

Courrier des statistiques

14



Rédaction en chef

Emmanuelle Nauze-Fichet

Contribution

Insee : Camille Beaurepaire,
Thierry Boucher, Pierrette Briant,
Élodie Kranklader, Pierre Lamarche,
Hervé Piffeteau, Pascal Rivière
DGAFP : Gaël de Peretti
Eurostat : Pierre Bischoff
SSP : Pauline Cazaban,
François Chevalier, Vincent Marcus

Directeur de la publication

Fabrice Lenglart

Directeur de la collection

Pascal Rivière

Rédaction

Solenn Ily, Emmanuelle Nauze-Fichet,
Yann Quélenne, Pascal Rivière,
avec la collaboration de Marine Le Roux

Composition

Agence Efil
90 boulevard Heurteloup
37 000 Tours
02 47 47 03 20
www.efil.fr

Photo de couverture

Getty Images

Éditeur

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

www.insee.fr

© Insee 2025 « Reproduction partielle
autorisée sous réserve de la mention
de la source et de l'auteur ».

Courrier des statistiques N14

SOMMAIRE

Présentation du numéro <i>Emmanuelle Nauze-Fichet</i>	4
La construction du système statistique européen – 1952-2003 : un grand élan <i>Pierre Bischoff et Hervé Piffeteau</i>	7
Des données de papier - Une histoire des échanges de publications statistiques <i>Camille Beaurepaire et Thierry Boucher</i>	27
Fonction publique : des chiffres et des lettres au cœur du débat public <i>Gaël de Peretti</i>	51
Prophyl : de la parcelle à la statistique, un projet pour simplifier la récolte de données agricoles <i>Pauline Cazaban, François Chevalier et Vincent Marcus</i>	77
À propos du temps de production des statistiques publiques <i>Pierre Lamarche et Pascal Rivière</i>	95
Des démarches qualité pour sécuriser et optimiser la production de statistiques publiques <i>Pierrette Briant et Élodie Kranklader</i>	119

PRÉSENTATION DU NUMÉRO

En décembre 2018, dans l'éditorial du N1, Jean-Luc Tavernier écrivait :

« Le Courrier des statistiques est [...] notre bien commun. Sa pérennité repose sur la bonne volonté de tous. »

Cette bonne volonté ne faiblit pas : avec ce numéro N14, le Courrier des statistiques « nouvelle formule » franchit le cap des cent articles publiés !

Dans cette nouvelle édition, nous vous proposons trois papiers éclairant l'histoire de la statistique publique et trois articles plus contemporains. Commençons donc par plonger dans le passé...

Nous sommes en 1951. Les six pays fondateurs de l'Union européenne viennent de signer le traité de Rome, instituant la Communauté européenne du charbon et de l'acier. Pour piloter sa mise en œuvre, des statistiques « harmonisées », comparables entre les six pays, sont indispensables. Comment les mettre en place alors que chaque pays a sa propre organisation statistique ? Ce qui peut paraître évident aujourd'hui est le fruit d'une longue histoire, marquée par des ralentissements et des accélérations au gré des crises et avancées de la construction européenne. Pierre Bischoff et Hervé Piffeteau nous en racontent les cinquante premières années. De 1952 à 2003, une succession d'élans ont permis de poser les fondements du système statistique européen (SSE), avec notamment : la création de l'Office statistique des Communautés européennes (qui adoptera en 1973 son surnom Eurostat), la mise en place d'un programme statistique pluriannuel, la création de comités structurant les échanges entre les partenaires nationaux et communautaires, et les fondements juridiques du futur SSE. Ce « grand élan », comme l'appellent les auteurs, sera coupé net par la révélation en 2004 de la falsification des comptes nationaux grecs... Une histoire que les auteurs vous réservent pour un prochain numéro...

