référentiel. Par exemple, l'enquête sur l'usage de l'informatique et des technologies de la communication dans les entreprises de moins de 10 personnes en 2016, dite « TIC TPE 2016 », est une opération rattachée à la série « TIC TPE » qui compose avec la série « TIC entreprises » et « TIC ménages » la famille « TIC ».

L'initialisation du référentiel a été effectuée à partir des métadonnées contenues dans DDS. Il reste désormais à le compléter par les métadonnées associées à des opérations nouvelles ou qui ne figuraient pas dans DDS. Dans le même temps, une réflexion de fond sur la qualité et la gouvernance des métadonnées doit être conduite pour s'assurer de la validité du référentiel dans la durée, de sa pérennité.

L'OFFRE DE SERVICES ASSOCIÉE AU RÉFÉRENTIEL

RMÉS n'est pas une base figée : afin que le référentiel vive et fonctionne de façon efficace, il est indispensable de disposer d'outils de gestion pour le mettre à jour et y accéder. C'est donc tout un environnement (applications de gestion, services) qui a été peu à peu mis en

Encadré 2. RMÉS a remplacé le Dispositif de Documentation Structurée (DDS)

À l'Insee, l'attention portée aux métadonnées n'est pas nouvelle. Dès les années 80, l'Insee avait conçu un dictionnaire de données statistiques innovant, qui a été rénové au début des années 2000 pour devenir le Dispositif de Documentation Structurée (DDS). Ce système a joué un rôle essentiel pour gérer et stocker la documentation de nombreuses opérations statistiques. DDS permettait également d'alimenter insee.fr et de produire le dictionnaire des codes, très utile pour les producteurs et les chargés d'études. Cependant il n'imposait pas de modèle unique contraignant, d'où une diversité des modélisations qui s'est avérée un frein aux mutualisations. DDS hébergeant en outre un grand nombre de bases autonomes, l'absence de liens entre elles faisait peser des risques d'incohérences. Enfin, l'émergence de standards internationaux, combinée à la croissance des échanges entre organismes, a encouragé à repenser ce dispositif.

place autour du référentiel (*figure 1*). Grâce à cela, le producteur peut aujourd'hui renseigner une seule fois⁵ dans le référentiel les métadonnées relatives à ses opérations pour satisfaire une diversité d'usages. En aval, le service le plus élémentaire du référentiel est de restituer l'information qu'il contient à des utilisateurs variés (humains ou machines), sous des formats divers.

L'un des premiers utilisateurs est le producteur lui-même. Il a en effet besoin des métadonnées statistiques de son opération pour produire de façon automatique certains documents nécessaires au processus de production dans le contexte de la statistique publique : demande d'avis d'opportunité au Cnis, contribution au dossier pour le comité du label

“ *Le référentiel est susceptible d'alimenter tout système qui utilise des métadonnées statistiques.* ”

(ces fonctionnalités sont prévues à terme dans RMÉS). Le producteur peut aussi constituer le dictionnaire des variables (« dico des codes ») à partir des métadonnées correspondantes figurant dans RMÉS, qu'il avait préalablement saisi.

Les métadonnées statistiques pouvant être renseignées tout au long du cycle de vie d'une opération, lorsque celle-ci est achevée, RMÉS contient toutes les informations nécessaires

pour produire un rapport qualité, notamment pour Eurostat lorsqu'il s'agit de statistiques européennes, ou tout type de documentation sur l'opération. Ainsi, le référentiel vise à documenter les fichiers de diffusion, notamment les micro-données mises à disposition par le Centre d'accès sécurisé aux données⁶ (CASD). Le portail d'accès aux fichiers détenus par le CASD reprend en effet la structure adoptée par RMÉS.

En alimentant les pages « Définitions, méthodes et qualité » du site internet de l'Insee, RMÉS permet une large mise à disposition des métadonnées statistiques, conformément au principe d'accessibilité du code de bonnes pratiques de la Statistique européenne. Cette alimentation s'effectue de façon automatique, pour faciliter la publication coordonnée des données ou publications et de la documentation de l'opération statistique concernée. Une rubrique spécifique est consacrée aux concepts et une aux nomenclatures. La rubrique « Sources statistiques et indicateurs » permet d'accéder à la description des opérations et le cas échéant à l'image du questionnaire d'enquête.

Plus généralement, le référentiel est susceptible d'alimenter tout système qui utilise des métadonnées statistiques, sous réserve du développement d'interfaces adaptées. Par exemple, il met d'ores et déjà à disposition la nomenclature d'activités *via* une API publiée conjointement à l'API Sirene sur le portail API de l'Insee. D'autres applications clientes pourront « se brancher » sur le référentiel pour récupérer des métadonnées transverses (par exemple les nomenclatures et concepts) ou propres à une opération : cette connexion, en évitant des duplications et des gestions parallèles autonomes, favorisera la cohérence et l'actualisation du système d'information statistique.

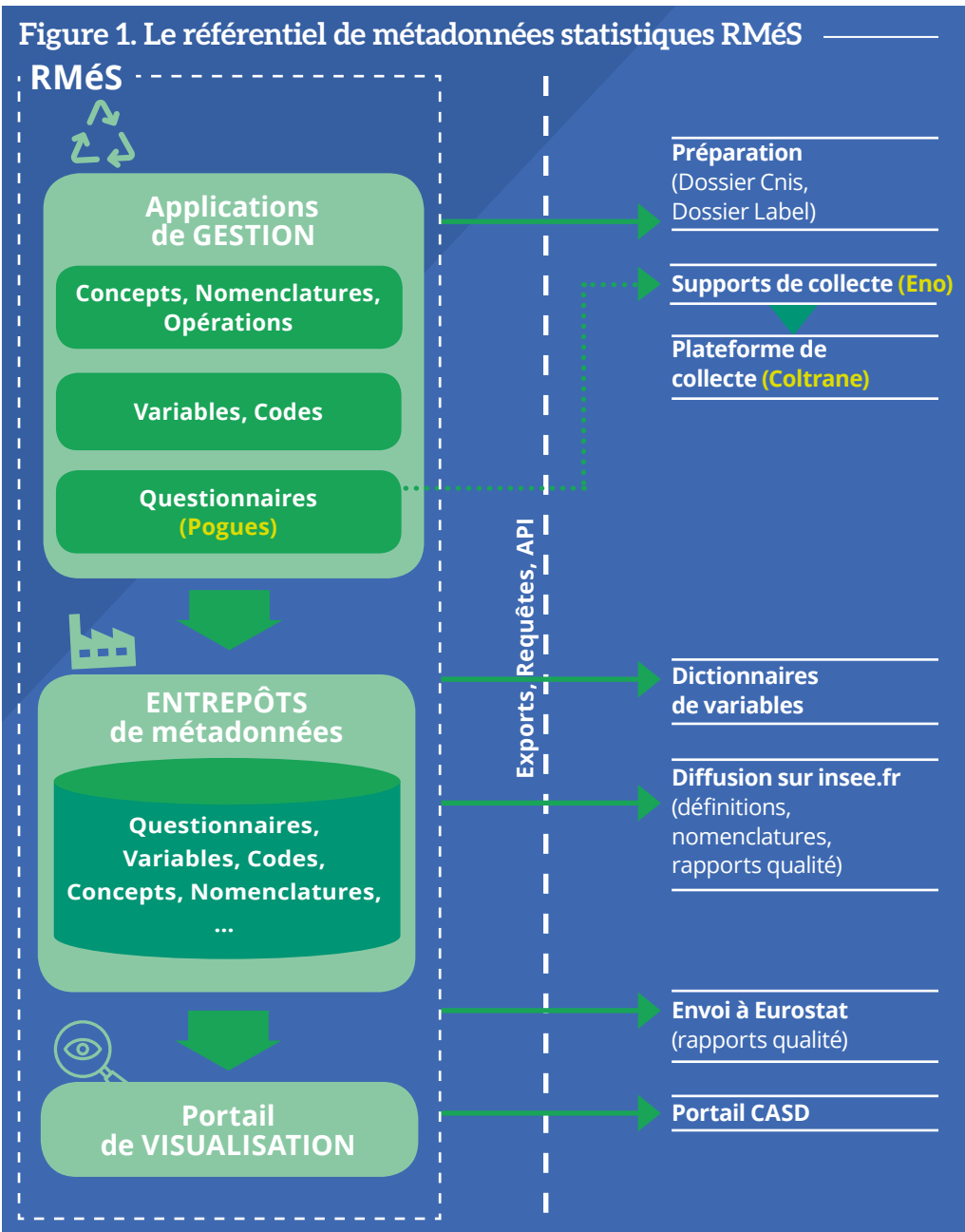
5. L'adoption en 2015 de la norme européenne de description des opérations statistiques mettait en avant le principe du « *once for all purposes* ».

6. Le CASD est un GIP (Groupement d'Intérêt Public) qui fournit à des chercheurs un outil sécurisé pour permettre l'accès à des données individuelles très détaillées après avis favorable du comité du secret statistique, et donc le plus souvent soumises à des conditions de sécurité élevées.

Pour parfaire ces services de dépôt ou d'extraction de métadonnées statistiques, il reste à développer un portail de consultation du contenu du référentiel.

VERS DES MÉTADONNÉES RENDUES ACTIVES

Le recours aux métadonnées contenues dans le référentiel peut aller au-delà de leur rôle de description et d'aide à la compréhension, avec accès *via* une interface dédiée ou des applications clientes. De nouveaux outils cherchent ainsi à tirer parti du caractère exhaustif et normalisé des métadonnées pour produire automatiquement des composants du processus statistique.



Elles acquièrent ainsi en quelque sorte un statut nouveau, passant du stade d'« informations facilitant la compréhension des statistiques », au stade de « données participant au processus de production » ; d'où l'idée de métadonnées actives.

Dans cette logique, en amont d'un processus d'enquête, une interface graphique de conception d'un questionnaire (dénommée « Pogues ») a été développée pour permettre aux concepteurs d'enquêtes de spécifier un questionnaire selon les standards internationaux,

« [Les métadonnées] acquièrent ainsi en quelque sorte un statut nouveau, passant du stade d'« informations facilitant la compréhension des statistiques », au stade de « données participant au processus de production ». »

sans avoir à connaître la technicité de ces formats. Très concrètement, les concepteurs d'enquêtes définissent la structure de leur questionnaire en séquences et modules. Ils saisissent les questions qu'ils souhaitent poser aux unités enquêtées et les modalités de réponse. Avec l'application Pogues, ils peuvent produire à la demande les métadonnées correspondantes respectant les standards⁷ (variables et listes de codes, le cas

échéant nomenclatures). Un deuxième outil, Eno, permet de générer les supports de collecte (questionnaire en PDF, support de collecte par internet) et de visualiser le support ainsi obtenu, de façon transparente pour le concepteur. Ce générateur produit des supports adoptant une mise en forme standardisée, dans un format approprié. Les supports peuvent ensuite être déposés sur une plateforme dédiée (par exemple « Coltrane », plateforme de collecte pour les enquêtes auprès des entreprises) et la collecte peut alors être lancée. Dans cet exemple, c'est le générateur de supports de collecte qui rend les métadonnées « actives ».

Cette interface de conception de questionnaires s'appuyant sur les métadonnées statistiques montre que ce type de démarche engendre des bénéfices immédiats pour le concepteur. Le gain est aussi collectif : gain en qualité avec la cohérence, la traçabilité et la transparence des métadonnées mais aussi gain du fait des mutualisations possibles. Une telle approche contribue à gagner en efficience et plus généralement en qualité. Des travaux sont en cours pour examiner dans quelle mesure le principe de métadonnées actives pourrait être utilisé pour décrire des contrôles ou spécifier des produits de diffusion.

📍 UN RÉFÉRENTIEL POUR L'ENSEMBLE DU SERVICE STATISTIQUE PUBLIC ?

L'attention portée aux métadonnées est largement partagée au sein de la statistique publique. D'ores et déjà, les services statistiques ministériels ont accès en lecture, comme tout internaute, à l'entrepôt qui héberge les concepts, et les nomenclatures, une vue du référentiel étant accessible sur insee.fr. Une étroite coopération est ainsi nouée entre l'Insee et les SSM pour harmoniser les concepts communs et valider la documentation des opérations communes.

7. En l'occurrence le standard DDI, déjà évoqué.

L'objectif à terme est cependant plus ambitieux. Il vise à ouvrir aux services statistiques ministériels un accès en écriture, pour plusieurs raisons :

- ❶ Les SSM sont propriétaires de certains concepts et ont à ce titre un rôle de validation de l'information intégrée par l'unité qualité dans le référentiel ;
- ❷ Ils peuvent aussi gérer des nomenclatures spécifiques à leur domaine mais utilisées plus largement au sein du SSP (nomenclature des infractions, nomenclature des familles professionnelles...);
- ❸ Certains d'entre eux souhaitent pouvoir décrire leurs sources dans le référentiel et en extraire des rapports qualité au format européen ;
- ❹ Enfin, quelques-uns souhaitent avoir accès à l'application de spécification de questionnaires pour leurs enquêtes.

Une réflexion est amorcée pour identifier l'ensemble des besoins des SSM, afin d'examiner quelles réponses pourraient y être apportées et dans quel cadre. Il s'agit de conforter l'utilisation des métadonnées comme un des leviers de la stratégie qualité du Service Statistique Public. Dans cette optique, et sur la base d'une étroite concertation qui demeure à définir, RMÉS sera probablement amené à jouer un rôle central dans les années qui viennent.

COTTON, Franck et DUFFES, Guillaume, 2010. *SDMX : Un standard pour l'échange de données et de métadonnées statistiques*. Document interne Insee.

COTTON, Franck, MARTIN, Mélanie et TAILHURAT Romain, 2018. *Report on the implementation of three statistical services – ESSnet SCFE (Sharing common functionalities in the ESS) – Deliverable D3-1* [en ligne]. [Consulté le 8 avril 2019]. Disponible à l'adresse : https://ec.europa.eu/eurostat/cros/system/files/scce_-_d3-1_-_report_on_the_implementation_of_three_statistical_services.pdf

CROSNIER, Dominique, 2000. Le nouveau DDS de l'Insee. In : *Courrier des statistiques* [en ligne]. Mars 2000. n° 93, pp. 10-17. [Consulté le 24 mai 2019]. Disponible à l'adresse : <https://www.epsilon.insee.fr/jspui/bitstream/1/8513/1/cs93c.pdf>

DESROSIERES, Alain, 2008. Les catégories socioprofessionnelles. In : *Courrier des statistiques* [en ligne]. Novembre 2008. n° 125, pp. 13-15. [Consulté le 24 mai 2019]. Disponible à l'adresse : <https://www.epsilon.insee.fr/jspui/bitstream/1/8513/1/cs93c.pdf>

DUBOIS, Thomas et KEROUANTON, Marie-Hélène, 2018. The French statistical Metadata Repository, RMÉS : managing metadata throughout the whole statistical process. In : *European conference on quality in official statistics, Krakow* [en ligne]. [Consulté le 8 avril 2019]. Disponible à l'adresse : https://www.q2018.pl/wp-content/uploads/Sessions/Session%2028/Marie%20H%C3%A9l%C3%A8ne%20K%C3%A9rouanton/Session%2028_Marie%20H%C3%A9l%C3%A8ne%20K%C3%A9rouanton.docx

DUFFES, Guillaume, 2014. *Use of standards at Insee*. Document interne Insee : note pour le Workshop of the Modernisation Committee on Standards : International Collaboration for Standards-Based Modernisation, Genève, Suisse, 5 – 7 mai 2015.

EUROSTAT, 2015. *ESS handbook for quality reports*. In : *Theme 1 : General and regional statistics* [en ligne]. Édition 2014, Luxembourg, Publications Office of the European Union, Manuals and guidelines. [Consulté le 8 avril 2019]. Disponible à l'adresse : <https://ec.europa.eu/eurostat/documents/3859598/6651706/KS-GQ-15-003-EN-N.pdf/18dd4bf0-8de6-4f3f-9adb-fab92db1a568>

GREGORY, Arofan, 2011. The Data Documentation Initiative (DDI) : An Introduction for National Statistical Institutes. In : *Open Data Foundation* [en ligne]. Juillet 2011. [Consulté le 8 avril 2019]. Disponible à l'adresse : http://odaf.org/papers/DDI_Intro_forNSIs.pdf

GREGORY, Arofan et HEUS, Pascal, 2007. DDI and SDMX : Complementary, Not Competing, Standards, Version 1.0. In : *Open Data Foundation* [en ligne]. [Consulté le 8 avril 2019]. Disponible à l'adresse : http://odaf.org/papers/DDI_and_SDMX.pdf

GUIBERT, Bernard, LAGANIER, Jean et VOLLE, Michel, 1971. Essai sur les nomenclatures industrielles. In : *Économie et statistique* [en ligne]. Février 1971. n° 20, pp. 23-36. [Consulté le 24 mai 2019]. Disponible à l'adresse : https://www.epsilon.insee.fr/jspui/bitstream/1/17169/1/estat_1971_20_3.pdf

POULAIN Claude, 1983. Le dictionnaire des données d'une production statistique. In : *Courrier des statistiques* [en ligne]. Juillet 1983. n° 27, pp. 25-29. [Consulté le 24 mai 2019]. Disponible à l'adresse : <https://www.epsilon.insee.fr/jspui/bitstream/1/14105/1/cs27.pdf>

ROUPPERT, Benoît et KEROUANTON, Marie-Hélène, 2014. *Projet de Référentiel de Métadonnées Statistiques*. Document interne Insee : note pour le comité des investissements N°181/DG75/MHK.

ÉLIRE, UN PROJET AMBITIEUX AU SERVICE DU RÉPERTOIRE ÉLECTORAL UNIQUE

Magali Demotes-Mainard*

Augmenter la participation électorale en retardant autant que possible la date limite d'inscription sur les listes électorales, tel était l'objectif premier de la loi de 2016 qui a institué le Répertoire électoral unique et en a confié la gestion à l'Insee. La construction de ce nouveau répertoire et de son système de gestion, menée en moins de trois ans dans un calendrier contraint, a constitué pour l'institut un projet original et porteur de forts enjeux, du fait de la multiplicité des partenaires, du travail spécifique d'initialisation du répertoire et de l'attention portée à la sécurité des données. Au-delà du projet, la gestion courante du répertoire électoral fait évoluer le rôle de l'Insee au sein du processus électoral, en lui confiant une responsabilité et une visibilité accrue.

 *Increasing voter turnout by delaying the deadline for registration on the electoral rolls as far as possible was the main objective of the 2016 Act that instituted the Single Electoral Register and entrusted its management to Insee. The construction of this new register and its management system, carried out in less than three years within a constrained timetable, constituted for Insee an original project with high stakes, due to the multiplicity of partners, the specific work of initializing the directory and the attention given to data security. Beyond the project, the current management of the electoral register is changing Insee's role within the electoral process, giving it greater responsibility and visibility.*

Après la seconde guerre mondiale, du fait des mouvements de population et de l'élargissement du corps électoral aux femmes et aux militaires, il est apparu que les communes pouvaient difficilement assurer la mise à jour des listes électorales sans un dispositif organisé d'échange d'information. Dans la même période, il avait été demandé à l'Insee de créer puis de gérer le Répertoire national d'identification des personnes physiques (RNIPP). Et dès 1946 (Godinot, 2013), le Fichier général des électeurs (FE), destiné à organiser les flux d'information pour l'établissement des listes électorales des communes, a intégré les missions « régaliennes¹ » de l'institut, avec la tenue des référentiels de nomenclatures, des répertoires d'entreprises ou de personnes. En modernisant le dispositif, la loi de 2016 a accru la responsabilité de l'Insee. Cet article décrit les opérations conduites dans le cadre du programme Élire (Élaboration des Listes à partir du Répertoire Électoral) depuis trois ans et les enjeux pour l'institut.