Remontons encore plus loin dans le passé. Nous sommes cette fois en 1853. Le premier Congrès international de statistique se tient à Bruxelles. Des statisticiens publics de différents pays se réunissent à la fois pour échanger sur leurs méthodes scientifiques et pour réaliser (déjà) des travaux de comparaison internationale. À une époque où Internet n'existe pas, les publications statistiques « papier » sont un bien précieux pour ces discussions et travaux, et les bibliothèques statistiques sont au cœur de l'univers de la donnée. Camille Beaurepaire et Thierry Boucher nous racontent l'histoire des échanges de publications statistiques entre les pays au XIX^e siècle et pendant la première moitié du XX^e siècle. Les questions sont alors celle de l'organisation matérielle des échanges de publications (par voie postale !), mais aussi celle de leur classement. Les statisticiens publics tentent de résoudre ce problème, central pour les bibliothèques, pour le cas particulier du « savoir statistique ». Grâce en particulier au fond ancien étranger de la bibliothèque Alain Desrosières de l'Insee (environ 75 000 documents, soit près d'un kilomètre linéaire), les auteurs nous font revivre cette période passionnante... tout en nous initiant à l'histoire de la bibliothéconomie.

Nous voici à présent au début des années 1970, pour voir naître le Service statistique ministériel de la fonction publique (SSM FP). Gaël de Peretti, son responsable actuel, a fouillé les archives (et notamment celles du « premier » Courrier des statistiques)

pour nous raconter l'histoire de ce service. Depuis son origine, celui-ci fait partie de la Direction générale de l'administration et de la fonction publique (DGAFP), un ancrage qui lui permet d'être au cœur de l'organisation où se conçoivent les politiques des ressources humaines de la fonction publique. Ce positionnement lui permet aussi de discuter de son programme dans des instances de concertation spécifiques. Les grands travaux du SSM FP démarrent avec la production de statistiques sur les effectifs et les rémunérations des agents de la fonction publique de l'État, à l'appui de remontées des fichiers de paie. Ils s'étoffent progressivement, pour aborder d'autres thématiques et couvrir également les fonctions publiques territoriale et hospitalière. L'évolution du rapport annuel sur l'état de la fonction publique, publication phare du SSM FP, témoigne de tous les enrichissements réalisés. Ce rapport, originellement « politique », adopte aujourd'hui tous les critères d'une publication « statistique ».

Nous voici maintenant « au présent » avec le récit d'une expérimentation menée en 2024 par le service de la statistique et de la prospective (SSP) du ministère de l'Agriculture, de l'Agro-alimentaire et de la Souveraineté alimentaire. De plus en plus d'agriculteurs s'équipent de « logiciels de gestion parcellaire » pour piloter leur exploitation et simplifier leurs démarches administratives, y consignant notamment les traitements phytosanitaires (pesticides) appliqués sur leurs cultures. Afin de moderniser la collecte de ces informations et réduire la charge d'interrogation statistique, le SSP a imaginé que les agriculteurs puissent transmettre directement les données issues de ces logiciels. C'est de cette idée qu'est né le projet « Prophyl ». Pauline Cazaban, François Chevalier et Vincent Marcus en racontent la genèse et les suites envisagées. La première étape a été la mise en place d'un cadre de dialogue et de coopération entre acteurs publics et privés. Il a fallu penser tous les aspects organisationnels, juridiques, informatiques, avec une attention particulière aux questions de confidentialité, puis mettre en place des actions de communication auprès des agriculteurs. Au total, des informations ont été recueillies sur près de 5 000 traitements phytosanitaires. Le projet Prophyl constitue un premier pas vers une collecte de données phytosanitaires plus moderne, sécurisée et simplifiée. Il préfigure l'enquête Phyto 2026, qui intégrera ce dispositif à grande échelle.

Dans l'article suivant, Pierre Lamarche et Pascal Rivière s'intéressent au temps de production des statistiques publiques, pour répondre notamment à ceux qui le trouveraient trop long. Ils proposent une grille qui permet d'analyser en parallèle les processus fondés sur des enquêtes (auprès de ménages ou d'entreprises), sur des sources administratives (par exemple les fichiers fiscaux) ou sur des sources privées (par exemple les données collectées par les enseignes de la grande distribution sur les achats des ménages). Tous ces processus peuvent se décomposer en trois phases : préparation, construction des données, traitement statistique et diffusion. Pour chaque phase, les auteurs analysent les facteurs principaux qui jouent sur le temps. C'est sur la dernière phase qu'il semble y avoir le plus de marge pour gagner du temps en automatisant. Mais même là, une intervention humaine est souvent nécessaire, celle du statisticien, pour vérifier par exemple la cohérence interne ou externe des données, comprendre l'origine des éventuelles incohérences, décider des traitements correctifs à appliquer. À travers l'identification des facteurs influents, les auteurs mettent en avant où peuvent se faire d'éventuels « compromis » pour gagner du temps, quand cet enjeu est crucial, ce qui implique alors de réduire l'ambition sur d'autres aspects.

Pour clôturer ce numéro, Pierrette Briant et Élodie Kranklader nous expliquent le principe et les enjeux des « démarches qualité » menées par l'Insee pour sécuriser et optimiser la production de statistiques publiques. Par exemple, le processus d'élaboration de l'indice des prix à la consommation a fait l'objet d'une telle démarche. S'inscrivant dans les standards internationaux, au premier rang desquels le code de bonnes pratiques de la statistique européenne, l'approche est celle des petits pas, de l'amélioration continue. Une première démarche permet ainsi de cerner des axes d'amélioration et de décider d'actions ciblées à mettre en œuvre, une démarche ultérieure (une « revue ») permet de faire le point et de définir d'éventuelles nouvelles actions, etc. La démarche, en elle-même, se déroule en trois temps : description globale du processus, analyse des risques à surveiller, plan d'actions à mettre en œuvre. Elle est pilotée par des agents de l'unité Qualité de l'Insee, appelés « experts qualité ». La démarche est néanmoins collective, menée en interaction étroite avec les équipes en charge des processus. C'est d'ailleurs la force de ces démarches : les experts qualité aident les équipes à échanger sur le fonctionnement de leur processus et les difficultés de chacun pour établir ensemble un plan d'actions consensuel.

Bonne lecture !

Emmanuelle Nauze-Fichet
Rédactrice en chef, Insee

La construction du système statistique européen

1952-2003 : un grand élan



Pierre Bischoff* et Hervé Piffeteau**

Ce premier volet de l'*histoire du système statistique européen (SSE)* couvre les années 1952 à 2003. C'est au cours de cette période qu'ont été posés les grands piliers du système actuel.

Les statistiques officielles harmonisées ont d'abord été développées sur la base d'une coopération entre les six pays fondateurs des premières Communautés européennes : la Communauté européenne du charbon et de l'acier, la Communauté économique européenne et la Communauté européenne de l'énergie atomique. En 1959, un service statistique dédié à ces Communautés est créé : l'Office statistique des Communautés européennes (OSCE), plus connu aujourd'hui sous le nom d'Eurostat. Ce dernier s'attelle à faire progresser l'harmonisation statistique.

Cependant, les avancées deviennent plus difficiles, avec l'élargissement à d'autres États et à d'autres domaines politiques. Pour relever les défis, la statistique européenne s'organise plus formellement : programmation des travaux, création des grands comités. Les valeurs partagées par les statisticiens facilitent les avancées, mais les demandes vont toujours croissant. La construction juridique de l'Union européenne se propage alors à la sphère statistique : en 1997, le Conseil adopte le règlement n° 322/97 relatif à la statistique communautaire. Cette « loi statistique de 1997 » préfigure le règlement n° 223/2009 relatif aux statistiques européennes qui régit aujourd'hui le SSE.

 This first chapter of the history of the European Statistical System (ESS) covers the years 1952 to 2003. This period saw the construction of the current system's foundation.

Harmonised official statistics were first developed on the basis of cooperation between the six founding members of the first European Communities: the European Coal and Steel Community, the European Economic Community, and the European Atomic Energy Community. In 1959, a statistics department dedicated to these communities was born: the European Statistical Office, more commonly known as Eurostat nowadays. Its goal was to further improve statistical harmonisation.

However, progress became difficult with the enlargement of the European Union to other states and other political fields. To take on those challenges, European statistics started organising in a more formal manner: planification of work, creation of the big comities. The values statisticians have in common made progress easier, but demand kept on growing. The European Union legal construction then spread to the statistical sphere: in 1997, the Council passed Regulation n° 322/97 regarding Community statistics. This "statistics law of 1997" announces Regulation n° 223/2009 regarding European statistics. This regulation still rules the ESS to this day.