* Directrice du programme Élire, Insee,
magali.demotes-mainard@insee.fr

1. Godinot qui mentionne que « peu de pays semblent [avoir fait le choix de confier la tenue d'un fichier électoral à leur institut statistique], la solution la plus généralement adoptée étant la mise en place d'une Commission nationale électorale en principe indépendante du pouvoir politique, chargée de l'ensemble des opérations de préparation et de contrôle des élections. »

📍 L'INSEE, ACTEUR DES ÉCHANGES D'INFORMATION ENTRE COMMUNES DEPUIS 70 ANS

Après avoir construit en 1947-1948 un fichier image des listes communales (Insee, 1950), l'Insee enregistrait dans ce fichier tous les mouvements électoraux dont il était informé pour en aviser les communes concernées afin qu'elles les reportent sur leurs propres listes (**encadré 1**). Ainsi par exemple, à chaque nouvelle inscription d'un électeur, la commune devait en informer l'Insee qui recherchait dans le FE la précédente commune d'inscription et notifiait à cette dernière qu'elle devait radier l'électeur de sa propre liste. De même, l'Insee informait les communes des décès et des radiations d'office. À partir de 1998, les jeunes majeurs ont pu bénéficier d'une procédure d'inscription d'office² et c'est encore l'Insee qui en transmettait la liste aux communes. L'Insee jouait ainsi le rôle de plaque tournante de l'information dans un dispositif qui restait, malgré les progrès de l'informatisation et de la dématérialisation, à la fois lourd dans sa mise en œuvre et fragile dans ses résultats. Car les listes électorales utilisées pour les scrutins restaient celles que chaque commune mettait à jour, sans autre vérification que celle des commissions électorales annuelles.

Encadré 1. Trois types de listes électorales

Listes électorales principales : il s'agit des listes électorales établies par les communes, recensant les personnes de nationalité française qui remplissent les conditions pour être électeur dans la commune.

Listes électorales complémentaires : il s'agit des listes électorales établies par les communes, recensant les ressortissants d'un autre État membre de l'Union européenne qui remplissent les conditions pour être électeur dans la commune soit pour les élections municipales (liste complémentaire municipale) soit pour l'élection des représentants français au Parlement européen (liste complémentaire européenne).

Listes consulaires : il s'agit des listes électorales établies par les ambassades et les postes consulaires, recensant les personnes de nationalité française qui remplissent les conditions pour être électeur dans l'ambassade ou le poste consulaire.

En avril 2019, les 47,1 millions d'électeurs inscrits sur les listes électorales françaises se répartissent ainsi* :

- sur liste communale principale : 45,5 millions,
- sur liste communale complémentaire : 330 000,
- sur liste consulaire : 1,3 million.

* Voir (Demotes-Mainard, 2019).

📍 LA LOI DE 2016 MODIFIE SENSIBLEMENT LA NATURE DE SA RESPONSABILITÉ

Dans son rapport de 2014, la représentation parlementaire a cherché comment « Mieux établir nos listes électorales pour revitaliser notre démocratie » (Pochon et Warsmann, 2014).

2. Depuis 1997, le ministère de la Défense collecte des informations relatives aux jeunes lors de journées « Défense et citoyenneté ».

Encadré 2. Le programme Élire : deux projets en un, l'agilité pour tenir les délais

Le programme Élire comporte formellement deux projets de nature différente mais conduits en pratique de façon conjointe et par la même équipe de spécificateurs et de développeurs : un projet visant à construire le « Système de gestion » et un projet d'« Initialisation » du répertoire.

La conduite du programme Élire était très fortement contrainte dans ses délais de réalisation : la loi qui institue le REU a été promulguée le 1^{er} août 2016 et est entrée en vigueur le 1^{er} janvier 2019. Le recours à une méthode agile de conduite de projet a permis de prioriser les développements en fonction des jalons calendaires impératifs.

Pour le développement du système de gestion, l'objectif était de livrer l'API* destinée aux communes en octobre 2018 pour la version permettant la vérification de l'initialisation et début janvier 2019 pour la version cible : même si toutes les fonctions n'étaient pas utiles dès le début de l'année, on a voulu éviter de déployer des versions successives.

Par ailleurs, comme beaucoup de communes utilisent des logiciels de gestion électorale, il fallait permettre aux éditeurs d'anticiper leurs propres développements : un environnement de test leur avait donc été ouvert dès le début de l'année 2018 et enrichi progressivement au fur et à mesure des avancées du côté de l'Insee.

Pour accéder au REU, les utilisateurs communaux s'authentifient par une adresse électronique. Les préfetures avaient recueilli au cours de l'été 2018 l'adresse électronique du compte primaire de chaque commune qui, pour sécuriser les accès, avait ensuite été créé centralement. La période d'initialisation de fin 2018 a été mise à profit pour résoudre tous les problèmes techniques de connexion que les communes pouvaient rencontrer.

En complément des développements fonctionnels, différents tests ont été conduits pour valider la capacité du système à répondre aux exigences du projet : tests des circuits d'échange, tests de performance afin d'offrir un service de qualité, tests de charge prenant en compte les périodes de pic d'activité et tests de sécurité. Le « bac à sable » mis à disposition des éditeurs a permis de tester et de fiabiliser le système mis en place avant le déploiement en production.

Un outil de supervision du fonctionnement de l'application a été mis en place. À partir notamment de représentations graphiques, il permet de surveiller le fonctionnement et les performances de l'API et, en suivant la façon dont ils l'utilisent, de guider les éditeurs dans leurs développements.

En parallèle aux développements du système de gestion et aux travaux d'initialisation ont été menées au sein de l'Insee une démarche qualité du futur processus, une analyse organisationnelle du devenir des anciens sites de gestion électorale de l'Insee et une étude de sécurité du système d'information placé sous la responsabilité de l'institut. Cette dernière démarche s'est inscrite dans un dispositif d'homologation « chapeau » de l'ensemble de l'« écosystème REU », mis en place par le ministère de l'Intérieur et destiné à vérifier la cohérence des objectifs de sécurité et des mesures de prévention des risques portés par les différents systèmes d'information intervenant dans le processus.

* Application Programming Interface : interface de programmation applicative qui permet à un logiciel d'offrir des services à un autre logiciel.

Ses travaux ont débouché sur le vote de la loi n° 2016-1048 du 1^{er} août 2016 : celle-ci vise à moderniser le dispositif, en instituant un Répertoire électoral unique (REU) d'où sont extraites les listes qui font désormais référence pour les scrutins. Ce répertoire, dont la gestion a été à nouveau confiée à l'Insee, est l'un des deux leviers introduits pour permettre une mise à jour des listes électorales rapide et en continu ; l'autre levier est le transfert au maire ou au responsable d'un poste consulaire des décisions d'inscription et de radiation, en lieu et place des commissions administratives annuelles. Le nouveau processus permet ainsi d'arrêter ces listes au plus près du scrutin : la date limite d'inscription, dictée essentiellement par le délai de recours, a été fixée en régime courant au 6^e vendredi précédant le scrutin.

Dans le nouveau dispositif, l'Insee n'est donc plus seulement en charge d'assurer la performance des échanges d'information : il doit garantir que le contenu des listes électorales, retraçant les décisions individuelles des maires et des responsables consulaires ainsi que les mouvements d'office prévus par la loi, est conforme aux principes du code électoral.

❶ LES ENJEUX D'UN SYSTÈME INTÉGRÉ : QUALITÉ DE LA DONNÉE ET AUTOMATISATION

Un répertoire n'est pas une simple liste d'observations, et son établissement suppose un système de gestion qui peut être complexe. Il s'agit en effet d'assurer la cohérence interne des données et de définir les règles permettant de contrôler la pertinence des mises à jour ; il faut également organiser les échanges d'information entre les différents acteurs.

Pour construire le système d'information du REU répondant aux nouvelles exigences du code électoral³, l'Insee a mis en place le programme Élire (**encadré 2**). Le principe du nouveau système est de permettre aux communes et aux postes consulaires de notifier directement au REU leurs décisions d'inscription et de radiation ; en complément, il intègre automatiquement les données provenant des administrations disposant d'informations sur la capacité électorale des individus.

“*l'Insee doit garantir la qualité de l'identification des électeurs.*”

Pour assurer l'unicité de l'inscription sur les listes électorales et la bonne affectation des événements électoraux notifiés par les différents partenaires, l'Insee doit garantir la qualité de l'identification des électeurs. Pour cela, le système de gestion d'Élire dispose tout d'abord d'une fonction de recherche sur état civil, simplifiée et limitée au champ du REU. Si

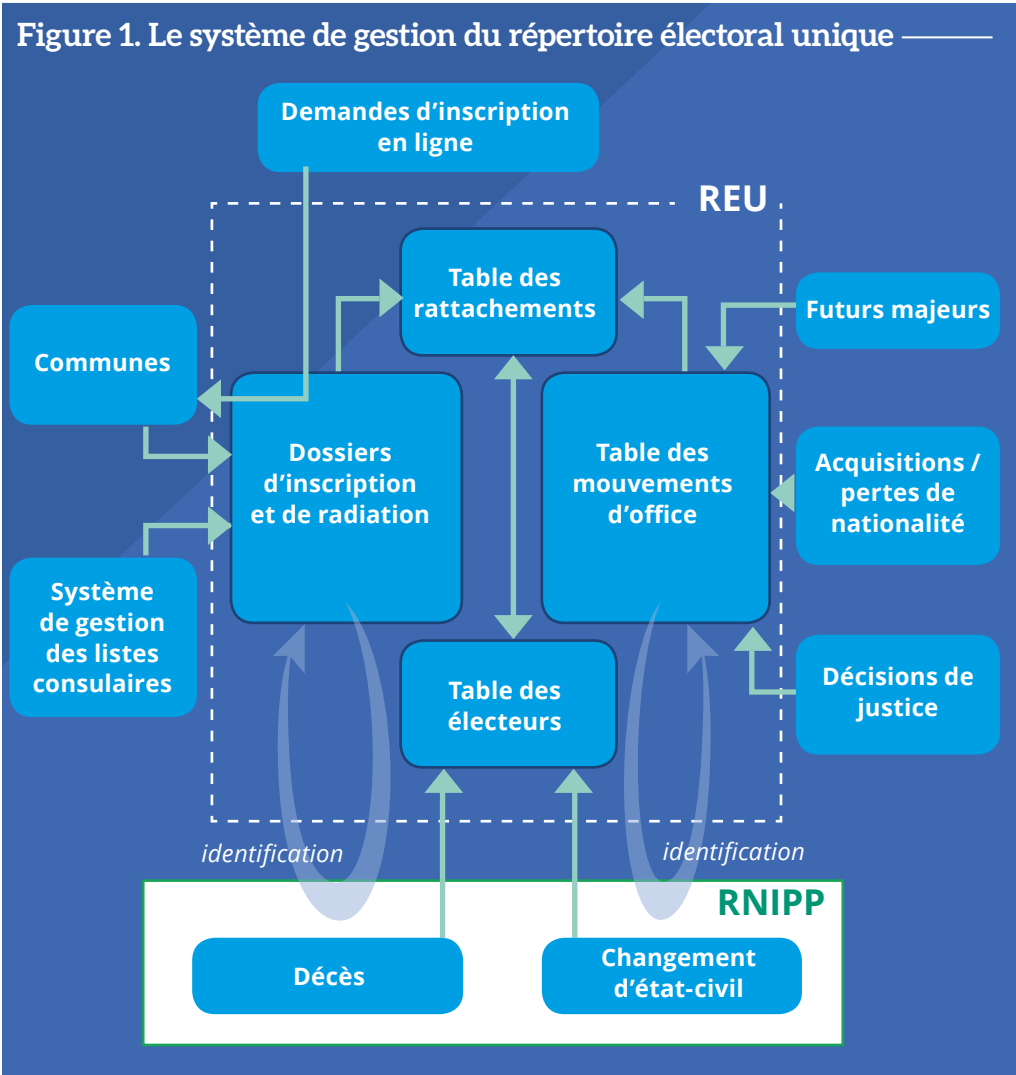
l'électeur ne peut pas être identifié de cette façon, le système interroge automatiquement le Répertoire national d'identification des personnes physiques (RNIPP) par le biais d'un service d'identification : en cas d'échec de la recherche automatique, le cas est soumis à l'expertise d'un gestionnaire de l'Insee. Au final, le système renvoie un identifiant de l'individu et son état civil tel qu'il est connu au RNIPP, et donc tel qu'il figure désormais sur les documents électoraux.

3. Dans le nouveau processus, le rôle dévolu à l'Insee est de prendre en compte dans le REU les décisions d'inscription des maires, des responsables consulaires et des commissions de contrôle, de procéder aux inscriptions et radiations d'office, de notifier aux communes et aux postes consulaires les mouvements affectant leur liste électorale et de mettre à disposition les listes électorales extraites du REU.

Par ailleurs, le système de gestion du RNIPP notifie automatiquement au REU les décès des électeurs ainsi que les changements d'état civil : toute modification de l'état civil des électeurs est ainsi directement reportée dans le REU.

Enfin, le système de gestion du REU récupère les demandes d'inscription déposées en ligne par les électeurs eux-mêmes sur le site www.service-public.fr : le dossier, y compris les pièces justificatives, est alors mis à la disposition des communes pour instruction. Ce dispositif permet d'ouvrir le service d'inscription en ligne à tous les électeurs potentiels, et évite aux communes d'investir dans un raccordement technique spécifique.

Le dispositif mis en place s'interface ainsi avec plus de 35 000 communes et 100 préfectures, mais aussi avec les systèmes d'information du ministère de l'Europe et des Affaires étrangères (MEAE) pour les listes consulaires qu'il gère centralement, et de différentes autres administrations (*figure 1*). Il doit par ailleurs s'adapter à l'organisation spécifique aux collectivités d'outre-mer (*encadré 3*).



GROS VOLUMES D'ÉCHANGES ET FORTES CONTRAINTES DE SÉCURITÉ

Le système de gestion du REU prend également en compte des contraintes non fonctionnelles. D'une part, il doit gérer des volumes de mouvements très importants, qui peuvent atteindre par an 2 millions d'inscriptions volontaires, fortement concentrées à l'approche de la date limite d'inscription, 800 000 inscriptions d'office (notamment des jeunes majeurs), 500 000 décès et 500 000 radiations pour perte d'attache.

D'autre part, il porte des enjeux de sécurité importants. Il doit être disponible en permanence pour ses nombreux utilisateurs, avec une attention particulière au moment des périodes cruciales du calendrier électoral. Il gère des données personnelles, notamment l'adresse des électeurs, d'où le besoin de garantir un haut niveau de confidentialité. Enfin, une altération des listes électorales peut être préjudiciable à la validité d'un scrutin, ce qui exige de se prémunir contre les risques d'intrusion et de corruption à tous les stades du dispositif.

À tous ses utilisateurs et partenaires, le système a cherché à proposer des outils de gestion et des modalités d'interaction simples et performantes, tout en assurant la sécurité et la confidentialité des échanges.

Pour que les communes et le système de gestion des listes consulaires dialoguent avec le REU, l'option retenue a été d'offrir des interfaces de programmation applicatives (API) accessibles par internet : cette solution apporte en effet aux utilisateurs une grande souplesse dans le raccordement à leur propre système d'information. Les communes peuvent manipuler l'API soit par une interface homme-machine (IHM) développée par l'Insee (« Portail Élire ») soit par un logiciel de gestion électoral, qui peut alors offrir des fonctionnalités complémentaires et s'adapter à l'organisation interne de chacune. Les autres administrations partenaires

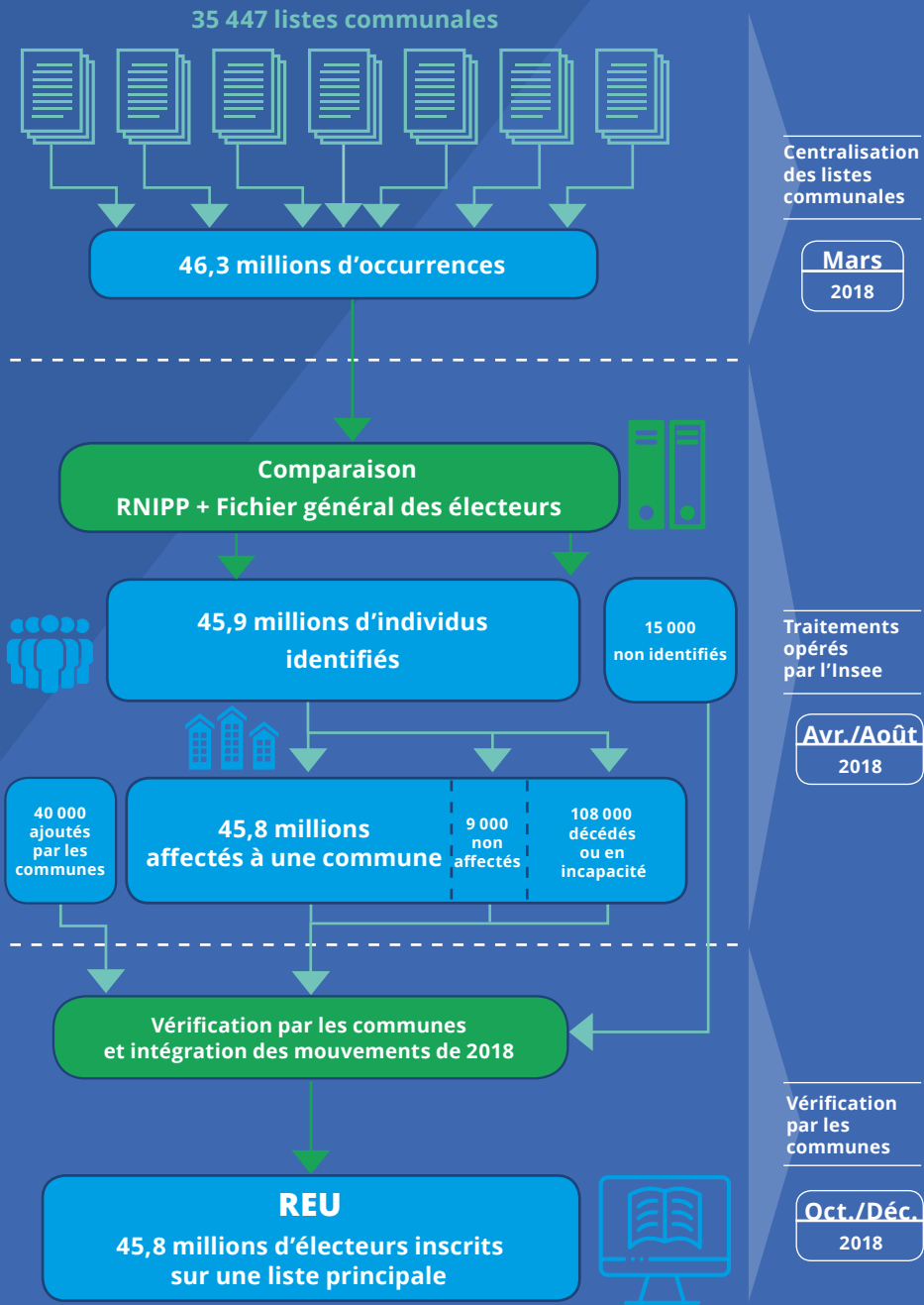
Encadré 3. Le Répertoire électoral unique en Outre-Mer

La loi rénovant les modalités d'inscription sur les listes électorales s'applique de plein droit dans les départements d'outre-mer, mais aussi à Saint-Pierre-et-Miquelon, Saint-Barthélemy et Saint-Martin. Dans les autres collectivités d'outre-mer, les règles applicables en métropole doivent y être expressément étendues par la loi. En l'espèce, la loi n°2016-1048 s'applique également en Polynésie française et à Wallis-et-Futuna, mais pas en Nouvelle-Calédonie qui conserve le principe de listes annuelles tenues par les commissions administratives communales.

Selon la loi, la responsabilité de retracer dans le REU les inscriptions et radiations des électeurs relève de l'Insee pour Saint-Pierre-et-Miquelon, Saint-Barthélemy et Saint-Martin, mais de l'Institut statistique de Polynésie française (ISPF) pour la Polynésie et de l'Administrateur supérieur pour Wallis-et-Futuna. En pratique, pour assurer en temps réel l'unicité d'inscription et pour éviter de dupliquer les fonctionnalités du système de gestion, l'ISPF et l'Administrateur supérieur de Wallis-et-Futuna ont délégué par convention à l'Insee la mise en œuvre de leur responsabilité : le système de gestion développé par l'Insee couvre donc aussi le champ de la Polynésie française et de Wallis-et-Futuna et prend en charge la mise à jour du REU y compris pour ces collectivités.