* Conseiller auprès de la directrice générale d'Eurostat.
pierre.bischoff@ec.europa.eu

Note : Les opinions exprimées dans le présent article sont celles de l'auteur et ne reflètent pas nécessairement la position de la Commission européenne.

** Membre de l'Inspection générale de l'Insee.
herve.piffeteau@insee.fr

Le pilotage de l'Union européenne (UE) nécessite des statistiques fiables et harmonisées entre les pays. Les besoins d'information sont liés à la mise en œuvre et au suivi des politiques européennes, lesquelles concernent des domaines variés : agriculture, emploi et affaires sociales, marché unique, etc.¹ Pour répondre à ces besoins, les statistiques officielles élaborées dans le cadre du système statistique européen (SSE)² jouent un rôle prépondérant. Il importe en effet que tous les décideurs puissent disposer de données impartiales et objectives. Ces statistiques officielles contribuent aux capacités d'information requises par le fonctionnement de l'UE, pour appuyer ses objectifs stratégiques, ainsi que les politiques et instruments de soutien correspondants. Elles constituent dès lors un bien public qui permet à la démocratie de fonctionner harmonieusement.

Un système efficace de production de statistiques européennes nécessite un cadre juridique à la fois stable et transparent, en mesure de garantir l'indépendance, l'intégrité et le sens des responsabilités des autorités qui en sont chargées. Dans le même temps, il doit être assez souple pour pouvoir relever les défis actuels et à venir. Il doit notamment pouvoir faire face à l'évolution rapide des besoins des politiques, aux contraintes financières, à l'objectif de réduction de la charge pesant sur les répondants et à l'évolution des technologies de l'information.

“
Les avancées et crises politiques ayant façonné l'Europe ont évidemment marqué la manière dont les statistiques officielles ont été construites.
”

Le SSE actuel et ses fondements juridiques, mais aussi les règlements européens sous-jacents à la production des statistiques officielles européennes, ont été élaborés au gré de la construction de l'Europe. Ce premier volet de l'histoire du SSE couvre les années 1952 à 2003, une période marquée par une succession d'élans qui ont permis de poser les piliers du système³. Les avancées et crises politiques ayant façonné l'Europe ont évidemment marqué la manière dont les statistiques officielles ont été construites (*figure 1*). La première étape a été la

création d'un office statistique européen, dont la production de statistiques a d'abord reposé essentiellement sur la coopération entre les six États membres fondateurs de l'Europe. L'élargissement progressif à d'autres pays, à partir de 1973, a cependant rendu de plus en plus manifeste le besoin d'une procédure décisionnelle impliquant des comités pour mieux réguler les échanges entre, d'une part, l'office statistique européen (désormais appelé Eurostat) et, d'autre part, les instituts nationaux de statistique (INS) et les autres autorités statistiques nationales⁴. Ce n'est cependant qu'au milieu des années 1980 qu'émergera l'idée d'un « système statistique européen », doté de structures et d'une organisation de concertation commune. Enfin, comme pour les autres champs de la construction européenne, les statistiques officielles ont été développées sur la base de dispositions juridiques, indispensables à la consolidation d'une Europe dont le suivi nécessitait des indicateurs robustes et fiables. Ce sont tous ces élans qui ont façonné entre 1952 et 2003 le SSE, avec notamment la création du comité du programme

¹ Sur les domaines d'action de l'UE, voir le lien : https://european-union.europa.eu/priorities-and-actions/actions-topic_fr.

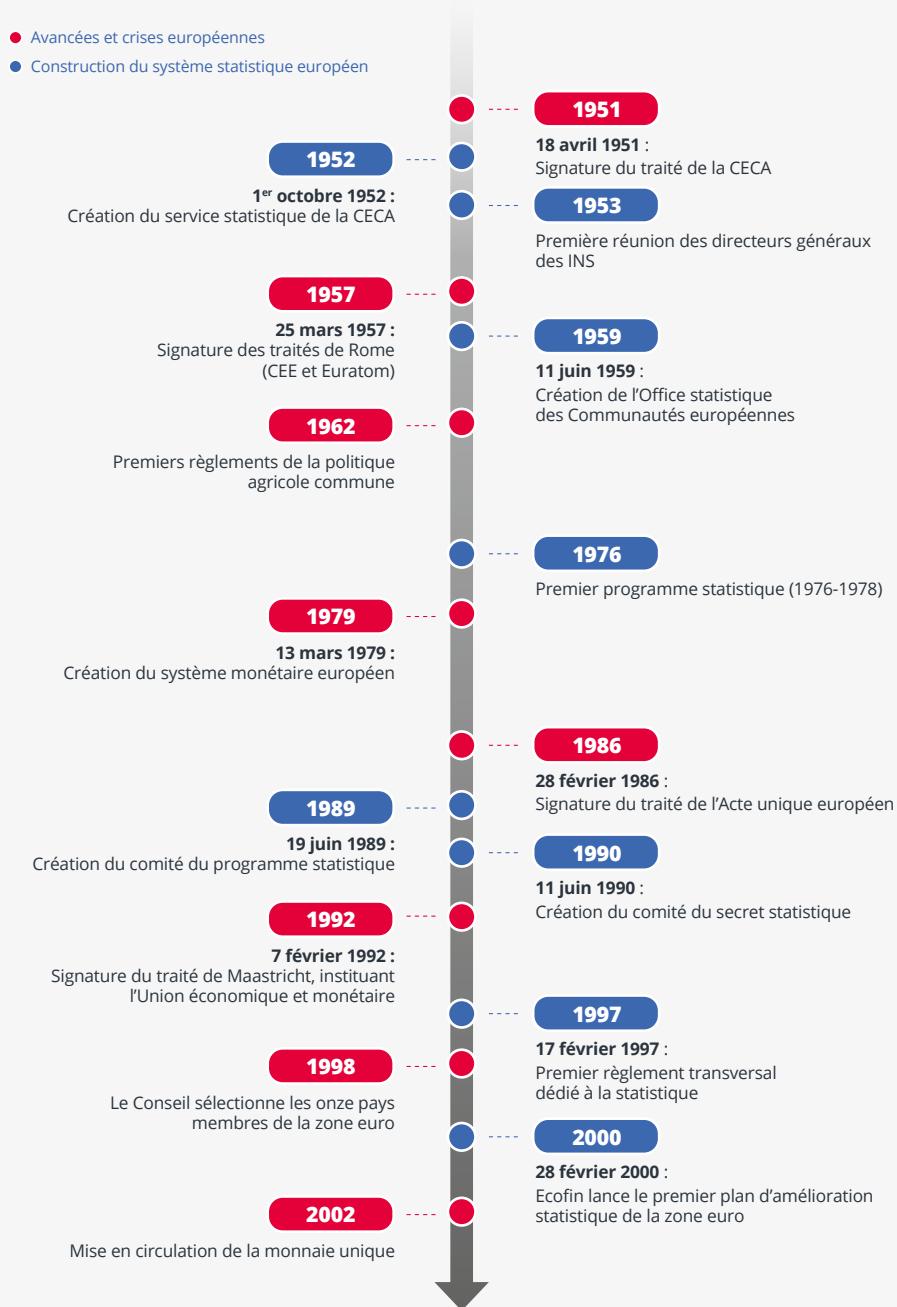
² Voir le lien : <https://ec.europa.eu/eurostat/fr/web/european-statistical-system/overview>.

³ Pour de nombreux détails historiques, cet article s'appuie sur l'histoire d'Eurostat retracée par deux de ses anciens directeurs, Alberto De Michelis et Alain Chantraine (De Michelis et Chantraine, 2003).

⁴ Les autres autorités statistiques nationales recouvrent par exemple, en France, des services statistiques ministériels.

► Figure 1 - L'élaboration du système statistique européen au gré de la construction de l'Europe

- Avancées et crises européennes
- Construction du système statistique européen



CECA : Communauté européenne du charbon et de l'acier.

CEE : Communauté économique européenne.

Ecofin : Conseil pour les affaires économiques et financières, rassemblant les ministres des Finances des États membres.

Euratom : Communauté européenne de l'énergie atomique.