L'ISPF et l'Administrateur supérieur doivent néanmoins disposer de fonctionnalités spécifiques leur permettant, sur leur champ de compétence, de procéder à des inscriptions et des radiations d'office dans le REU, pour compléter si besoin l'information détenue par l'Insee, et de se substituer à leurs communes ou circonscriptions pour la gestion des listes électorales.

Figure 2. Les étapes du traitement des listes communales principales



ont privilégié des échanges d'information par transfert automatisé de fichiers, solution qui impactait moins l'existant de leurs propres systèmes d'information.

❶ INITIALISER LE RÉPERTOIRE : UN PROJET EN SOI

La mise au point d'un système de gestion permet de faire vivre un répertoire, en permettant sa mise à jour, mais il faut aussi constituer le « stock » de départ. L'initialisation du REU a donc constitué un projet en soi, où la finalité du traitement des données ne pouvait pas être seulement statistique : l'enjeu de qualité d'un répertoire est que chaque donnée individuelle soit traitée avec autant de soin que possible.

Pour cela, la démarche a consisté à procéder d'abord à des traitements permettant de régler les cas les plus nombreux, puis à affiner progressivement l'examen des cas particuliers subsistants, aussi loin qu'il était possible dans le temps imparti (*figure 2*).

L'option prise pour initialiser le REU a été de partir des listes communales et consulaires issues de la révision électorale 2017/2018⁴ : en maîtrisant les évolutions par rapport aux listes électorales qui faisaient référence avant la réforme, cette solution devait limiter les contestations tant de la part des communes que des électeurs en assurant une meilleure continuité avec l'existant.

Les listes consulaires sont gérées par une application centralisée : la récupération de ces listes ne posait donc pas de difficulté. Pour les listes communales, on a mobilisé les fichiers départementaux normalisés issus du système de dépôt du ministère de l'Intérieur, sauf pour Paris et Marseille qui ont transmis directement leurs listes à l'Insee. Grâce à l'implication des préfetures dans le suivi des dépôts des communes, la centralisation a été d'excellente qualité, tant en nombre qu'en contenu : seules une vingtaine de petites communes n'ont pas fourni de listes exploitables, ce qui a conduit dans ce cas à utiliser l'extrait du Fichier général des électeurs pour initialiser le répertoire.

❷ UNE BASE DE 48 MILLIONS D'ENREGISTREMENTS À VALIDER

La compilation des fichiers de 2018 a constitué une première base de plus de 48 millions d'observations⁵. Mais cette base ne pouvait pas telle quelle constituer un répertoire, qui doit respecter des principes d'exhaustivité et de cohérence interne des données. Il fallait donc repérer et traiter tous les cas où les données des listes rassemblées soit étaient insuffisantes soit portaient une impossibilité logique (doublons, électeurs décédés ou en incapacité, etc.).

Dans les listes électorales communales, les électeurs ne sont connus que par les éléments de leur état civil, qui peuvent ne pas être repris de façon strictement identique en cas d'inscription multiple : il était donc indispensable de confronter ces éléments avec un état

4. Une solution alternative aurait été de soumettre aux communes le contenu du Fichier général des électeurs (ce qui aurait évité la phase d'identification mais aurait fortement augmenté la charge de travail des communes), éventuellement après avoir utilisé pour les plus grandes communes le dispositif existant de mise en concordance.

5. Les données de la figure 2 ne concernent que les enregistrements présents sur les listes principales (46,3 millions d'occurrences), ce qui représente toutefois 97 % des cas.

civil de référence, en l'occurrence celui du RNIPP. La création d'un lien entre un électeur du REU et un individu du RNIPP est de toute façon indispensable en régime courant pour pouvoir ensuite attribuer à chaque électeur les événements (inscription, radiation, décès, incapacité électorale, modification d'état civil) qui le concernent.

Le premier travail réalisé par l'équipe de projet a donc consisté à identifier les 48 millions d'observations, c'est-à-dire à rechercher un individu connu de la base du RNIPP à partir des cinq éléments d'état civil de la liste électorale (nom, prénoms, sexe, date et lieu de naissance) dont certains pouvaient être approchés, erronés ou manquants. Ce travail s'est fait en plusieurs étapes, de façon à pouvoir traiter efficacement un volume d'observations très important (**encadré 4**) : en effet, même après une étape de codification du lieu de naissance, une simple comparaison stricte sur les cinq champs n'aurait permis d'associer que 70 % des électeurs à un individu de la base du RNIPP, ce qui laissait de côté près de 14 millions de personnes.

Les traitements automatiques d'identification ont permis de repérer les électeurs présents deux fois ou plus dans la compilation des listes communales. Dans l'ancien dispositif de tenue des listes électorales, l'absence d'automatisme dans les échanges comportait en effet un risque de rupture dans le circuit d'information, d'où la possibilité qu'un électeur reste inscrit simultanément sur deux listes (voire plus) centralisées à la même date.

Encadré 4. La première identification des électeurs

L'identification des électeurs présents dans la compilation des listes électorales récupérées a commencé par un premier travail de nettoyage et de normalisation de l'information.

Le point le plus délicat a été la codification du lieu de naissance : il n'était pas possible en effet de fonder la suite du traitement sur de simples libellés, du fait de la grande variété de leur retranscription. On a utilisé pour cela une table de passage préexistante à l'Insee ; en complément, l'interrogation systématique d'une API de Google a permis de coder des lieux de naissance à l'étranger qui ne figuraient pas dans cette table.

On a ensuite procédé à un appariement entre les listes communales et consulaires et le Fichier général des électeurs (FE) par comparaison des états civils, en relâchant progressivement certaines contraintes de correspondance stricte, mais en rajoutant comme donnée de contrôle la commune d'inscription, information présente dans le FE. Cette première phase a ainsi permis d'identifier 97 % des électeurs inscrits sur les listes principales et 75 % des électeurs inscrits sur les listes complémentaires et consulaires.

Les électeurs non identifiés après ce premier appariement ont été transmis au service d'identification automatique du RNIPP, plus performant mais plus strict et plus coûteux que le programme développé pour la première phase : cela a produit un écho valide pour 64 % des électeurs des listes complémentaires restant à identifier, et pour un peu plus de 40 % de ceux des listes complémentaires.

Ces résultats ont été complétés par la mobilisation d'autres sources, notamment la « base d'apprentissage » constituée par le travail d'expertise individuelle réalisé entre juin 2017 et février 2018 sur le reliquat non identifié des listes communales 2017 et la conclusion des identifications manuelles réalisées dans le cadre de la révision électorale 2017/2018. S'est enclenchée à partir de juin 2018 une expertise individuelle par les gestionnaires du RNIPP, qui a porté sur près de 100 000 électeurs.

Pour trancher ces cas de double inscription, il fallait retenir le lieu de l'inscription la plus récente, ce qui supposait de connaître les dates d'inscription. Cette information ne figurait pas dans les fichiers des listes communales centralisées mais était souvent présente dans le FE : quand ce n'était pas le cas, l'Insee ne pouvait pas conclure, ce qui l'a conduit à mettre en place en 2017 et 2018 des enquêtes auprès des communes pour compléter l'information manquante.

📍 46,8 MILLIONS D'ÉLECTEURS ACTIFS DANS LA BASE SOUMISE AUX COMMUNES...

Ces traitements centralisés d'identification et de résolution des inscriptions multiples ont permis de construire fin août 2018 une base du REU contenant 46,8 millions d'électeurs actifs, dont 355 000 présents sur plusieurs listes communales mais pour lesquels un rattachement unique était proposé par l'Insee. Pour faciliter la vérification par les communes, la base reprenait également 100 000 électeurs décédés et 13 000 individus privés du droit de vote présents sur les listes électorales centralisées.

Il restait par ailleurs 30 000 électeurs (soit moins d'un pour mille) qui ne pouvaient pas être insérés dans le répertoire soit parce qu'ils se trouvaient dans une configuration complexe d'inscription multiple (12 600 électeurs), soit parce qu'on n'avait pas réussi à les identifier (17 400 électeurs) : une part importante des électeurs non identifiés fin août venaient de Mayotte et de Polynésie française, zones où un dispositif spécifique avait été mis en place dès le début de l'été 2018 avec les acteurs locaux pour récupérer rapidement les actes d'état civil nécessaires. Les travaux d'identification se sont poursuivis en parallèle entre la constitution du fichier à charger (fin août) et la soumission de la version initiale du REU aux communes (octobre 2018).

📍 ... ELLES-MÊMES INVITÉES À VÉRIFIER LEURS LISTES

À ce stade, l'Insee était allé au bout de ce qu'il pouvait faire centralement. Il a alors passé la main aux communes et au ministère de l'Europe et des Affaires étrangères en leur soumettant le résultat de son travail pour qu'ils complètent et corrigent si besoin la proposition, mais aussi pour qu'ils s'approprient le contenu du REU : les communes et les postes consulaires restent en effet responsables de la tenue de leurs listes électorales.

La version initiale du REU a donc été chargée dans le système de gestion et, à compter d'octobre 2018, le portail Élire a été ouvert aux communes : des fonctionnalités spécifiques avaient été développées pour faciliter la comparaison entre le REU et les listes initiales (*figure 3*). Les communes devaient notamment examiner le cas de leurs électeurs figurant parmi les individus non intégrés dans le REU : après vérification, elles pouvaient le cas échéant rajouter ces électeurs en indiquant une date d'inscription (ce qui permettait de résoudre les cas d'inscription multiple résiduels) ou en modifiant ou complétant l'état civil (ce qui pouvait permettre de faire aboutir l'identification).

Les communes étaient aussi invitées à s'assurer de la pertinence des radiations opérées pour les électeurs intégrés au répertoire, qu'il s'agisse de radiation pour décès ou incapacité ou pour cause d'inscription plus récente dans une autre commune : dans ce dernier cas, les communes pouvaient, directement à travers le portail, contester la radiation en indiquant une date d'inscription plus récente que celle de l'autre commune.

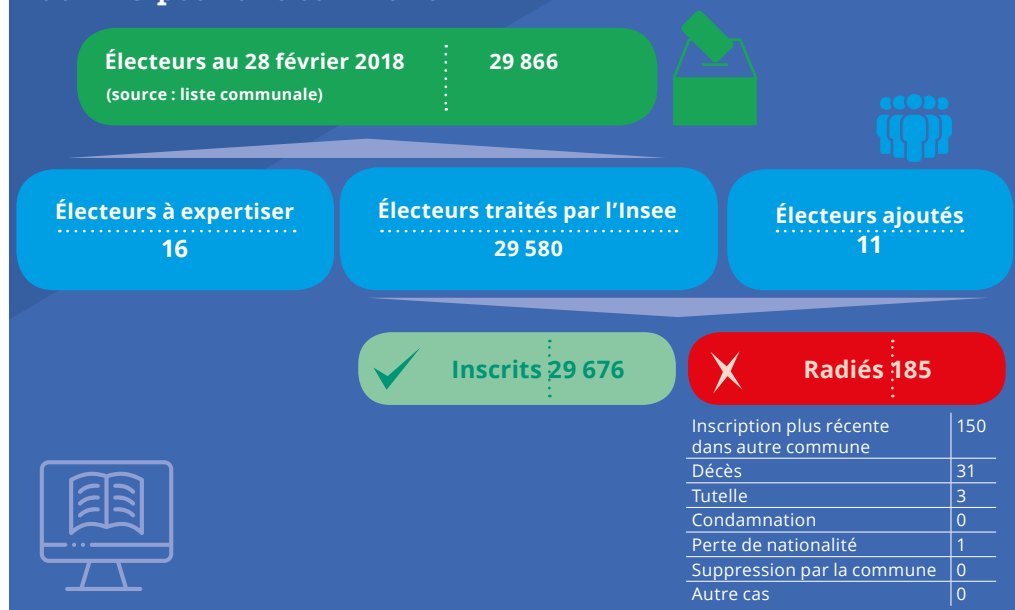
LEVER LES DOUTES SUR L'IDENTITÉ D'UN ÉLECTEUR : UN TRAVAIL MINUTIEUX

Il était enfin demandé aux communes de vérifier les écarts importants entre l'état civil figurant sur la liste électorale et l'état civil du REU (soit celui du RNIPP), parce qu'ils pouvaient recouvrir des erreurs d'identification ; ils pouvaient aussi empêcher la commune de faire le lien avec ses précédentes listes électorales, voire faire porter un doute sur l'identité de l'électeur le jour du vote. Pour éviter aux communes un balayage systématique de leurs listes, l'écart entre les cinq champs d'état civil avait été qualifié en trois catégories selon son importance. Seule une très faible part des écarts d'état civil paraissait de nature à soulever un doute sur l'identification de l'électeur ; mais les communes ont signalé plusieurs dizaines de milliers de différences, plus ou moins importantes, dont certaines correspondaient bien à des défauts de mises à jour du RNIPP.

Les listes consulaires devaient de la même façon être vérifiées, mais l'opération était de bien plus faible ampleur : sur 1,3 million d'inscrits sur les listes consulaires, seuls 343 électeurs n'avaient pas pu être insérés dans le REU à l'issue des travaux centralisés. La centralisation préexistante assurait en effet une excellente coordination entre listes consulaires ; par ailleurs, l'inscription simultanée sur une liste consulaire et sur une liste principale restait admissible jusqu'à fin mars 2019, si bien que ces cas de double inscription n'avaient pas à être traités. Mais 20 000 écarts entre l'état civil du RNIPP et celui de la base des listes consulaires ont été signalés, écarts qui ont été également expertisés et traités par les gestionnaires du RNIPP.

Cette phase de contrôle décentralisé a contribué à faire évoluer la version initiale du REU. Fin décembre 2018, les communes avaient ainsi ajouté 40 000 électeurs au REU et en avaient supprimé 13 000. Les observations des communes et du ministère de l'Europe et des Affaires étrangères ont permis de déceler et de corriger environ 1 700 erreurs d'identification,

Figure 3. Exemple de tableau de synthèse de l'initialisation du REU pour une commune



et de lever un peu plus de 300 incapacités qui n'étaient pas d'actualité. La prise en charge des signalements relatifs à l'état civil s'est prolongée sur tout le début de l'année 2019.

L'Insee a enfin repris la main pour procéder au premier trimestre 2019, après la vérification faite par les communes, à différents contrôles qui l'ont amené à revenir individuellement vers les communes concernées ou vers les électeurs : vérification de l'état civil du reliquat de 3 500 électeurs qui restaient non identifiés et de la pertinence de reprendre ces électeurs dans le REU, signalement pour réinscription éventuelle des 2 700 électeurs inscrits au départ sur plusieurs listes mais en janvier 2019 revendiqués sur aucune et des 2 500 électeurs européens dont la nationalité n'était pas précisée. Les cas d'une grande partie de ces électeurs ont ainsi pu être régularisés.

DÉBUT 2019, INJECTER LES MOUVEMENTS 2018 EN PARALLÈLE AU FONCTIONNEMENT COURANT

La version du REU proposée pour validation aux communes en octobre 2018 se voulait être l'image des listes électorales au 28 février 2018. Il fallait donc y rajouter les inscriptions et les radiations intervenues depuis cette date, et pour l'essentiel décidées par les commissions administratives qui se sont réunies une dernière fois fin 2018/début 2019.

Ces mouvements ont été intégrés au REU à l'initiative des communes à l'aide des fonctionnalités standard d'inscription et de radiation de l'API d'Élire, et en parallèle à la notification des mouvements courants puisque le REU était opérationnel dès le 1^{er} janvier 2019. Certaines communes ont intégré les mouvements 2018 dès janvier 2019, mais dans l'ensemble leurs envois se sont étalés jusqu'à mi-mars 2019, le temps que soit finalisée la mise au format requis par Élire du stock d'enregistrements constitué avec leur ancien système d'information. Les mouvements des listes consulaires ont été chargés par le ministère de l'Europe et des Affaires étrangères début février 2019. L'Insee devait de même intégrer les mouvements d'office au titre de 2018, ce qui s'est fait entre mi-janvier et début mars 2019 : comme pour les mouvements à l'initiative des communes et des postes consulaires, il a fallu remettre au format d'entrée du REU les informations présentes dans le FE et le RNIPP.

Entre janvier et mars 2019, le REU a ainsi enregistré 1 million d'inscriptions volontaires, 600 000 radiations pour perte d'attache communale ou consulaire, 900 000 inscriptions d'office, 600 000 décès, 40 000 autres radiations d'office et 80 000 modifications d'état civil.

Le référentiel des communes a aussi dû prendre en compte les 238 créations de communes nouvelles intervenues après le 1^{er} mars 2018 : les fusions de listes dans le REU ont été réalisées au cours du mois de février 2019.

Quelques autres opérations de gestion importantes ont été menées au premier semestre 2019 : suppression à compter d'avril 2019 de la possibilité d'être inscrit simultanément sur liste consulaire et liste communale par radiation centralisée de la liste communale de tous les doubles inscrits à cette date, levée fin mars 2019 de toutes les incapacités pour tutelle suite à la mise en œuvre de la loi sur la réforme de la justice qui a aboli cette mesure, opération d'échanges de données avec les États membres de l'Union européenne et de gestion des mentions à apposer sur les listes d'émargement pour éviter la possibilité de double vote.

Le contenu du REU était ainsi globalement stabilisé en mars 2019, et prêt à être rendu public. Le service en ligne permettant à tout électeur de vérifier son inscription sur la liste électorale de sa commune a été ouvert le 11 mars. Les communes ont envoyé à tous leurs électeurs, en général en avril ou en mai, leur nouvelle carte électorale, sur laquelle figure le numéro national d'électeur qui les identifie désormais dans ce répertoire.

La première échéance nationale pour la mise en œuvre du nouveau dispositif électoral défini par la loi de 2016 a été celle des élections européennes du 26 mai 2019, mais le REU est la référence pour tous les scrutins organisés depuis le 17 mars 2019, et de nombreuses élections municipales partielles se tiennent chaque semaine depuis cette date, utilisant les listes électorales extraites du nouveau répertoire.

Maintenant que le fonctionnement courant est enclenché, il s'agit pour l'Insee de mettre en place le processus bien spécifique de gestion de ce répertoire. Le REU est en effet plus qu'un simple fichier : c'est un outil de travail quotidien pour 35 500 communes, de tailles très hétérogènes donc aux attentes potentiellement variées. Il faut donc que non seulement le système mis à disposition mais aussi la communication et l'accompagnement soient dans la durée à la hauteur de cet enjeu.

Par ailleurs, l'application développée, notamment parce qu'elle associe différentes fonctionnalités, paraît être plus qu'un dispositif d'échange d'information entre l'Insee et les communes : elle ressemble par certains aspects à un outil de gestion interne du processus électoral. Dès lors, si s'expriment à l'avenir des souhaits de compléments à l'application, l'Insee aura à se demander jusqu'où doit aller sa prestation de service en la matière.

Encadré 5. L'accompagnement des communes

La mise en œuvre de la loi de 2016 modifie profondément la façon dont les communes tiennent leurs listes électorales. Elles ont donc dû se familiariser avec un nouveau processus en même temps qu'elles découvraient de nouveaux outils.

Pour les communes, la première manifestation concrète de la réforme a été la phase de vérification et de validation du REU menée entre octobre et décembre 2018. Pour cette opération, dont il coordonnait le processus, l'Insee a mobilisé d'une part son service Insee Contact, qui assurait le point d'entrée, et d'autre part les sites de gestion des répertoires de personnes qui traitaient sur le fond les observations remontées des communes. Cette charge d'assistance a été d'une ampleur exceptionnelle pour l'Insee sur les trois derniers mois de 2018.

À compter du 1^{er} janvier 2019, avec l'entrée en vigueur du nouveau processus électoral et la mise à disposition des nouvelles versions des logiciels, s'est manifesté un besoin d'accompagnement d'une autre nature.

Malgré les textes et la documentation produits, les formations proposées et les actions de communication menées, de nombreux agents communaux n'avaient pas une vision claire du nouveau processus et notamment du rôle respectif de leurs différents interlocuteurs : préfectures, Insee, éditeurs de logiciel. L'Insee a alors continué à être fortement sollicité, de façon encore accentuée dans les semaines précédant le scrutin du 26 mai 2019, et bien souvent sur des domaines où il n'avait pas de réponse à apporter : dans ce contexte, assurer le service d'assistance s'est révélé un exercice difficile, mais que les acteurs impliqués ont assuré avec beaucoup de professionnalisme.

BIBLIOGRAPHIE

DEMOTES-MAINARD, Magali, 2019. 47,1 millions d'électeurs présents en avril 2019 sur les listes électorales françaises, In : *Insee Focus* [en ligne]. 7 mai 2019. N° 154.
[Consulté le 28 mai 2019]. Disponible à l'adresse : <https://www.insee.fr/fr/statistiques/4134308>

GODINOT, Alain, 2013. Faut-il confier à des statisticiens des responsabilités en matière électorale ?. In : *Statistique et société* [en ligne]. 28 mai 2013. Vol. 1, n° 1, pp. 23-28.
[Consulté le 28 mai 2019].
Disponible à l'adresse : <http://statistique-et-societe.fr/article/view/147/144>

INSEE, 1950. Étude sur le fichier général des électeurs. *Supplément au Bulletin mensuel de statistique*, avril-juin 1950, pp. 149-181.

POCHON, Élisabeth et WARSMANN, Jean-Luc, 2014. Mieux établir nos listes électorales pour revitaliser notre démocratie. In : *Site de l'Assemblée nationale* [en ligne]. 17 décembre 2014. N° 2473. [Consulté le 28 mai 2019].
Disponible à l'adresse : <http://www.assemblee-nationale.fr/14/pdf/rap-info/i2473.pdf>

STATISTIQUES CONJONCTURELLES D'ENTREPRISES,

DES IMPULSIONS EUROPÉENNES, DES AVANCÉES FRANÇAISES

Philippe Scherrer*

La construction européenne a connu une accélération au moment de la création de l'euro, le 1^{er} janvier 1999. Pour doter la Banque centrale européenne d'indicateurs économiques conjoncturels, l'Union européenne avait voté en 1998 un règlement sur les statistiques d'entreprise de court terme instaurant un cadre harmonisé pour mesurer l'offre et la demande, les facteurs de production et les prix à la production. Ce règlement a été modifié à plusieurs reprises pour améliorer la pertinence et la couverture de ces indicateurs. De nouvelles évolutions sont en discussion depuis 2011, notamment pour mieux couvrir les secteurs en croissance comme les services et le commerce, au sein d'un nouveau règlement cadre englobant toutes les statistiques d'entreprises, qui a été récemment adopté par le Conseil et le Parlement européens. Les débats se sont orientés depuis 2018 vers la construction d'un indice mensuel de la production totale marchande et, en réponse à la crise de 2008, vers la nécessité de couvrir l'immobilier commercial. Le service statistique public français, qui s'appuie sur un dispositif de production éprouvé, a su s'adapter ou anticiper ces évolutions européennes. Il s'est de ce fait mis en mesure de respecter le plus tôt possible les nouvelles dispositions sur les statistiques conjoncturelles qui s'appliqueront à partir du 1^{er} janvier 2021.

 *European integration accelerated with the creation of the euro on 1st January 1999. To provide the European Central Bank with short-term economic indicators, the European Union adopted a regulation on short-term business statistics in 1998, which established a harmonised framework for measuring supply and demand, production factors and producer prices. This Regulation has been amended several times to improve the relevance and coverage of these indicators. New developments have been under discussion since 2011, in particular to better cover growing sectors such as services and trade, and in the framework of a new Regulation, covering all business statistics, which has just been adopted by the European Council and Parliament. Discussions focused since 2018 on the construction of a monthly index of total market production and, in response to the 2008 crisis, on the need to cover commercial real estate. The French public statistical service, which relies on a proven production system, has been able to adapt or anticipate these European developments. It has taken steps to comply as soon as possible with the new provisions on short-term statistics that shall apply in 2021.*

* Chef du département des statistiques de court terme, Insee,
philippe.scherrer@insee.fr

La construction européenne a connu une accélération au moment de la création de l'euro, le 1^{er} janvier 1999. La création d'une union monétaire intégrée et la mise en place d'une coordination des politiques économiques et budgétaires gérées par les États membres ont nécessité de doter la Commission européenne et la Banque centrale européenne d'indicateurs statistiques adaptés au pilotage de l'Union économique et monétaire. En France, les statistiques conjoncturelles d'entreprises, déjà bien développées, ont bénéficié de cet élan européen. Cette dynamique, qui va connaître un nouveau développement début 2021, est appelée à se poursuivre pour les années futures.

❶ CONSTRUIRE UN TABLEAU DE BORD CONJONCTUREL DE LA ZONE EURO...

En mai 1998, le cadre européen définissant la production de statistiques conjoncturelles d'entreprise a été posé. Il instaurait des statistiques conjoncturelles harmonisées de l'offre et de la demande, des facteurs de production (main-d'œuvre) et des prix à la production. Ce cadre a évolué de manière progressive à partir des années 2000, à la suite d'un plan d'actions pour la poursuite des travaux dans le domaine des statistiques nécessaires au bon fonctionnement de l'Union économique et monétaire (UEM). L'Union européenne s'est dès lors attachée à la construction d'un véritable tableau de bord conjoncturel de l'Union économique et monétaire, fondé sur les principaux indicateurs économiques européens (PIEE) introduits en 2002 (**encadré 1**), dont la moitié est issue de la statistique d'entreprise. Les délais de mise à disposition de nombreux indicateurs ont été ainsi réduits pour se rapprocher des standards américains et, dans le domaine des statistiques d'entreprise, la couverture des facteurs de production et celle des prix de la production ont été étendues aux services et celle des prix aux importations de produits industriels.

Ces indicateurs conjoncturels représentent une information primordiale pour le suivi à court terme de l'économie, de l'identification de points de retournement du cycle économique à un stade précoce, parallèlement ou en association avec d'autres grands indicateurs macroéconomiques. Ils sont utilisés dans l'élaboration des comptes nationaux trimestriels,

Encadré 1. Les principaux indicateurs économiques européens (PIEE)

Ce sont un ensemble d'indicateurs économiques relatifs à l'Union européenne et à ses États membres et essentiels pour la surveillance de la zone euro. Le concept de PIEE a été introduit pour la toute première fois dans la communication de la Commission européenne en 2002. Ces statistiques visent à décrire la conjoncture économique, la situation du marché du travail, ainsi que l'évolution des prix, trois éléments qui revêtent une importance particulière pour la politique économique et monétaire. Les indicateurs couvrent les sujets suivants : la balance des paiements (investissements directs étrangers, etc.), les enquêtes de conjoncture et de consommation, l'industrie et les services (production industrielle, commerce de détail, chiffre d'affaires dans les services, prix, etc.), le commerce extérieur, le marché du travail (chômage, coût de la main-d'œuvre, vacance d'emploi, etc.), les indicateurs monétaires et financiers (taux d'intérêt, taux de change), les comptes nationaux (PIB, dette et déficit, investissements, épargne, consommation, etc.), les prix (taux d'inflation, prix des logements, etc.).

Près de la moitié des PIEE proviennent des statistiques conjoncturelles. Il s'agit : de l'emploi, du coût de la main-d'œuvre, des prix à la production dans l'industrie, des prix à l'importation de produits industriels, de la production industrielle, de la production dans la construction, du chiffre d'affaires déflaté des ventes au détail, du chiffre d'affaires dans les services et des permis de construire.

au travers des différentes composantes du produit intérieur brut. Les indicateurs de prix de la production ou des importations sont des signaux avancés de l'inflation (prix à la consommation); ils sont aussi largement utilisés pour le partage volume/prix de nombreuses grandeurs économiques (indices de la production industrielle, indices de volume des chiffres d'affaires, comptes nationaux annuels et trimestriels, etc.). L'ensemble de ces indicateurs conjoncturels revêt de ce fait un intérêt particulier pour la Banque centrale européenne (BCE) et de l'Eurogroupe dans leur suivi de l'économie de la zone euro et pour des organismes internationaux, comme le Fonds monétaire international (FMI) et l'Organisation de coopération et de développement économique (OCDE).

... INSCRIT EN 1998 DANS UN RÈGLEMENT EUROPÉEN SUR LES STATISTIQUES CONJONCTURELLES...

Le cadre définissant la production de statistiques communautaires à court terme sur le cycle conjoncturel a été défini par le règlement du Conseil européen sur les statistiques conjoncturelles en date du 19 mai 1998.

Ce dernier instaurait une normalisation des statistiques conjoncturelles pour répondre aux besoins d'information européens en matière de convergence économique. Cet objectif d'harmonisation a été renforcé par la rédaction de différents manuels méthodologiques. Le règlement cherchait aussi, dès le départ, à alléger les procédures pour les entreprises, entre autres par la promotion de nouvelles technologies pour la collecte des données et l'élaboration des statistiques ainsi que par l'utilisation des sources administratives existantes, de sorte à ne pas constituer une charge disproportionnée par rapport aux résultats que les utilisateurs de ces statistiques étaient en droit d'attendre.

Le règlement définit très précisément les activités spécifiques pour lesquelles les statistiques sont à élaborer, les types d'unités statistiques à utiliser, la liste des variables, leur forme (données brutes, données corrigées des variations saisonnières et/ou des jours ouvrables), leur période de référence, leur niveau de détail, les délais de transmission des données, etc.

Ce règlement européen sur les statistiques conjoncturelles est toujours en vigueur aujourd'hui. Il couvre le champ de l'industrie, de la construction, du commerce et des services. Les statistiques demandées sont multiples et portaient initialement sur :

-  la production ;
-  le chiffre d'affaires (distinguant selon le cas marché intérieur et marchés extérieurs) ;
-  les entrées de commandes (distinguant selon le cas marché intérieur et marchés extérieurs, bâtiments et génie civil) ;
-  le nombre de personnes occupées ;
-  le nombre d'heures travaillées ;
-  les salaires et traitements bruts ;
-  les prix à la production (marché intérieur, marchés extérieurs) ;
-  les coûts de construction (distinguant selon le cas les coûts des matériaux et les coûts salariaux) ;
-  les permis de bâtir (distinguant selon le cas le nombre de logements et la superficie utile) ;
-  et le déflateur des ventes.

Ces statistiques ne sont pas tant à utiliser en niveau qu'en évolution pour mesurer les variations à court terme de l'activité, des facteurs de production et des prix à la production. Les périodes de référence de ces variables sont généralement le mois voire le trimestre. Selon leur nature, les États membres doivent transmettre à Eurostat ces variables peu ou prou entre 35 et 90 jours calendaires à compter de la fin de la période de référence.

Certains de ces indicateurs conjoncturels font partie des principaux indicateurs économiques européens, d'où une pression des utilisateurs pour un raccourcissement des délais de production de ces indicateurs.

1... QUI S'EST PROGRESSIVEMENT ADAPTÉ AUX BESOINS

En juin 2000, le Conseil pour les affaires économiques et financières (Ecofin) a invité la Commission européenne, en coopération étroite avec la BCE, à établir un plan d'action pour l'UEM (Ecofin, 2000). L'objectif de ce plan était de définir les domaines où des progrès urgents étaient nécessaires. Dès août 2000, la BCE a exprimé « les besoins dans le domaine des statistiques économiques générales » (Banque centrale européenne, 2000). Ce plan d'action concernant les obligations statistiques dans le cadre de l'UEM a aussitôt été validé par le Conseil Ecofin en septembre 2000.

Dès lors, le règlement européen sur les statistiques conjoncturelles a fait l'objet de plusieurs évolutions (tableau 1).

Les changements les plus notables sont :

1 en 2005 : l'ajout de la variable prix à l'importation de produits industriels, qui en plus de la disponibilité de la variable prix de la production exportée, permet de s'affranchir des indices de valeur unitaire du commerce extérieur, puis l'ajout de la variable prix à la production dans les services, compte tenu du poids de ce secteur dans le PIB, ensuite l'ajout de la ventilation zone euro et hors zone euro pour les variables portant sur les marchés extérieurs, et enfin le raccourcissement des délais de transmission pour les variables production industrielle et nombre de personnes occupées ;

Tableau 1. Indicateurs conjoncturels demandés au titre du règlement européen de 1998				
Variables	Industrie	Construction	Commerce de détail	Autres services*
Production	X	X		
Chiffre d'affaires total	X		X	X
Chiffre d'affaires domestique	X			
Chiffre d'affaires non domestique	X			
Déflateur des ventes			X	
Nombre de personnes employées	X	X	X	X
Heures travaillées	X	X	X	X
Salaires et coûts du travail	X	X	X	X
Prix de production total	X			
Prix de production marché domestique	X			X
Prix de production marché non domestique	X			
Prix à l'importation	X			
Coûts de production		X		
Permis de construire		X		

* hors services financiers, hors services non marchands...

Tableau 2 . Évolutions sur les indicateurs conjoncturels prévus dans le règlement cadre FRIBS

	CONSTRUCTION	COMMERCE (détail, gros, réparation d'automobiles et de motocycles)	SERVICES
PRODUCTION	Passage de la nomenclature CC à la nomenclature NACE		Nouvelle variable mensuelle sur la production dans les services
CHIFFRE D'AFFAIRES		Fréquence désormais mensuelle pour le commerce de gros et le commerce et réparation d'automobiles et de motocycles	Extension de champ aux activités immobilières, activités de location et location-bail, services relatifs aux bâtiments et aménagement paysager
VOLUME DES VENTES		Extension de champ au commerce de gros et commerce et réparation d'automobiles et de motocycles	
DÉFLATEUR DES VENTES		Suppression de la variable mensuelle déflateur des ventes du commerce de détail	
NOMBRE DE PERSONNES EMPLOYÉES HEURES TRAVAILLÉES SALAIRES ET COÛTS DU TRAVAIL			Extension de champ aux activités immobilières, activités de location et location-bail, services relatifs aux bâtiments et aménagement paysager
PRIX DE PRODUCTION	Nouvelle variable trimestrielle sur les prix de production des logements neufs		Extension du B2B au B2All (prise en compte du B2Others) Extension de champ à diverses activités*
COÛTS DE PRODUCTION	Suppression de la variable trimestrielle coût de production dans la construction		
PERMIS DE CONSTRUIRE	Passage de la nomenclature CC à la nomenclature CPA		

N.B. : Les évolutions concernent très marginalement l'industrie et ne sont donc pas mentionnées dans ce tableau.

*Extensions au transport terrestre et transport par conduites, transport par eau, entreposage et les services auxiliaires des transports, hébergement, restauration, édition, production de films cinématographiques, de vidéo et de programmes de télévision, enregistrement sonore et édition musicale, programmation et diffusion, activités immobilières, autres activités spécialisées, scientifiques et techniques, activités de location et de crédit-bail, activités des agences de voyages, voyagistes, services de réservation et activités connexes, services relatifs aux bâtiments et aménagement paysager et activités administratives et autres activités de soutien aux entreprises.

- ❶ **en 2006** : la suppression des variables entrées de commandes dans le bâtiment et les travaux publics ;
- ❷ **en 2008** : l'ajout des variables nombre d'heures travaillées et salaires et traitements bruts dans le commerce et les services ;
- ❸ **en 2012** : la suppression des variables relatives aux entrées de commandes industrielles. Celles-ci étaient destinées à servir d'indicateur précurseur concernant la production future. Toutefois, leur capacité prédictive s'est avérée limitée et, dans la mesure où leur caractère précurseur n'a pas été démontré de manière stable dans tous les États membres, le comité du système statistique européen a convenu de mettre un terme à la collecte de ces données en vue d'alléger la charge pesant sur les entreprises et le système statistique européen.

❶ FRIBS, UN NOUVEAU RÈGLEMENT POUR TOUTES LES STATISTIQUES D'ENTREPRISES...

En 2011, l'Union européenne a entrepris une démarche plus globale qui couvre l'ensemble des statistiques d'entreprises. Depuis cette date, elle a travaillé à un règlement cadre sur l'intégration des statistiques des entreprises (FRIBS¹) qui a été adopté récemment par le Conseil et le Parlement européens. Celui-ci entrera en vigueur au 1^{er} janvier 2021 et constituera un cadre juridique commun pour le développement, la production et la diffusion de statistiques d'entreprises

« Un projet qui vise à accroître la cohérence des statistiques produites dans ces différents domaines tout en poursuivant leur développement. »

européennes, tant conjoncturelles que structurelles, et pour le réseau européen des registres de statistiques d'entreprises et le répertoire européen des groupes.

Il ne s'agit pas d'une simple consolidation d'actes juridiques existants

dans les différents domaines des statistiques d'entreprise mais d'un projet qui vise à accroître la cohérence des statistiques produites dans ces différents domaines tout en poursuivant leur développement.

Pour les indicateurs conjoncturels, l'idée maîtresse du nouveau règlement est de rééquilibrer les attendus entre l'industrie et la construction d'une part, et le commerce et les services d'autre part, de manière à refléter la part de plus en plus importante prise par le secteur des services dans les économies européennes au cours des dernières décennies. L'unité statistique sera désormais la seule unité d'activité économique (subdivision d'une entreprise) alors que le règlement sur les statistiques conjoncturelles de 1998 considérait deux unités statistiques, différentes selon le secteur d'activité, pouvant être ou l'unité d'activité économique ou l'entreprise (**encadré 2**). Par ailleurs, le nouveau règlement recommande la mise à jour annuelle des pondérations pour les indicateurs de prix et de volume, en association avec des méthodes de chaînage (**encadré 3**), sans pour autant que cela devienne une obligation formelle.

Les principales attentes en matière de statistiques conjoncturelles sont au nombre de trois : développer un indice mensuel de production dans les services, étendre l'indice mensuel du volume des ventes au commerce de gros et à la réparation d'automobiles et de motos et étendre le champ et la couverture de l'indice trimestriel de prix à la production dans les services. Par ailleurs, les indicateurs de la construction seront réorganisés et des innovations méthodologiques seront introduites (**tableau 2**).

1. Framework Regulation Integrating Business Statistics (FRIBS).

1... AVEC UNE MEILLEURE COUVERTURE DU COMMERCE ET DES SERVICES...

Dans le règlement de 1998, un indice de volume des ventes n'est pas demandé directement pour le commerce de détail, mais il est obtenu indirectement en rapportant l'indice de chiffre d'affaires par l'indice du déflateur des ventes correspondant, tous deux requis pour le commerce de détail. Le nouveau règlement cadre étendra cet indice de volume des ventes à l'ensemble du commerce (par l'ajout du commerce et de la réparation de véhicules et motocycles et du commerce de gros). Le nouvel indicateur complet devra également être disponible deux mois après le mois de référence. Cet indice de volume des ventes pourra aussi être obtenu en déflétant les indices élémentaires de chiffre d'affaires du commerce par des indices de prix appropriés. Les indices de prix à la production du commerce n'existant pas à ce jour, Eurostat permet aux instituts nationaux de statistique de tirer le meilleur parti des informations disponibles au niveau national pour la construction de déflateurs du commerce (prix à la production, prix à la consommation, etc.). Ces indices élémentaires sont ensuite agrégés à partir de pondérations fondées sur le chiffre d'affaires, ce qui permet *in fine* de calculer un indice de volume des ventes.

De manière générale aux différents indicateurs conjoncturels, la couverture des services sera améliorée dans le nouveau règlement cadre. Elle inclura notamment les activités immobilières, les activités de location et location-bail et les services relatifs aux bâtiments et paysages pour les chiffres d'affaires, le nombre de personnes employées, les heures travaillées et salaires et coûts du travail.

Un indice mensuel de production dans les services sera instauré par ce règlement cadre. Cet indice pourra être obtenu en déflétant les indices élémentaires de chiffre d'affaires des services par des indices élémentaires de prix à la production des services correspondants. Les indices de prix à la production dans les services étant généralement produits à fréquence trimestrielle et selon un délai de deux mois, des techniques statistiques doivent être mises en œuvre pour les mensualiser et selon le mois les estimer pour s'en servir comme déflateurs. Les indices élémentaires correspondants sont ensuite combinés pour obtenir un indice de production en volume, en s'appuyant sur des pondérations en valeur

Encadré 2. Des évolutions importantes dans les unités statistiques

L'entreprise correspond, selon le règlement européen de 1993, à la plus petite combinaison d'unités légales qui constitue une unité organisationnelle de production de biens et de services jouissant d'une certaine autonomie de décision, notamment pour l'affectation de ses ressources courantes. Une entreprise exerce une ou plusieurs activités dans un ou plusieurs lieux. Une entreprise peut correspondre à une seule unité légale.

L'unité d'activité économique (UAE) regroupe au sein d'une entreprise l'ensemble des parties qui concourent à l'exercice d'une activité du niveau classe (quatre chiffres) de la NACE. Il s'agit d'une entité qui correspond à une ou plusieurs subdivisions opérationnelles de l'entreprise. L'entreprise doit disposer d'un système d'information permettant de fournir ou de calculer pour chaque UAE au moins la valeur de la production, des consommations intermédiaires, des frais de personnel, de l'excédent d'exploitation, ainsi que l'emploi et la FBCF. Cette unité est construite pour l'observation du processus de production, pour améliorer l'homogénéité des résultats sectoriels d'enquêtes statistiques sur les entreprises, observables par branche d'activité. L'UAE n'a pas nécessairement d'existence légale. L'UAE est une unité statistique qui n'est pas forcément une unité de collecte (l'unité de collecte pouvant être l'entreprise dont elle dépend).

ajoutée comme recommandé par Eurostat pour ce type d'indicateur (**encadré 4**). Ce nouvel indicateur est à produire deux mois après le mois de référence. Il est le pendant des actuels indices de production dans l'industrie et la construction.

Pour les indices de prix à la production dans les services, les extensions de champ sont

“ Pour les indices de prix à la production dans les services, les extensions de champ sont encore plus importantes puisque les obligations liées au règlement de 1998 ne couvrent qu'environ la moitié du champ des services. ”

encore plus importantes puisque les obligations liées au règlement de 1998 ne couvrent qu'environ la moitié du champ des services. Comme la disponibilité des informations s'est considérablement améliorée par rapport au moment où les indices de prix à la production dans les services ont été introduits, avec le nouveau règle-

ment cadre qui entrera en vigueur, les indices de prix à la production dans les services seront requis pour tous les secteurs des services pour lesquels des données sur les chiffres d'affaires seront collectées. Autre innovation d'importance, le suivi des transactions sera désormais élargi à l'ensemble des transactions. En effet, le règlement de 1998 requiert des indices de prix à la production dans les services pour les transactions entre partenaires commerciaux entreprises (« B2B ») mais aucunement pour les transactions avec les autres partenaires (services vendus aux ménages, exportations de services : « B2Others »). Le nouveau règlement cadre pallie ce manque et prévoit désormais une ventilation en catégories de clients complète (« B2All », « B2B », « B2Others ») pour les indices de prix à la production dans les services, avec l'objectif de s'en servir comme déflateurs des indices de chiffre d'affaires (eux aussi en « B2All ») dans le calcul de l'indice de production dans les services.

Encadré 3 . Les indices chaînés à pondérations annuelles —————

Les indices peuvent être calculés avec des pondérations fixes sur tout ou partie d'une période (base fixe) ou avec des pondérations mises à jour annuellement associées à des techniques de chaînage d'indices. Eurostat préconise le recours à des indices chaînés à pondérations annuelles pour les indicateurs conjoncturels en volume et en prix et a publié un manuel méthodologique sur le sujet en 2012. La France a commencé à calculer des indices chaînés à pondérations annuelles en 2013 (base 2010) pour les indices de prix de production et a élargi le recours à ces techniques en 2018 pour le changement de base 2015. Cette méthodologie permet de mieux prendre en compte la déformation de structure de l'économie et la déformation des prix relatifs et de réduire nettement les problèmes de révision lors des changements de base, ce qui est souhaitable en particulier dans le cas de séries longues ou de composantes connaissant une évolution rapide. Mais cette meilleure représentation économique a un coût : l'additivité disparaît.

Sont potentiellement concernés : les indices de production dans l'industrie et dans la construction, les indices de prix de production dans l'industrie et dans la construction, les indices de prix à l'importation de produits industriels, les indices de volume des ventes dans le commerce de détail et les indices de prix de production des services.

● ... ET DES INNOVATIONS POUR LA CONSTRUCTION

Dans le secteur de la construction, le nouveau règlement cadre prévoit une inversion de logique entre coût de production et prix de production. Le règlement actuellement en vigueur inscrit comme la variable d'intérêt le coût de construction qui peut être approchée par un prix producteur. Dans le nouveau règlement cadre, ce sera le prix de production, considéré comme de meilleure mesure des pressions inflationnistes, qui peut être approximé par le coût de production. Une autre innovation pour ce secteur consistera à harmoniser les nomenclatures. La classification des produits de la construction (CC), spécifique à ce secteur, sera ainsi remplacée, selon le cas, par la nomenclature d'activité (NACE) ou de produits (CPA).

● UN DISPOSITIF FRANÇAIS ADAPTÉ AUX ÉVOLUTIONS DU RÈGLEMENT DE 1998...

En France, l'Insee et plusieurs services statistiques ministériels² contribuent au règlement sur les indicateurs conjoncturels d'entreprises.

Ils s'appuient sur un dispositif de production éprouvé, qui s'est adapté ou a su anticiper les évolutions européennes en matière de statistiques conjoncturelles, et qui notamment produisait déjà de nombreux indicateurs conjoncturels demandés par le règlement de 1998.

“ Les évolutions intervenues au niveau national, avant ou après la mise en place du règlement sur les indicateurs conjoncturels, ont principalement concerné le dispositif de collecte des prix. ”

Les évolutions intervenues au niveau national, avant ou après la mise en place du règlement sur les indicateurs conjoncturels, ont principalement concerné le dispositif de collecte des prix. En matière de prix à la production, le dispositif est assez ancien : il est l'héritier des « indices de prix de gros » calculés en

France depuis 1911 puis des « indices de prix de vente industriels ». La mensualisation des indices de prix de production dans l'industrie a été mise en œuvre en France à partir de 1989, et elle anticipait de près de dix ans les dispositions du règlement de 1998 d'une demande d'une fourniture mensuelle d'indices de prix à la production dans l'industrie au niveau européen. Ce dispositif d'observation de prix de vente sur le marché intérieur a été complété par des indices de prix à la production de l'industrie pour les marchés extérieurs à partir de 2001. En 2004, des indices de prix d'achat à l'importation de produits industriels ont été créés, à un rythme également mensuel, après une nouvelle évolution du règlement. Le suivi des prix à la production des services a débuté en 1992 et a abouti en 1994 à la publication de premiers indices de prix à la production de services restreints aux seuls échanges entre entreprises (marché « B2B »). La collecte et la diffusion de ces indices sont depuis l'origine trimestrielles.

Les autres indicateurs conjoncturels n'ont pas requis ou nécessité d'évolutions particulières du dispositif de collecte ou de production au niveau national, du moins jusqu'au début des années 2010.

2. Il s'agit notamment des services statistiques ministériels suivants : le service de la donnée et des études statistiques (SDES) au ministère de la Transition écologique et solidaire, la direction de l'animation de la recherche, des études et des statistiques (Dares) au ministère du Travail, le service de la statistique et de la prospective (SSP) au ministère de l'Agriculture et de l'alimentation (pour la partie IAA, l'agriculture ne relevant pas de ce règlement européen mais d'un accord informel).

... ET QUI ANTICIPE DEPUIS 2013 LES DISPOSITIONS DE FRIBS

Le service statistique public français s'est mis en mesure de respecter le plus tôt possible les nouvelles dispositions sur les statistiques conjoncturelles du règlement cadre FRIBS qui s'appliqueront au 1^{er} janvier 2021. Engagés dès 2013 (Gallais, 2013), les travaux d'adaptation au projet de règlement cadre seront achevés fin 2019. À cette date, s'agissant des indicateurs conjoncturels, la France sera en conformité avant même l'entrée en vigueur de FRIBS.

En matière de prix à la production dans les services, la mise à jour annuelle des pondérations

associée à des techniques de chaînage était déjà effective depuis la diffusion des indices de la base 2010 intervenue en 2013, tout comme d'ailleurs pour les prix à la production dans l'industrie et les prix à l'importation de produits industriels. En 2013, les travaux ont également été engagés pour étendre progressivement la couverture du champ des indices de prix de produc-

« En 2013, les travaux ont également été engagés pour étendre progressivement la couverture du champ des indices de prix de production dans les services. »

tion dans les services à l'ensemble des nouveaux secteurs demandés et prendre en compte les transactions avec tous les partenaires, par catégorie de clients complète, en distinguant les transactions entre partenaires commerciaux, celles avec des ménages et même les transactions avec les marchés étrangers (Gallais, 2014). Ces extensions de la couverture des indices de prix dans les services seront achevées en 2019.

Le changement de base (passage à la base 2015) intervenu en 2018 a permis de poursuivre ces innovations méthodologiques, en étendant cette fois l'usage d'indices chaînés annuellement aux indices de production industrielle, rétrogradés jusqu'en 1990 pour ces derniers, et aux indices de production dans la construction. Les pondérations sont désormais quinquennales (en 2010 et en 2015) et non plus constantes pendant toute la période d'intérêt pour les indices de chiffre d'affaires et les indicateurs de volume dérivés, les indices de production dans les services et les indices du volume des ventes dans le commerce.

Depuis fin mars 2017, les indices mensuels de chiffres d'affaires sont publiés par grand secteur d'activité et en volume dans le secteur tertiaire, avec un nouvel indice mensuel de volume des ventes dans le commerce et un nouvel indice mensuel de production dans les services (**encadré 4**), tous deux attendus au titre du nouveau règlement cadre. Tous ces indices sont désormais calculés à partir de l'exhaustivité des déclarations mensuelles de TVA et prennent en compte de manière continue la démographie d'entreprise.

S'agissant de l'indice des prix de production des logements neufs, l'Insee dispose déjà d'un tel indicateur, produit à partir des données collectées par le SDES.

Enfin, le SDES, maître d'ouvrage des enquêtes mensuelles d'activité dans le bâtiment et les travaux publics, a conduit, en lien avec la Fédération Française du Bâtiment (FFB) et la Fédération Nationale des Travaux Publics (FNTP), maîtres d'œuvre de ces enquêtes, les travaux pour faire évoluer les deux dispositifs de collecte, de manière à pouvoir calculer les indices de production dans la construction selon la nomenclature NACE et les publier lors de l'entrée en vigueur du règlement cadre.

DEUX NOUVELLES ORIENTATIONS EUROPÉENNES ABORDÉES EN 2018

Les orientations du nouveau règlement cadre ayant été définies en 2013, notamment en ce qui concerne les statistiques conjoncturelles, de nouveaux besoins se sont exprimés en 2018 au niveau européen.

Les attentes sont doubles : d'une part, en réponse à la crise immobilière américaine de 2008, doter l'Union européenne d'un suivi des différents marchés immobiliers grâce à la

Encadré 4. Un nouvel indice de production dans les services depuis 2017

L'indice de production dans les services a pour objectif de retracer les évolutions de la production dans le secteur des services comme le fait l'indice de la production dans l'industrie.

Au niveau européen, le champ de cet indice de production dans les services comprend : les transports et entreposage, l'hébergement et restauration, l'information et communication, les activités immobilières, les activités spécialisées scientifiques et techniques, les activités de services administratifs et de soutien.

Au niveau national, le champ est un peu plus large et inclut aussi les arts, spectacles et activités récréatives et les autres activités de services. Depuis 2017, l'Insee produit cet indice de production dans les services mensuellement 60 jours après la fin du mois de référence.

Un indice calculé par déflation des indices de chiffre d'affaires en valeur

À un niveau fin (classe de la NAF), les indices de production dans les services sont calculés comme le rapport de l'indice de chiffre d'affaires par un indice de prix approprié, ce qui permet de mesurer l'évolution de l'activité des entreprises à prix constants.

Les indices de chiffre d'affaires dans les services sont produits mensuellement (et disponibles à M+60) alors que les indices de prix de production dans les services le sont trimestriellement (disponibles à T+60).

Pour obtenir un indice de volume mensuel, les indices de prix de production trimestriels doivent donc être mensualisés. Pour ce faire, on associe chaque donnée trimestrielle au deuxième mois du trimestre correspondant puis on mensualise la série *via* des méthodes adaptées fondées sur la continuité des dérivées première et seconde de la fonction prix. On recale ensuite les données mensualisées de façon que la moyenne des trois mois du trimestre soit égale à la donnée trimestrielle de départ.

Compte tenu du décalage temporel des indices de prix de production dans les services, il est également nécessaire d'extrapoler les indices de prix de production pour le trimestre suivant. L'extrapolation est réalisée soit à l'aide de modèles Arima, soit à l'aide d'une régression faisant intervenir l'indice de prix à la consommation (mensuel) de la classe correspondante.

Un test de saisonnalité est ensuite effectué pour chaque série de prix de production. S'il est positif, la série fait l'objet d'une désaisonnalisation. Celle-ci s'effectue au niveau classe à l'aide de Demetra+ (méthode X12).

L'indice de volume brut (respectivement CVS) est le rapport de l'indice de chiffre d'affaires brut (respectivement CVS) et de son prix brut (respectivement CVS) associé.

Les pondérations sont calculées à partir des valeurs ajoutées de chaque secteur : l'utilisation de pondérations en valeur ajoutée plutôt qu'en production totale permet d'éviter les doubles comptes et de rendre les résultats indépendants du degré d'intégration verticale des branches. Actuellement, les pondérations sont celles de 2015. L'indice n'est pas chaîné.

disponibilité de nouveaux indicateurs portant sur l'immobilier commercial ; d'autre part, disposer d'un suivi mensuel de la production totale marchande, couvrant l'industrie, le commerce et les services et qui pourra nécessiter de disposer à terme d'indices de production et de prix dans le commerce.

❶ Suivre l'immobilier commercial :

La mise en perspective de l'immobilier commercial résulte d'initiatives multiples, notamment du Fonds monétaire international et du Financial Stability Board en 2015, du Conseil

« Ces démarches font suite à la crise immobilière américaine de 2008 qui s'est transformée en une crise bancaire et financière mondiale. »

européen du risque systémique en 2016, du Conseil Ecofin et du Forum statistique européen en 2017, d'Eurostat, de la BCE et du Comité du système statistique européen en 2018. Ces démarches font suite à la crise immobilière américaine de 2008 qui s'est

transformée en une crise bancaire et financière mondiale³. Les attentes en la matière concernent à la fois les prix (qui relèvent d'un autre règlement cadre européen que FRIBS) et les indicateurs d'activité qui, eux, relèvent du champ des indicateurs conjoncturels d'entreprise. Relayant les recommandations internationales et celles de la BCE, Eurostat prévoit d'engager des travaux pour développer la production d'indicateurs de mises en chantiers, d'achèvement des travaux de construction et de taux de vacance. Le sujet de l'immobilier commercial a été abordé pour la première fois dans le cadre des réunions européennes sur les indicateurs conjoncturels en 2018. Le SDES sera en charge du suivi de ce dossier, à la fois pour les prix et les indicateurs de volume de l'immobilier commercial.

❷ Vers un indice mensuel de production totale marchande :

Le sujet sur l'indice mensuel de production totale marchande a également été abordé pour la première fois en 2018 dans le cadre des réunions européennes sur les indicateurs conjoncturels. Eurostat a présenté une méthodologie standard pour élaborer un indice mensuel de production totale marchande, agrégation des indices de la production dans l'industrie, la construction et les services, et d'une repondération de l'indice du volume des ventes dans le commerce, en prenant comme variable de pondération la valeur ajoutée et non les chiffres d'affaires. Dans ses conclusions, Eurostat a indiqué vouloir travailler plus avant avec les États membres sur l'élaboration de cet indice mensuel de production totale marchande, et invité les États membres à procéder selon cette méthodologie standard dans un premier temps, pour dans un second temps juger de la faisabilité d'un « vrai » indice de production (et de prix selon le cas) dans le commerce fondé sur les marges commerciales. L'Insee entreprend aujourd'hui de premiers travaux pour définir une méthodologie robuste pour cet indicateur mensuel de production totale marchande.

3. Comme le précise le Comité européen du risque systémique, « le secteur de l'immobilier joue un rôle important dans l'économie et ses évolutions peuvent avoir une influence significative sur le système financier. Les crises financières passées ont montré que les évolutions non durables des marchés immobiliers pouvaient avoir de graves répercussions sur la stabilité du système financier et de l'économie dans son ensemble. Par le passé, les évolutions défavorables survenues dans certains États membres, tant sur les marchés de l'immobilier résidentiel que sur ceux de l'immobilier commercial, ont engendré de lourdes pertes et ont eu des effets négatifs sur l'économie réelle. »

BIBLIOGRAPHIE

BANQUE CENTRALE EUROPÉENNE, 2000. Les besoins de la banque centrale européenne dans le domaine des statistiques économiques générales [en ligne]. In : *site de la Banque centrale européenne*. 3 août 2000 [Consulté le 29 mai 2019]. Disponible à l'adresse : https://www.ecb.europa.eu/pub/pdf/other/statreqgen_fr.pdf?4e22fce67ebddb9043e83b8d-1c802aaf

ECOFIN, 2001. Obligations statistiques dans le cadre de l'UEM – Conclusions. In : *2382^e session du Conseil – ECOFIN* [en ligne]. 6 novembre 2001, pp. 8-11. [Consulté le 29 mai 2019]. Disponible à l'adresse : http://europa.eu/rapid/press-release_PRES-01-401_fr.pdf

GALLAIS Alain, 2013. Linking services turnover / output prices to the national macroeconomic framework. In : *28^e rencontre du groupe de Voorburg sur les statistiques des services, Tokyo, 7-11 octobre 2013* [en ligne]. [Consulté le 29 mai 2019]. Disponible à l'adresse : <https://www.voorburggroup.org/Documents/2013%20Tokyo/Papers/1011.pdf>

GALLAIS Alain, 2014. Institutional sectoring – turnover and prices for various sectors. In : *29^e rencontre du groupe de Voorburg sur les statistiques des services, Dublin, septembre 2014*. [Consulté le 29 mai 2019]. Disponible à l'adresse : <https://www.voorburggroup.org/Documents/2014%20Dublin/Papers/018.pdf>

GLAUDE Michel, 2007. Enjeux et défis de la statistique européenne. In : *Courrier des Statistiques* [en ligne]. Mai-décembre 2007. N°121-122, pp. 19-22. [Consulté le 29 mai 2019]. Disponible à l'adresse : <http://www.epsilon.insee.fr/jspui/bitstream/1/8295/1/cs121-122d.pdf>

FONDEMENTS JURIDIQUES

RÈGLEMENTS (CE). [en ligne]. [Consultés le 23 avril 2019]. Disponibles à l'adresse : <http://data.europa.eu/eli/reg>

- ❶ N°1165/1998 du Conseil du 19 mai 1998 concernant les statistiques conjoncturelles. 5 juin 1998.
- ❶ N° 1882/2003 du Parlement européen et du Conseil du 29 septembre 2003 portant adaptation à la décision 1999/468/CE du Conseil des dispositions relatives aux comités assistant la Commission dans l'exercice de ses compétences d'exécution prévues dans des actes soumis à la procédure visée à l'article 251 du traité CE. 29 septembre 2003.
- ❶ N° 1158/2005 du Parlement européen et du Conseil du 6 juillet 2005 modifiant le règlement (CE) n° 1165/98 du Conseil concernant les statistiques conjoncturelles. 6 juillet 2005.
- ❶ N° 1503/2006 de la Commission du 28 septembre 2006 relatif à l'application et à la modification du règlement (CE) n° 1165/98 du Conseil sur les statistiques conjoncturelles en ce qui concerne la définition des variables, la liste des variables et la fréquence d'élaboration des données. 28 septembre 2006.

- ❶ N° 1893/2006 du Parlement européen et du Conseil du 20 décembre 2006 établissant la nomenclature statistique des activités économiques NACE Rév. 2 et modifiant le règlement (CEE) n° 3037/90 du Conseil ainsi que certains règlements (CE) relatifs à des domaines statistiques spécifiques Texte présentant de l'intérêt pour l'EEE. 20 décembre 2006.
- ❶ N° 1178/2008 de la Commission du 28 novembre 2008 modifiant le règlement (CE) n° 1165/98 du Conseil concernant les statistiques conjoncturelles et les règlements (CE) n° 1503/2006 et (CE) n° 657/2007 de la Commission en raison des adaptations effectuées après la révision des nomenclatures statistiques Nace et CPA (Texte présentant de l'intérêt pour l'EEE). 28 novembre 2008.
- ❶ N° 329/2009 de la Commission du 22 avril 2009 modifiant le règlement (CE) n° 1165/98 du Conseil concernant les statistiques conjoncturelles en ce qui concerne la mise à jour de la liste des variables, la fréquence de l'élaboration des statistiques et les niveaux de ventilation et d'agrégation à appliquer aux variables (Texte présentant de l'intérêt pour l'EEE). 22 avril 2009.
- ❶ N° 596/2009 du Parlement européen et du Conseil du 18 juin 2009 portant adaptation à la décision 1999/468/CE du Conseil de certains actes soumis à la procédure visée à l'article 251 du traité, en ce qui concerne la procédure de réglementation avec contrôle Adaptation à la procédure de réglementation avec contrôle – Quatrième partie. 18 juin 2009.
- ❶ N° 461/2012 de la Commission du 31 mai 2012 modifiant le règlement (CE) n° 1165/98 du Conseil concernant les statistiques conjoncturelles et les règlements (CE) n° 1503/2006, (CE) n° 657/2007 et (CE) n° 1178/2008 de la Commission en ce qui concerne les adaptations relatives à la suppression des variables sur les entrées de commandes industrielles Texte présentant de l'intérêt pour l'EEE. 31 mai 2012.

LE PROFILAGE À L'INSEE

UNE IDENTIFICATION PLUS PERTINENTE DES ACTEURS ÉCONOMIQUES

Olivier Haag*

En statistique, la notion d'entreprise a souvent été associée à sa définition purement juridique, i.e. l'unité légale inscrite au répertoire Sirene. Cette approche permet de bénéficier de nombreuses données administratives et réduit de fait la charge d'enquête qui pèse sur les entreprises. Toutefois, avec l'émergence des groupes de sociétés, les unités légales détenues par d'autres peuvent perdre tout ou partie de leur autonomie et donc leur statut d'entreprise. On nomme « profilage » la technique consistant à identifier au sein des groupes les entreprises au sens économique, puis à collecter et calculer des statistiques sur ces nouveaux contours. Il est aujourd'hui appliqué à l'Insee pour les groupes ayant un poids économique important en France. Cela requiert une collaboration directe entre le groupe et un « profileur », nouveau métier apparu au début des années 2000. Les monographies sont indispensables pour obtenir des statistiques robustes sur ces grands groupes. Cependant, coûteuses en moyens humains, elles ne pourraient être appliquées aux 100 000 groupes¹ présents sur le sol français. C'est pourquoi une nouvelle méthode de profilage dit « automatique » a été élaborée par l'Insee et mise en œuvre pour la première fois en 2018.

 *In statistics, the notion of enterprise has often been associated with its purely legal definition, i.e. the legal unit entered in the Sirene register. This approach will allow the benefit from a large amount of administrative data and effectively reduces the statistical burden on companies. However, with the emergence of corporate groups, legal units owned by others may lose all or part of their autonomy and therefore their status of enterprise. « Profiling » is the technique of identifying enterprises in the economic sense within groups and then collecting and calculating statistics on these new profiles. It is now applied at Insee to groups with significant economic weight in France. This requires direct collaboration between the group and a « profiler », a new profession appeared at Insee in the early 2000s. These monographs are essential to obtain robust statistics on these major groups. However, since they are costly in terms of human resources, they could not be applied to the 100,000 groups² present in France. This is why a new so-called « automatic » profiling method has been developed by Insee and implemented for the first time in 2018.*

* Coordinateur des projets structurants de la Direction des statistiques d'entreprises (DSE), Insee, olivier.haag@insee.fr

1. Un groupe de sociétés, au sens statistique, est un ensemble d'unités légales, le plus souvent liées entre elles par des liens financiers, de telle sorte qu'elles sont toutes contrôlées à plus de 50 %, directement ou indirectement, par l'une d'entre elles (appelée tête de groupe), cette dernière n'étant pas contrôlée par une autre. La tête de groupe peut être située à l'étranger.

2. A group of companies, in the statistical sense, is a set of legal units, most often linked together by financial ties, so that they are all more than 50 % directly or indirectly controlled by one of them (called the group head), the latter not being controlled by another. The group leader can be located abroad.

La plupart des noms des sociétés du CAC40 et plus généralement d'entreprises qui font quotidiennement la une des journaux sont en réalité des noms de « groupes ». Mais les groupes, qui réalisent les deux tiers de la valeur ajoutée des entreprises³ - c'est dire leur importance - n'ont pas de véritable existence juridique et ne constituent pas des personnes morales. C'est la raison pour laquelle, jusqu'à des années récentes, la statistique publique les ignorait en pratique et ne s'appuyait dans la plupart des pays que sur les unités légales (sociétés et entreprises individuelles), même si de plus en plus de textes légaux ou réglementaires y faisaient référence. Ceci avait pour conséquence de prendre en compte, dans les statistiques, des transactions entre unités du même groupe qui la plupart du temps ne respectent pas les règles du marché et n'ont donc pas une réelle signification économique. Constatant qu'il n'était plus pertinent de nos jours de se baser uniquement sur les unités légales pour répondre aux exigences réglementaires françaises ou européennes relatives aux statistiques d'entreprises, l'Insee a donc décidé de diffuser des statistiques structurelles fondées sur des entreprises au sens économique, dont la définition prend en compte la notion de groupe. Procéder ainsi nécessite de concevoir et de mettre en application une méthodologie élaborée : le profilage.

Suite à des réflexions générales sur le profilage dans les années 90, les premiers travaux à visée opérationnelle ont été conduits en France en 2006 et 2007 *via* un groupe de travail du Cnis (Depoutot, 2008). Ceci a conduit l'Insee, dans le cadre de la LME⁴, à prendre la décision de modifier à terme la définition des entreprises dans le système statistique public français (**encadré 1**).

Ces préoccupations sont également présentes au niveau européen, sous l'égide d'Eurostat, qui a entrepris une vaste refonte de l'ensemble des règlements qui régissent la statistique d'entreprise européenne, sous la forme du règlement cadre « FRIBS⁵ », désormais adopté : un groupe de travail en réseau (ESSnet) a ainsi été lancé en 2014, pour une durée de deux ans, sur le sujet du profilage et de la définition de « l'entreprise ».

À cette occasion, Eurostat a constaté que la définition officielle de « l'entreprise » était appliquée de façon souvent imparfaite par nombre de pays. La définition issue du règlement européen (N° 696) de 1993, sera reprise ensuite dans la LME française en 2008 pour définir les catégories d'entreprises. En 2015, Eurostat a ainsi demandé aux pays de se conformer à cette définition pour l'élaboration des statistiques structurelles et d'élaborer un plan d'actions pour y parvenir au plus tard en 2020.

Encadré 1. Définition de l'entreprise au sens de la Loi de Modernisation de l'Économie du 4 août 2008

L'entreprise correspond à la plus petite combinaison d'unités légales qui constitue une unité organisationnelle de production de biens et de services jouissant d'une certaine autonomie de décision, notamment pour l'affectation de ses ressources courantes. Une entreprise exerce une ou plusieurs activités dans un ou plusieurs lieux. Une entreprise peut correspondre à une seule unité légale.

3. Voir (Béguin, Hecquet, 2015) et (Béguin, Hecquet, Lemasson, 2012).

4. Loi de Modernisation de l'Économie du 4 août 2008.

5. *Framework Regulation Integrating Business Statistics (FRIBS)*.

● LES PREMIÈRES EXPÉRIENCES DATENT DE LA FIN DES ANNÉES 90

La nécessité de « profiler »⁶ certains groupes est apparue dès la fin des années quatre-vingt-dix, à l'occasion d'une restructuration juridique des deux grands constructeurs automobiles français qui avaient décidé, à peu près à la même date, de donner la personnalité morale à chacun de leurs établissements de production. Ainsi de nouveaux flux monétaires, sans réelle consistance économique, sont apparus entre les unités légales de ces groupes, correspondant aux flux physiques des composants des voitures. Dans cette restructuration, les unités légales de production louaient la main-d'œuvre et les machines à certaines unités légales du même groupe et elles vendaient leur production à d'autres unités légales du groupe en charge de la commercialisation. Toutes ces nouvelles unités légales n'avaient pas de réelle autonomie de production, ni de décision, les décisions continuant bien évidemment d'être prises au niveau du groupe. Elles ne satisfaisaient donc pas aux critères de définition classiques d'une « entreprise » et n'étaient pas des « acteurs économiques » à part entière.

“ Ce genre de décision, consistant à créer plusieurs unités légales au sein d'une nouvelle entité, peut se produire à tout instant au sein des grands groupes. ”

La restructuration mentionnée ci-dessus conduisait à **multiplier « fictivement » par près de trois** le chiffre d'affaires, la sous-traitance, etc., de la filière automobile française alors qu'en réalité rien n'avait changé dans les capacités de production du fait de cette restructuration, de sorte qu'il n'était pas possible de laisser ces nouvelles données telles

quelles si l'on voulait se ramener à une situation économiquement significative.

Ce genre de décision, consistant à créer plusieurs unités légales au sein d'une nouvelle entité, peut se produire à tout instant au sein des grands groupes. L'assimilation de l'unité légale à l'entreprise n'est donc plus pertinente dans ce genre de cas : autonomie de décision et gestion de l'ensemble des facteurs de production sont au niveau du groupe et non plus au niveau des unités légales. De plus, continuer à raisonner en unités légales modifie non seulement les niveaux des agrégats mais également leur répartition sectorielle.

● L'INTÉRÊT DU PROFILAGE: UN EXEMPLE ILLUSTRATIF

On peut illustrer la démarche suivie par l'exemple fictif qui suit (**figure 1**), mettant en jeu une entreprise constituée d'une seule unité légale. Celle-ci fait l'année N un chiffre d'affaires de 300 k€ et crée une valeur ajoutée de 200 k€. Elle a plusieurs activités dans son processus de production et on note A son APE⁷ que l'on suppose industrielle.

Imaginons maintenant que pour des raisons fiscales les dirigeants de cette entreprise décident de créer en N+1 une structure de groupe dans laquelle il y a autant d'unités légales que d'activités différentes dans le processus de production initial. À fonctionnement économique constant, il se crée donc quatre unités légales au sein du groupe G :

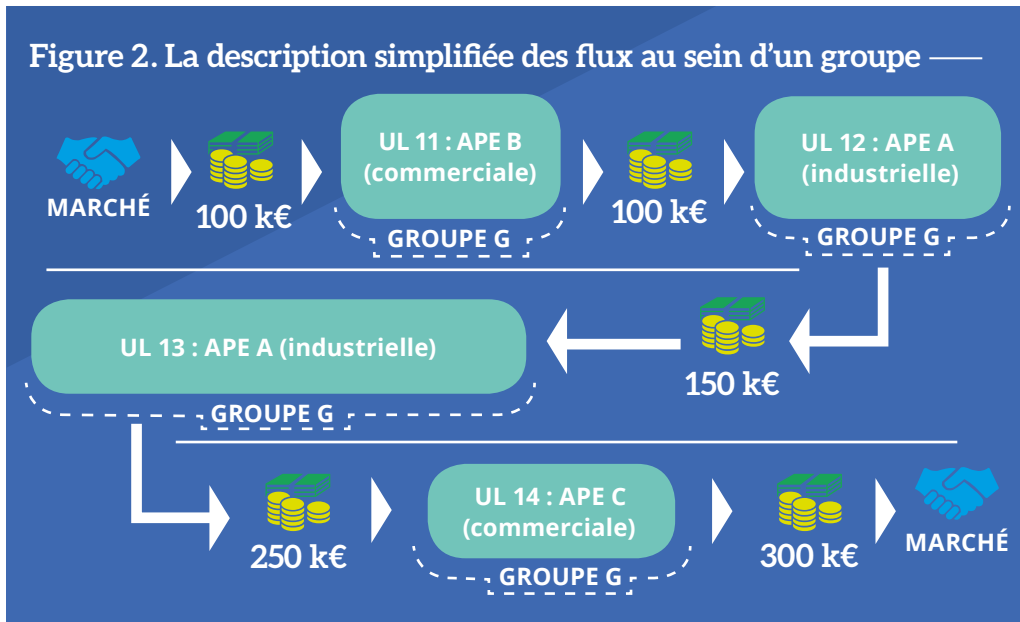
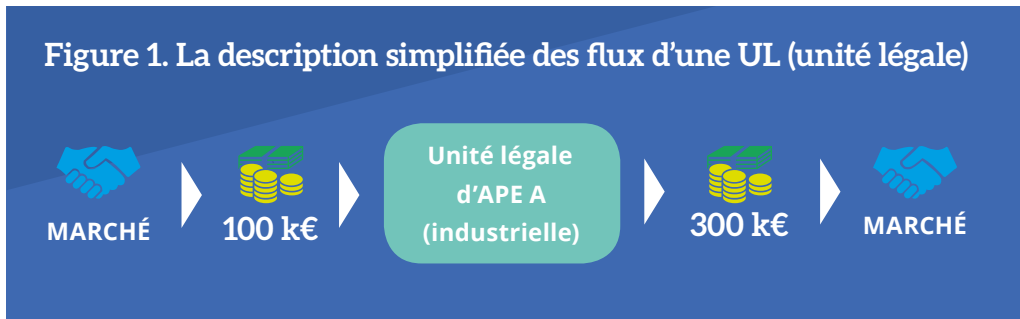
- UL11 est en charge de l'approvisionnement en matières premières et revend l'ensemble de ses achats au prix coûtant à une société industrielle ; elle a donc une APE commerciale B.

6. Il s'agit de l'action menée par des membres de l'INS (profileur) pour définir les entreprises au sein des groupes et collecter des données sur ces unités.

7. APE = Activité Principale Exercée.

- UL12 est un sous-traitant exclusif d'une autre unité légale industrielle (UL13) dont l'activité est le cœur de métier du groupe, c'est-à-dire l'activité A ;
- UL13 vend l'ensemble de sa production à une société commerciale du groupe (UL14) qui se charge de commercialiser les produits sur le marché.
- UL14 a donc une APE commerciale C.

L'APE du groupe reste industrielle : ce sont les activités industrielles qui génèrent le plus de valeur ajoutée⁸. Avec l'hypothèse simplificatrice qu'entre les deux années il n'y a aucune modification de l'activité économique de la société, on peut aboutir au résultat de la **figure 2** alors que rien ne change en réalité sur le plan économique, le fait de raisonner en unités légales (elles-mêmes modifiées) bouleverse le paysage statistique.



Voici quelques conséquences :

- faire passer le chiffre d'affaires de 300 lorsque l'unité légale était indépendante à 800 à cause des flux intragroupe sans qu'il y ait de modification de l'activité économique. Ce dernier point est d'ailleurs observable au niveau de la valeur ajoutée qui ne change pas

8. L'algorithme de calcul de l'activité principale se fonde en effet sur les valeurs ajoutées respectives générées par les différentes activités.

entre les deux années. La valeur ajoutée est en effet une variable additive qui est censée être indépendante de la structure des unités (mais ce n'est plus vrai si on s'intéresse à la ventilation de la valeur ajoutée par secteur) ;

- ❶ faire apparaître de nouvelles activités (une industrielle et deux commerciales) ;
- ❷ modifier les répartitions sectorielles du chiffre d'affaires mais aussi de la valeur ajoutée.

Au contraire, le fait de considérer « le groupe dans son ensemble » comme étant « une entreprise » permet de retomber sur la situation précédente qui, par hypothèse, n'a pas changé. En creux, notre exemple permet de mettre en évidence les modifications engendrées par le profilage des groupes :

- ❸ disparition des activités secondaires, internes au groupe ;
- ❹ réduction du chiffre d'affaires total (et des variables non additives⁹ en général) à cause de la suppression des flux intragroupe ; ceci est légitime dans la mesure où ces derniers ne respectent pas les règles du marché (les prix utilisés pour ces échanges communément appelés prix de transfert sont définis arbitrairement par le groupe sans avoir forcément un lien avec les prix en cours sur le marché pour le même type d'échanges) ; on suivra donc désormais des comptes « consolidés » ;
- ❺ conservation du niveau total de la valeur ajoutée (seule sa ventilation sectorielle sera modifiée).

Fort de ce constat, il a donc été décidé de prendre en compte la notion de groupe dans l'élaboration des statistiques d'entreprises *via* la méthode du profilage. Celle-ci recouvre essentiellement deux activités bien distinctes : la détermination du contour de ces « entreprises », et le calcul des statistiques sur les unités ainsi construites.

❶ UNE APPROCHE TOP-DOWN POUR DÉFINIR DES ENTREPRISES AU SENS ÉCONOMIQUE

Pour une année donnée, les contours¹⁰ de ces entreprises économiques sont définis à partir des contours des groupes contenus dans le répertoire Lifi (Liaisons Financières)

“ La dimension nationale des statistiques introduit une complexité supplémentaire. ”

(Mariotte, 2017). Le profilage d'un groupe commence par l'opération de délimitation d'une ou plusieurs entreprises au sens économique au sein du groupe (plusieurs lorsque celui-ci est constitué de plusieurs unités opérationnelles productrices de biens et services, jouissant d'une certaine autonomie).

La **figure 3** présente la situation qui prévaut, sur un exemple réel mettant en jeu 10 unités légales (UL1 à UL10) du répertoire Sirene : on y trouve les groupes (G1 et G2) constitués dans le répertoire Lifi à partir des liens financiers entre unités légales, et les entreprises (E1 à E6) créées par les profileurs dans la BCE (Base de création d'entreprise).

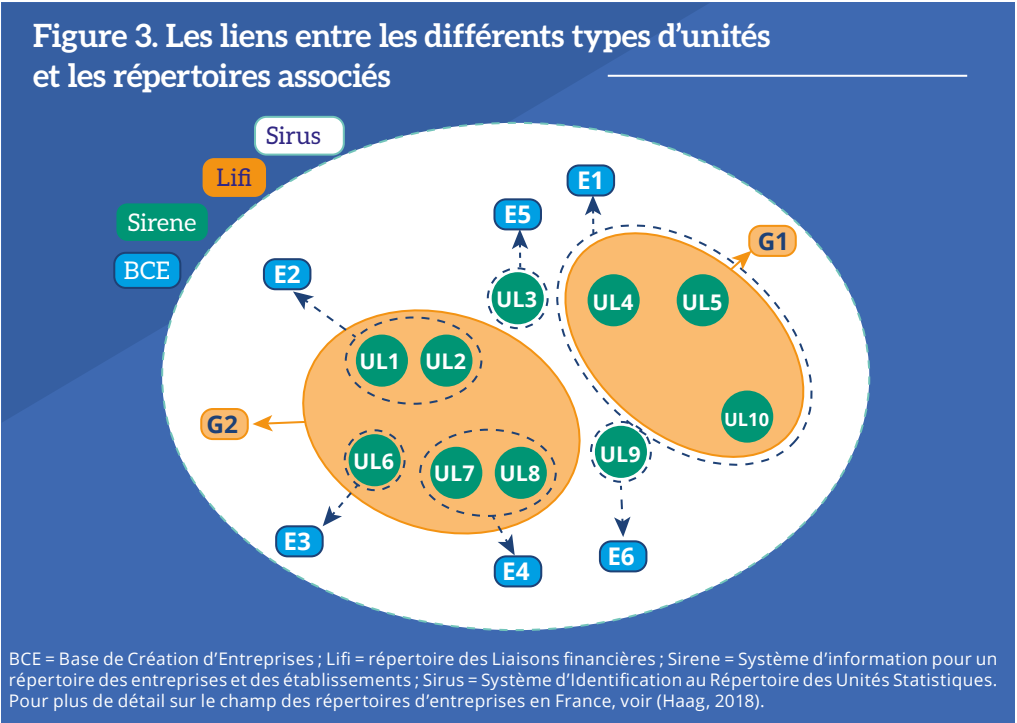
9. Les variables non-additives (chiffre d'affaires et sa ventilation en branches, achats, dividendes, titres de participations, capitaux propres, créances et dettes, etc.) résultent d'algorithmes de consolidation (voir plus loin).

10. Le contour dans ce contexte est la liste des unités légales qui composent l'entreprise.

La dimension nationale des statistiques introduit une complexité supplémentaire : l'élaboration des statistiques requiert de considérer comme entreprise en France un ensemble d'unités légales en France. Dès lors, dans le cas d'une entreprise au sens économique qui dépasse les frontières nationales¹¹, on est obligé de se limiter à la trace française de cette dernière. On entend par « **trace française** » la simple restriction de l'activité au sol national, indépendamment de la nationalité du ou des propriétaires (actionnaires par exemple). Une entreprise au sens économique en France sera donc au final :

- soit une seule unité légale non contrôlée par une autre unité légale (UL3 = E5 et UL9 = E6) ;
- soit la trace française d'un groupe d'unités légales (G1 = E1) ;
- soit une partie de la trace française d'un groupe d'unités légales (UL1 + UL2 = E2, UL6 = E3 et UL7 + UL8 = E4, 3 entreprises composant le groupe G2).

On notera que dans cette restriction au sol national, on peut perdre une partie de « l'autonomie » censée caractériser l'entreprise en tant qu'acteur économique. On touche là, de fait, à la difficulté de l'économie mondialisée actuelle : il n'est pas possible de décrire fidèlement son fonctionnement en se fondant exclusivement sur des « acteurs nationaux ». Les ensembles (nationaux) que l'on considère à la fin et qui sont utilisés pour produire des statistiques nationales (dont le PIB) n'ont pas toujours le caractère d'autonomie que l'on souhaiterait et qui caractérise normalement une « entreprise ». Mais seul un niveau supranational (européen, voire mondial) permettrait de ne prendre en compte que de véritables acteurs économiques indépendants. C'est d'ailleurs l'objectif du profilage européen¹². Ce dernier est piloté par Eurostat et la France y contribue grandement.



11. Ce qui arrive si cette « entreprise » appartient à un groupe multinational. C'est d'ailleurs une situation fréquente.
12. Le profilage européen repose sur le même principe que le profilage national, mais la trace est cette fois européenne (Xirouchakis, Hecquet, 2018).

Toutefois il n'est à l'heure actuelle qu'au stade expérimental et aucune statistique n'est produite à ce jour à partir des « *Global Enterprises* » (nom des unités statistiques définies par le profilage européen).

❶ UNE APPROCHE BOTTOM-UP POUR CALCULER DES STATISTIQUES SUR CES NOUVEAUX CONTOURS

Dans de nombreux Instituts Nationaux de Statistique (INS), le profilage se limite à cette première étape de définition et d'enregistrement des contours dans les répertoires nationaux.

Dans d'autres INS (l'Insee mais aussi les instituts néerlandais, canadien¹³, australien¹⁴), on considère que le profilage n'a de sens que si on l'utilise pour élaborer des statistiques structurelles. Ainsi, après avoir identifié le contour de l'entreprise¹⁵, il convient de :

- ❶ lui attribuer des caractéristiques du répertoire statistique (identifiant, APE, catégorie, etc.);
- ❷ calculer des données consolidées sur son périmètre ;
- ❸ disposer d'une réponse à son niveau pour les enquêtes où elle a été identifiée comme unité statistique. Cette réponse peut être obtenue soit directement au niveau de l'entreprise soit en agrégeant les réponses obtenues au niveau des unités légales qui la composent.

Dans le cadre du dispositif ESANE¹⁶ la méthode utilisée pour le calcul des caractéristiques fiscales et d'enquête est une méthode de type bottom-up (**figure 4**). Ainsi, l'unité de collecte est l'unité légale et les valeurs des caractéristiques au niveau de l'entreprise sont obtenues par consolidation des données recueillies au niveau des unités légales¹⁷. Il faut cependant distinguer deux situations :

- ❶ Les variables dites « additives » (effectif, valeur ajoutée, etc.) : la valeur de la variable au niveau entreprise est égale à la somme des valeurs de cette variable pour toutes les unités légales qui la composent. Par exemple l'effectif de l'entreprise est la somme des effectifs de ses unités légales.
- ❷ Les variables dites « non additives » (chiffre d'affaires, achats, créances, dettes, dividendes, titre de participation) : il est nécessaire de retirer du cumul des unités légales, les flux internes à l'entreprise. On verra dans l'**encadré 2** l'exemple du chiffre d'affaires.

❷ PLUSIEURS TYPES DE PROFILAGE

Différentes méthodes peuvent être mises en œuvre. Elles vont du profilage « manuel » ou « sur mesure », qui se traduit par une monographie du groupe par un profileur, à un profilage totalement automatique, qui ne demande aucune intervention humaine *a priori*. Entre ces deux méthodes, il existe des méthodes alternatives dites semi-automatiques plus ou moins sophistiquées.

En France, les résultats du profilage manuel sont pris en compte dans le calcul des statistiques structurelles (ESANE) depuis l'exercice 2013 publié en juin 2015. Ceux du profilage automatique le seront pour la première fois pour l'exercice 2017 publié en juin 2019.

Les méthodes semi-automatiques ne sont qu'exploratoires pour le moment.

13. Voir (Willeboordse, 1997), (Brudieu, 1999) et (Hoogsteen, 2016).

14. Voir (Pietsch, 1995) et (Rivière, 1999), ainsi que le site de l'INS australien ABS :

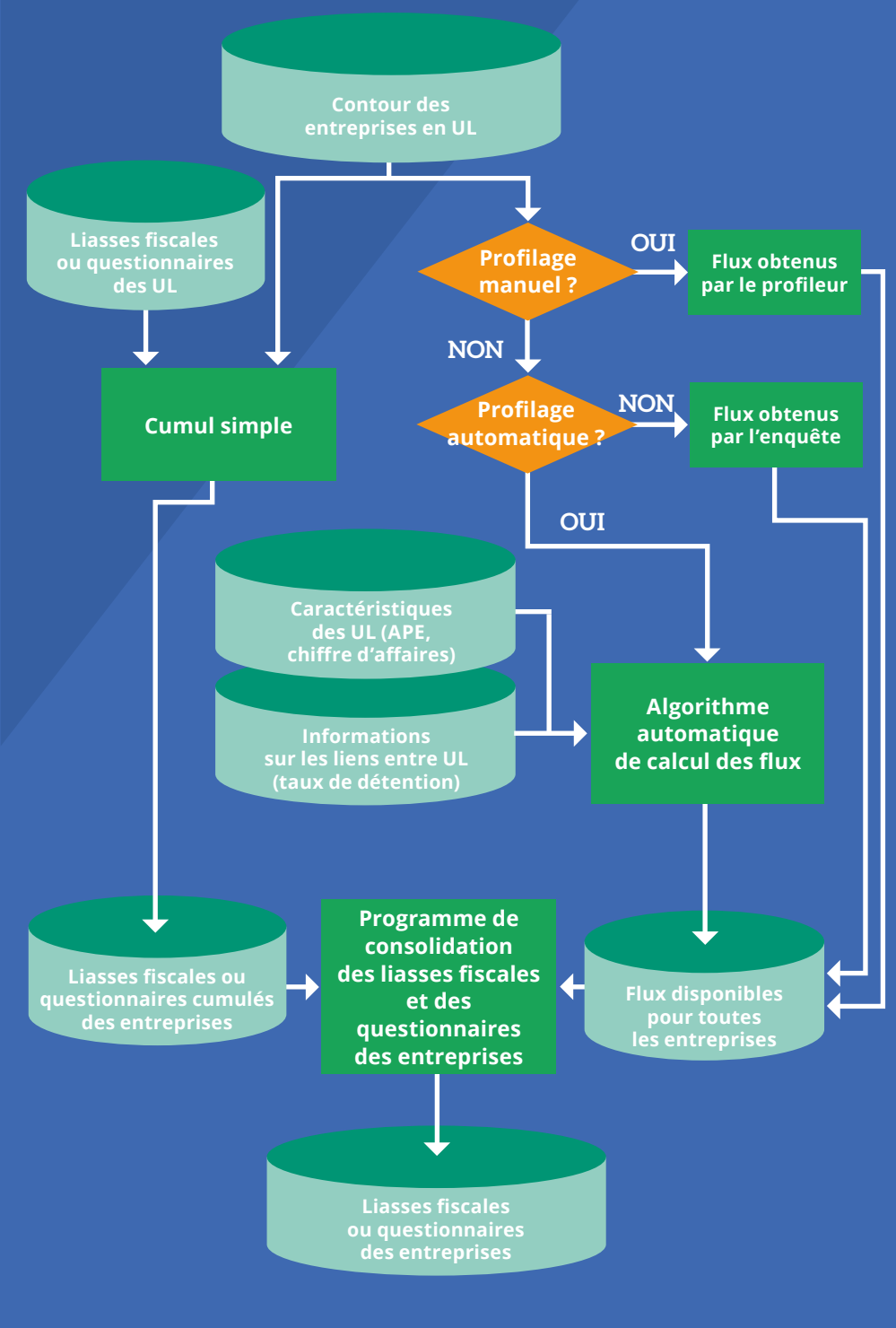
<http://www.abs.gov.au/ausstats/abs@.nsf/dossbytitle/AC79D33ED6045E88CA25706E0074E77A?OpenDocument>

15. Voir (Smith, 2013).

16. Élaboration des Statistiques Annuelles d'Entreprises.

17. Par données on entend ici les liasses fiscales, les données administratives sur l'emploi ou les réponses aux enquêtes statistiques.

Figure 4. Description de la méthode bottom-up



Le profilage manuel est ainsi le plus avancé. Il concerne aujourd'hui une cinquantaine des plus grands groupes (en termes de valeur ajoutée ou d'emploi) présents sur le sol français. La cible est d'atteindre 70 groupes d'ici 3 ans. Réalisé au sein d'une unité dédiée de l'Insee¹⁸, il a fait émerger de nouvelles façons de travailler et même un nouveau métier : profileur.

🕒 LE PROFILAGE MANUEL : PROFESSION PROFILEUR

Le travail des profileurs consiste principalement à prendre contact avec les groupes et à les rencontrer pour obtenir de leur part :

- 🕒 les contours en unités légales des entreprises au sein du groupe,
- 🕒 les données comptables pour ces entreprises,
- 🕒 la réponse à l'enquête sectorielle annuelle (ESA) pour ces entreprises.

Une fois ces données collectées, les profileurs doivent mettre en forme, contrôler et mettre en cohérence ces informations avec les données des unités légales dont ils disposent par ailleurs. Ce travail de validation est complexe vu le nombre d'unités légales en jeu et la difficulté dans certains cas d'obtenir, dès le premier envoi, des données fiables de leurs interlocuteurs.

“ Une fois ces données collectées, les profileurs doivent mettre en forme, contrôler et mettre en cohérence ces informations avec les données des unités légales dont ils disposent par ailleurs. ”

Ces activités, très originales pour l'Insee, demandent des compétences spécifiques : l'équipe comprend ainsi des profils très divers, attachés ou administrateurs de l'Insee mais aussi contractuels ayant travaillé dans le secteur privé (anciens comptables, ou

consolidateurs, etc.). Ces expériences de la comptabilité et du secteur privé permettent non seulement de communiquer plus facilement avec les groupes en parlant « leur langage » mais aussi de proposer des algorithmes de consolidation plus robustes.

🕒 DÉFINIR DES CONTOURS : APPROCHER AU MIEUX LE CRITÈRE D'AUTONOMIE

La première activité du profileur consiste à identifier, au sein des groupes, les sous-groupes qui respectent la définition de l'entreprise au sens de la LME. La notion la plus difficile à appréhender et que seul le groupe peut nous fournir de façon sûre est celle d'autonomie. Or celle dont ils disposent n'est pas absolue : un sous-groupe de sociétés ne constituera une entreprise que s'il jouit d'une « certaine » autonomie.

L'Insee s'est doté de critères pour l'objectiver. Cette restriction relative sur l'autonomie fait écho à la possession des titres des sociétés par la maison-mère : une entreprise ne disposera jamais (quasiment) de la propriété intellectuelle et des financements longs, ainsi que de la capacité de décider des investissements lourds (à l'échelle du groupe), toutes choses qui caractérisent le pouvoir d'une holding. L'autonomie dont nous parlons ici, c'est la possibilité de prendre toutes les autres décisions, relevant pour la plupart de la gestion quotidienne, ou à plus ou moins court terme.

18. La division Profilage et Traitement des Grandes Unités (PTGU).

Ainsi, pour qu'un sous-groupe de sociétés puisse constituer une entreprise, il faut d'abord qu'il ait les moyens de prendre ses décisions : dispose-t-il des effectifs nécessaires à son activité de production, de négoce ou de service, d'un directeur général en propre, de moyens immobiliers, etc. ? Il faut aussi qu'il forme un ensemble cohérent économiquement, donc apte à fonctionner dans un environnement économique, et que son activité soit majoritairement orientée vers le marché et non vers le groupe. Enfin il est indispensable que le groupe soit en mesure de fournir des données fiscales et statistiques sur ces ensembles. En effet, s'il est fréquent que les groupes disposent dans leur système d'information de données, ces dernières ne sont pas toujours déclinées sur le périmètre français du sous-groupe d'une part et elles ne sont pas disponibles au format du plan comptable général d'autre part. Les groupes doivent donc souvent mettre en place un système d'information *ad hoc* pour répondre aux besoins de l'Insee.

POUR OBTENIR DES DONNÉES DE QUALITÉ, L'IMPORTANCE DU DIALOGUE

Une fois les contours des entreprises définis, le profileur va collecter auprès du groupe les informations nécessaires pour constituer une liasse fiscale et un questionnaire ESA consolidé pour chacune des « entreprises ». Il devra aussi obtenir du groupe un ensemble de flux intragroupe qu'il devra soustraire du cumul des unités légales du groupe pour les variables non additives (**encadré 2**).

Pour les données fiscales, une liste de flux a été définie. Elle concerne :

-  le compte de résultat (ventes vs achats notamment) ;
-  le bilan (créances vs dettes, emprunts vs prêts) ;
-  les provisions ;
-  les résultats internes (dividendes-cessions immobilisations) ;
-  les titres de participations internes.

Pour les données de l'ESA, il s'agit de récupérer le flux de chiffre d'affaires ventilé par branche d'activité ainsi que la réponse aux questions sectorielles de l'ESA. Ce type d'information est complexe à récupérer car l'interlocuteur des profileurs au sein du groupe se situe le plus souvent dans la sphère financière et non dans la partie productive et il n'en dispose donc pas.

Dans certains cas, les profileurs jouent également un rôle d'intermédiaire pour faciliter la collecte d'enquêtes thématiques, lorsqu'une réponse au niveau de l'entreprise a été négociée (enquêtes TIC¹⁹ et CIS²⁰ par exemple).

LE PROFILAGE AUTOMATIQUE : COÛT MOINDRE, HYPOTHÈSES FORTES...

Contrairement au profilage manuel, le profilage automatique ne relève pas intégralement des profileurs. L'unité en charge du profilage gère certes la définition des contours, mais l'examen des données issues de la consolidation automatique est réalisé par le service chargé des enquêtes structurelles.

19. Enquêtes sur les Technologies de l'Information et de la Communication.

20. Enquête communautaire sur l'innovation.

Encadré 2. La consolidation du chiffre d'affaires (Chanteloup, 2018)

La consolidation du chiffre d'affaires est régie par un principe de précaution qui vise à éviter autant que faire se peut une sur-consolidation. En effet, si la diffusion en unités légales conduit à une surestimation du chiffre d'affaires total (tous les chiffres d'affaires intra-groupe sont pris en compte), il faut à l'inverse éviter que la prise en compte de l'entreprise se traduise par une sous-estimation du chiffre d'affaires qui serait la traduction d'une suppression à tort de chiffre d'affaires lors de la phase de consolidation automatique.

Pour respecter ce postulat, les principes de base qui régissent cet algorithme sont les suivants :

- En cas de doute, quand aucune hypothèse logique n'est possible, aucune consolidation n'est mise en œuvre, car on ne dispose d'aucune source sur les échanges intra-groupes.
- Lorsque l'on fait l'hypothèse qu'il y a un lien d'échange potentiel entre deux unités légales au sein du groupe, on consolide au maximum des capacités d'achat de l'unité en aval de la relation et des capacités de production de l'unité en amont (c'est-à-dire au minimum de ces deux capacités). Cette décision résulte du fait qu'il n'existe aucune source disponible sur l'ensemble des unités légales appartenant à un groupe fournissant la distinction entre les échanges intra-groupe et ceux qui transitent par le marché.
- La totalité du chiffre d'affaires des unités légales classées en auxiliaire est consolidé. On suppose que ces unités n'accèdent pas au marché.

Le calcul d'un chiffre d'affaires consolidé s'effectue par étapes successives.

- **Étape 1 :** pour chaque unité légale, on détermine si elle est auxiliaire (ou intra-groupe), commerciale ou productive, à partir de son APE.
 - Les unités légales dont l'APE les classe en commerce sont considérées comme commerciales.
 - On considère comme auxiliaires, ou intra-groupe, les unités légales classées en « holding » ou « sièges sociaux » ou « autres activités de soutien aux entreprises », et celles de tout autre secteur des services^a lorsque l'ensemble des UL du groupe et de ce secteur abrite moins de 10 % des effectifs salariés du groupe.
 - Les autres unités légales sont qualifiées de productives.
- **Étape 2 :** on consolide le chiffre d'affaires des unités productives et commerciales uniquement. Pour ce faire, des couples d'activités principales susceptibles d'échanger entre eux ont été identifiés, divisés en trois types de liens :
 - des unités productives amont et aval^b. Dans ce cas on consolide au maximum le chiffre d'affaires (les ventes) de l'unité légale en amont avec les achats de matières premières de l'unité légale en aval.
 - Des unités commerciales de gros et de détail^c. Dans ce cas on consolide au maximum le chiffre d'affaires de l'entreprise de commerce de gros avec les achats de marchandises de l'unité légale du commerce de détail.
 - Des unités productives qui vendent (respectivement achètent) aux commerciales^d. Dans ce cas on consolide au maximum le chiffre d'affaires (respectivement achats de matières premières) de la productive avec les achats (respectivement ventes) de marchandises de la commerciale.
- **Étape 3 :** enfin, on consolide le chiffre d'affaires des unités restantes, i.e. les unités auxiliaires, avec les sous-groupes consolidés dans l'étape 2, qui regroupaient déjà toutes les unités commerciales et productives.

a. Activités auxiliaires de service financier et assurance, programmation de conseil et autres activités informatiques, transport routier de fret de proximité, formation continue des adultes, etc.

b. Il s'agit d'une intégration verticale à savoir que la production de l'unité légale amont constitue un input pour la production de l'unité légale en aval. On peut citer l'exemple d'une unité légale qui fabrique des lacets et qui vend sa production à une unité légale du groupe qui fabrique des chaussures.

c. Il s'agit par exemple du cas de centrales d'achats qui achètent en gros et revendent en l'état cette marchandise à des détaillants du même groupe.

d. On peut citer l'exemple d'un groupe qui décide de créer une filiale spécifique pour centraliser toutes ses exportations.

Le point de départ pour constituer les entreprises au sens économique est le noyau dur du groupe tel que défini par le répertoire Lifi. On y trouve l'ensemble des unités légales détenues à plus de 50 % par la tête de groupe.

Comme il est impossible d'apprécier, *via* un algorithme, le critère d'autonomie, il a été décidé, dans un premier temps, pour le profilage automatique, de considérer l'ensemble du groupe comme une unique entreprise : c'est là une hypothèse forte, mais nécessaire en pratique (Chanteloup, 2017). Il a également été décidé de poser une seconde hypothèse : pour définir les contours des entreprises, on peut se restreindre aux unités légales des groupes qui appartiennent au champ des statistiques structurelles d'entreprises de l'Insee (ESANE). Sans cela, le risque eût été d'introduire des incohérences, avec des unités légales du champ appartenant à des entreprises hors-champ, et des unités légales hors-

champ appartenant à des entreprises du champ... ce qui aurait engendré des conséquences fâcheuses pour l'appareil statistique : doubles comptes difficiles à neutraliser²¹, et symétriquement unités non couvertes²².

Les choix effectués vont à l'encontre de la définition pure de l'entreprise : on peut ainsi être amené à scinder un groupe en plusieurs entreprises selon un critère non directement lié à la définition européenne de l'entreprise et en particulier

« Comme il est impossible d'apprécier, *via* un algorithme, le critère d'autonomie, il a été décidé, dans un premier temps, pour le profilage automatique, de considérer l'ensemble du groupe comme une unique entreprise. »

à l'autonomie de décision. Mais elle a le mérite de simplifier les choses et sans doute de faciliter l'adhésion des acteurs concernés au profilage automatique. Par ailleurs cette pratique est actuellement mise en œuvre pour le profilage manuel. En effet les profileurs, lorsqu'ils le peuvent, regroupent les activités financières des groupes au sein d'une entreprise à part entière.

In fine, les groupes profilés automatiquement se décomposent en :

- ❶ une entreprise contenant l'ensemble des unités légales du champ de la statistique structurelle d'entreprises ;
- ❶ et autant d'entreprises qu'il y a d'unités légales hors champ.

❶... ET MODALITÉS DE CALCUL ADAPTÉES

Comme pour le profilage manuel, la seconde phase du profilage automatique consiste à calculer les caractéristiques de ces entreprises (ou plutôt, cette fois, les estimer). Différents algorithmes ont ainsi été définis, adaptés selon les variables :

- ❶ L'identifiant de l'entreprise est obtenu à partir d'un algorithme de gestion de la continuité fondé sur les effectifs des unités légales.
- ❶ Le calcul de l'activité principale (APE) s'appuie sur le même algorithme que celui utilisé pour une unité légale : à partir de la ventilation du chiffre d'affaires en branche, mais appliqué aux activités principales et chiffre d'affaires de l'ensemble des unités. Par la suite,

21. Par exemple, les unités légales financières ou agricoles qui pourraient continuer d'être interrogées par la Banque de France ou le service statistique du ministère de l'agriculture (SSP) et qui seraient aussi prises en compte dans les résultats des groupes auxquelles ces unités légales appartiennent.

22. Si tous les services (Banque de France et SSP par exemple) ne prennent pas en compte l'entreprise dans l'élaboration de leurs statistiques, les unités légales hors de leur champ en unités légales mais dans leur champ en entreprises ne seraient plus couvertes.

si l'entreprise est interrogée dans l'ESA, son APE sera obtenue à partir de la ventilation en branches de son chiffre d'affaires consolidé.

- ❶ Les variables additives sont la somme simple des variables correspondantes dans les UL.
- ❷ Les variables non-additives (chiffre d'affaires et sa ventilation en branches, achats, dividendes, titres de participations, capitaux propres, créances et dettes, etc.) résultent d'algorithmes de consolidation. Ces variables sont concernées par de possibles échanges intragroupe non valorisés au prix du marché, qu'il convient d'éliminer pour avoir la meilleure estimation possible de l'activité économique. Le but de l'algorithme est donc d'estimer les flux intragroupe et de les retirer au total des unités légales afin d'obtenir la valeur consolidée de la variable.

Un exemple d'algorithme est présenté dans l'**encadré 2**.

❶ GROUPES DE TAILLE INTERMÉDIAIRE, SOLUTIONS INTERMÉDIAIRES

La décision de se limiter aux 70 plus grands groupes en profilage manuel a été prise en fonction des moyens disponibles au sein de l'unité dédiée au profilage. Ce seuil totalement arbitraire n'est pas le fait d'une cassure importante dans la hiérarchie des groupes

« Des réflexions sont en cours pour essayer de trouver des voies d'amélioration moins coûteuses que le profilage manuel pour traiter les groupes de taille intermédiaire. »

en France. Autrement dit, le 71^e groupe sur le sol français a un poids économique très proche du 70^e. Or, si le profilage automatique, qui se base sur des hypothèses relativement frustes, permet d'obtenir des résultats satisfaisants sur les groupes aux structures assez simples, il est moins performant face à des structures plus complexes ; la principale critique étant de systématiquement considérer tout le groupe comme ne formant qu'une

entreprise. C'est pourquoi des réflexions sont en cours pour essayer de trouver des voies d'amélioration moins coûteuses que le profilage manuel pour traiter les groupes de taille intermédiaire et notamment ceux cotés en Bourse.

Parmi ces travaux on peut citer :

- ❶ le profilage en bureau qui consiste à essayer de définir des entreprises, sans interaction directe avec les groupes, par une prise de connaissance d'informations qu'ils publient ou qui sont publiées sur eux (rapport annuel du groupe par exemple) ;
- ❷ un test pour collecter *via* une enquête statistique auprès du groupe les mêmes flux que ceux collectés par les profileurs lors du profilage manuel ;
- ❸ la prise en compte des résultats du profilage européen qui permet d'identifier des entreprises au niveau mondial et dont les traces françaises peuvent être considérées comme des entreprises nationales.

❶ QUELLES CONSÉQUENCES SUR LA VISION DE L'ÉCONOMIE ?²³

Les entreprises des groupes profilés manuellement représentent environ 12 % des effectifs ETP²⁴ et 16 % de la valeur ajoutée des entreprises françaises. Les entreprises issues des groupes profilés automatiquement représentent environ 58 % des ETP et 56 % de la valeur ajoutée des entreprises françaises. L'impact du profilage est donc significatif sur les résultats relatifs aux caractéristiques comptables et financières des entreprises. Les principaux messages liés à la prise en compte de l'entreprise au sens économique sont :

- ❶ Une économie beaucoup plus concentrée sur les grandes entreprises et les entreprises de taille intermédiaire. Là où les unités légales de plus de 250 salariés concentraient 38 % de l'emploi, 41 % de la valeur ajoutée et 61 % des exportations, les entreprises au sens économique de même taille regroupent 52 % de l'emploi, 57 % de la valeur ajoutée et 86 % des exportations. Dans ces conditions, on peut de moins en moins dire que ce sont les PME qui font l'économie en France.
- ❶ Une réallocation sectorielle. L'industrie gagne 3 points en part, passant de 23 à 26 points au détriment des services principalement.
- ❶ Une vision plus réaliste des performances des entreprises. Ainsi, on constate une augmentation de la productivité du travail (valeur ajoutée par salarié) et des salaires moyens pour les entreprises manufacturières, la construction et les transports.

Les deux derniers résultats traduisent le fait que pour une entreprise donnée, les résultats de l'ensemble des unités légales qui la composent vont contribuer au secteur de l'entreprise là où dans la vision en unités légales elles contribuaient pour leur propre secteur. Dans l'exemple de la **figure 2**, dans la vision en entreprise, les résultats de l'UL11 commerciale vont contribuer au secteur de l'entreprise qui est industrielle alors que dans la vision en unités légales, ses résultats seraient comptabilisés dans son secteur commercial.

- ❶ Une réduction du chiffre d'affaires total de l'ordre de 7 %, celui des dividendes et capital social de l'ordre de 60 %. L'effet de la consolidation est donc très fort pour les variables du bilan.

❷ LE PROFILAGE, UNE LONGUE HISTOIRE DONT LA FIN RESTE À ÉCRIRE

L'histoire du profilage à l'Insee remonte à plus de 25 ans. Les premières réflexions datent du début des années quatre-vingt-dix, avec notamment les travaux d'Emmanuel Raulin (Raulin,

1996), sachant qu'en parallèle le règlement européen de 1993 donnait une définition de l'entreprise toujours en vigueur aujourd'hui²⁵. Une analyse poussée sur les unités statistiques et sur le profilage s'est ensuivie, notamment à travers l'étude des cas néerlandais, canadien, et australien, déjà opérationnels²⁶. Mais l'approche française demeurerait alors pour l'essentiel théorique. Les premières

« Les premières mises en œuvre concrètes datent du début des années 2000, avec le profilage de groupes automobiles. »

23. Voir (Deroyon, 2015), (Haag, 2016), (Bacheré, 2017) et (Balcone, Schweitzer, 2018).

24. Équivalent temps plein (ETP).

25. Voir aussi (Willeboordse, 1997), le travail qui avait été mené par le CBS à la demande d'Eurostat.

26. Voir (Brudieu, 1999) et (Rivière, 1999).

prises en œuvre concrètes datent du début des années 2000, avec le profilage de groupes automobiles dont les résultats ont été pris en compte dans les Enquêtes Annuelles d'Entreprises. Par la suite, les résultats d'autres groupes profilés manuellement ont été pris en compte dans les résultats d'ESANE.

L'année 2019 marque une étape importante avec la prise en compte de l'ensemble des groupes dans la statistique structurelle d'entreprise. En outre, des pistes d'amélioration sont envisagées : en plus des réflexions menées sur le profilage semi-automatique, d'autres sont en cours afin :

- d'améliorer les algorithmes de consolidation en mobilisant d'autres informations (nouvelles variables collectées dans l'ESA sur les flux intragroupe, prise en compte des comptes consolidés livrés par la DGFIP sur certains groupes) ;
- d'être plus efficace sur le profilage en bureau en ayant recours à des méthodes de *web scrapping* et d'analyse textuelle pour essayer d'extraire de façon plus automatisée les informations utiles des documents de référence, tableaux des filiales et participations des liasses sociales.

Par ailleurs, si la France reste moteur sur ce sujet à l'instar des Pays-Bas en Europe ou du Canada ou de l'Australie au niveau mondial, le profilage national commence à se développer dans les autres pays à la suite de recommandations faites par Eurostat.

En revanche, le profilage européen (Xirouchakis, Hecquet, 2018) reste pour le moment plus confidentiel et se limite au stade de l'expérimentation. Plus de 300 groupes ont certes été profilés dans ce cadre, mais il ne s'agit pas des plus importants. Par ailleurs, aucune statistique prenant en compte les entreprises obtenues par ces travaux n'a été diffusée. Toutefois, il y a fort à parier qu'avec la mondialisation croissante et l'interdépendance accrue des économies, cette démarche a de beaux jours devant elle.

BIBLIOGRAPHIE

BACHERÉ, Hervé, 2017. Une forte proportion des emplois créés entre 2009 et 2015 sont portés par les entreprises de taille intermédiaire. In : *Les entreprises en France, édition 2017, Insee Références* [en ligne]. 7 novembre 2017. pp. 25-34. [Consulté le 4 juin 2019]. Disponible à l'adresse :

https://www.insee.fr/fr/statistiques/fichier/3152792/ENTFRA17_c_D1-ETI.pdf

BALCONE, Thomas et SCHWEITZER, Camille, 2018. Profiling : a new way to increase the quality of statistics on research and development. In : *European conference on quality in official statistics*, Cracovie, 7 mars 2018 [en ligne]. [Consulté le 4 juin 2019]. Disponible à l'adresse :

https://www.q2018.pl/papers-presentations/?drawer=Sessions*Session%2019*Thomas%20Balcone

BÉGUIN, Jean-Marc et HECQUET, Vincent, 2015. Avec la définition économique des entreprises, une meilleure vision du tissu productif. In : *Les entreprises en France, édition 2015, Insee Références* [en ligne]. [Consulté le 4 juin 2019]. Disponible à l'adresse :

https://www.insee.fr/fr/statistiques/fichier/1906509/ENTFRA15c_D1_Entreprises.pdf

BÉGUIN, Jean-Marc, HECQUET, Vincent et LEMASSON, Julien, 2012. Un tissu productif plus concentré qu'il ne semblait – Nouvelle définition et nouvelles catégories d'entreprises. In : *Insee Première* [en ligne]. 27 mars 2012. N°1399. [Consulté le 4 juin 2019]. Disponible à l'adresse :

<https://www.insee.fr/fr/statistiques/fichier/version-html/1281373/ip1399.pdf>

BRUDIEU, Germaine, 1999. Synthèse sur les pratiques néerlandaise, canadienne et française. In : *Des unités statistiques pour représenter l'économie – Approche française et mise en perspective internationale, Insee-Méthodes* [en ligne]. Octobre 1999. N°90, pp. 89-95. [Consulté le 4 juin 2019]. Disponible à l'adresse :

<https://www.epsilon.insee.fr/jspui/bitstream/1/17930/1/imethode90.pdf>

CHANTELOUP, Guillaume, 2017. Calculation of the automatically profiled enterprises characteristics. In : *Meeting of the Group of Experts on Business Registers, Unece*, Paris, 27-29 septembre 2017 [en ligne]. [Consulté le 4 juin 2019]. Disponible à l'adresse :

https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.42/2017/Session_2_France_ppt.pdf

CHANTELOUP, Guillaume, 2018. Consolider les réponses des unités légales pour une statistique d'entreprise plus cohérente. In : *Journées de Méthodologie Statistique 2018*, Paris, 12-14 juin 2018 [en ligne]. [Consulté le 4 juin 2019]. Disponible à l'adresse :

http://www.jms-insee.fr/2018/S24_2_ACTE_CHANTELOUP_JMS2018.pdf

DEPOUTOT, Raoul, 2008. Rapport du groupe de travail sur les statistiques structurelles fondées sur les groupes d'entreprises et leurs sous-groupes. In : *site du CNIS* [en ligne]. Janvier 2008. Rapport N° 107. [Consulté le 4 juin 2019]. Disponible à l'adresse :

https://www.cnis.fr/wp-content/uploads/2017/10/RAP_2008_107_statistiques_structurelles.pdf

DEROYON, Julien, 2015. De nouvelles données issues du profilage des groupes : une part accrue de l'industrie, des entreprises plus performantes, mais une capacité à financer l'investissement et un endettement plus dégradés. In : *Les entreprises en France, édition 2015, Insee Références* [en ligne]. 28 octobre 2015. pp. 39-51. [en ligne] [Consulté le 4 juin 2019]. Disponible à l'adresse :

https://www.insee.fr/fr/statistiques/fichier/1906511/ENTFRA15d_D2_profilage.pdf

HAAG, Olivier, 2016. Profiling : a new and better way to apprehend the globalization. In : *European Conference on Quality in Official Statistics*, Madrid, 3 juin 2016. [en ligne]. [Consulté le 4 juin 2019]. Disponible à l'adresse : <http://www.ine.es/q2016/docs/q2016Final00033.pdf>

HAAG, Olivier, 2018. How to improve the quality of the statistics by combining different statistical units. In : *The Unit Problem and Other Current Topics in Business Survey Methodology* – Edited Proceedings of the European Establishment Statistics – Workshop 2017. Cambridge scholar Publishing – Registration Number : 04333775, pp. 31-46, Editor(s) : Boris Lorenc, Paul A. Smith, Mojca Bavdaž, Gustav Haraldsen, Desislava Nedyalkova, Li-Chun Zhang, Thomas Zimmermann. ISBN (10) 1-5275-1661-X, ISBN (13) 978-1-5275-1661-8.

HOOGSTEN, Tammy, 2016. Globalization and Profiling at Statistics Canada. In : *25th Meeting of the Wiesbaden Group on Business Registers, International Round table on Business Survey Frames*, Tokyo, 8-11 novembre 2016 [en ligne]. [Consulté le 4 juin 2019]. Disponible à l'adresse : https://www.stat.go.jp/english/info/meetings/wiesbaden2016/pdf/7-03_can_pap.pdf

MARIOTTE, Henri, 2017. LIFI : le répertoire français sur les groupes. In : *Meeting of the Group of Experts on Business Registers, Unece*, Paris, 27-29 septembre 2017 [en ligne]. [Consulté le 4 juin 2019]. Disponible à l'adresse : https://www.unec.org/fileadmin/DAM/stats/documents/ece/ces/ge.42/2017/France_2_FRE.pdf

PIETSCH, Leon, 1995. Profiling Large Business to define Frame Units. In : COX, Brenda G. et al.. *Business Survey Methods*. New York, Wiley, 16 février 1995, pp. 101-114. ISBN : 9780471598527.

RAULIN, Emmanuel, 1995. Observer autrement pour mieux représenter. In : *Observer et représenter un monde de plus en plus complexe – Un défi pour la statistique d'entreprise, Insee-Méthodes* [en ligne]. 12 octobre 1995. N°54, pp. 139-146. [Consulté le 4 juin 2019]. Disponible à l'adresse : <https://www.epsilon.insee.fr/jspui/bitstream/1/17910/1/imethode54.pdf>

RIVIÈRE, Pascal, 1999. Profilage et unités statistiques en Australie. In : *Des unités statistiques pour représenter l'économie – Approche française et mise en perspective internationale, Insee-Méthodes* [en ligne]. Octobre 1999. N°90, pp. 109-117. [Consulté le 4 juin 2019]. Disponible à l'adresse : <https://www.epsilon.insee.fr/jspui/bitstream/1/17930/1/imethode90.pdf>

SMITH, Paul, 2013. Sampling and Estimation for Business Surveys. In : SNIJKERS, Ger et al.. *Designing and conducting business surveys*. Wiley, septembre 2013. ISBN-13 : 978-0470903049. ISBN-10 : 047090304X.

WILLEBOORDSE, Ad, 1997. *Handbook on design and implementation of business surveys* [en ligne]. Eurostat, octobre 1997. [Consulté le 4 juin 2019]. Disponible à l'adresse : <http://ec.europa.eu/eurostat/ramon/statmanuals/files/Handbook%20on%20surveys.pdf>

XIROUCHAKIS, Ioannis et HECQUET, Vincent 2018. Improving the quality of Business Statistics through Profiling. In : *European conference on quality in official statistics*, Cracovie, 7 mars 2018 [en ligne]. [Consulté le 4 juin 2019]. Disponible à l'adresse : https://www.q2018.pl/wp-content/uploads/Sessions/Session%2019/Ioannis%20Xirouchakis/Session%2019_Ioannis%20Xirouchakis_paper.pdf



Présentation du numéro N2

Le dossier du numéro N2 du *Courrier des statistiques* s'inscrit dans une tradition de coopération de trente ans entre l'Insee et Destatis, l'office fédéral de la statistique allemand. Suivent deux articles illustrant l'impulsion donnée à la statistique publique par les institutions internationales : ainsi, le 16^e objectif du développement durable des Nations Unies a contribué à l'émergence rapide en France de mesures de la sécurité et de la satisfaction des citoyens vis-à-vis des institutions ; le dispositif européen de statistiques conjoncturelles d'entreprises, créé avec l'Euro, évolue encore avec le nouveau règlement FRIBS : la statistique conjoncturelle française s'y adapte et anticipe.

Avec son nouveau référentiel de métadonnées statistiques, RMÉS, la statistique publique franchit un cap : fondées sur des standards internationaux, les métadonnées des opérations statistiques seront activées tout au long du cycle de vie et gérées dans un système d'information commun, assorti de services mutualisés. C'est un autre référentiel que le projet Élire a mis en place en trois ans : le Répertoire électoral unique, adossé à un dispositif d'échanges de données de qualité et largement modernisé.

Enfin, le dernier article présente le « profilage » des entreprises réalisé par l'Insee pour mieux identifier les acteurs économiques pertinents. Ici encore, la statistique publique s'adapte aux évolutions de son environnement et construit des processus où automatisation des traitements et expertise se complètent utilement.

ISSN 2107-0903
ISBN 978-2-11-151235-1