INS : Institut national de statistique.

statistique, en 1989, et l'adoption de la première loi statistique européenne, en 1997. C'est sur la base, notamment, de ces deux piliers fondateurs que naîtra, à la fin des années 2000, le règlement n° 223/2009 relatif aux statistiques européennes⁵ qui régit aujourd'hui le fonctionnement du SSE.

► 1952-1985 : la statistique européenne émerge au gré de la construction de l'Europe

Le 18 avril 1951 marque le départ officiel de la construction de l'Union européenne. Six pays signent à Paris le traité instituant la Communauté européenne du charbon et de l'acier (CECA)⁶ : l'Allemagne de l'Ouest, la France, l'Italie, la Belgique, les Pays-Bas et le Luxembourg. Ce traité instaure un marché commun du charbon et de l'acier entre les six pays signataires : il supprime les droits de douane et les restrictions quantitatives entravant la libre circulation de ces marchandises, ainsi que toutes les aides ou subventions qui seraient accordées par les pays à leur production nationale. La gestion de la Communauté est placée sous le contrôle de la Haute Autorité, présidée initialement par Jean Monnet. Le traité entre en vigueur le 23 juillet 1952 pour une durée de cinquante ans (il expirera le 23 juillet 2002).

Les traités de Paris et de Rome, fondateurs de la construction européenne, ébauchent un service de la statistique européenne...

Dans son article 46, le traité de Paris fixe à la Haute Autorité des objectifs de suivi de l'activité dans les domaines de l'acier et du charbon qui requièrent la mise en place d'une documentation statistique⁷. L'article 47 prévoit par ailleurs que la Haute Autorité peut recueillir les informations nécessaires à l'accomplissement de sa mission. En 1952, afin d'assister les divisions de l'économie et de l'industrie qui viennent d'être créées, il est décidé de mettre en place à titre temporaire une division auxiliaire de la statistique, comprenant initialement sept personnes. Elle sera finalement pérennisée et prendra par la suite le nom de division des statistiques. Cette division va définir les thèmes qui seront le fil conducteur du programme statistique européen tout au long des années à venir : **l'harmonisation des statistiques visant à leur comparabilité, les relations avec les INS de la Communauté et les relations internationales.**

La division des statistiques commence à travailler en rassemblant des données issues de sources internationales existantes. Puis, en partenariat avec les INS de la Communauté et les organisations professionnelles, elle élabore des méthodologies communes afin de rapprocher des sources nationales existantes ou de créer de nouvelles sources. L'objectif est d'abord d'aboutir à des statistiques comparables dans les domaines du charbon et de l'acier. C'est le début de l'harmonisation européenne des statistiques publiques.

⁵ Voir les références juridiques en fin d'article. Ce règlement a depuis fait l'objet de révisions, relatives notamment à l'accès aux sources de données privées par le système statistique européen (Lesur, 2025).

⁶ Voir les références juridiques en fin d'article.

⁷ Par exemple, la Haute Autorité doit effectuer une étude permanente de l'évolution des marchés et des tendances des prix.

Au début, la coopération avec les INS de la Communauté se limite à la participation de leurs représentants dans des comités sectoriels, dominés essentiellement par les fonctionnaires de la Haute Autorité et des membres des ministères et organisations professionnelles concernés. En 1953, la division des statistiques réunit pour la première fois le « groupe de travail des directeurs généraux des INS ». Ce dernier fonctionnera ensuite régulièrement et sera renommé ultérieurement « conférence des directeurs généraux des INS » ou, dans le jargon des statisticiens, « conférence des DGINS ».

Le 25 mars 1957, l'intégration européenne se poursuit : les six pays pionniers de la construction européenne signent à Rome les traités instituant, d'une part, la Communauté économique européenne (CEE) et, d'autre part, la Communauté européenne de l'énergie atomique (Euratom)⁸. La CEE doit permettre la création d'une union douanière pour l'ensemble des biens et services (elle entrera en vigueur en juillet 1968), tandis que l'Euratom doit promouvoir la coopération dans l'énergie nucléaire.

Avec le traité instituant la CEE naît la législation européenne en matière de statistiques, qui permettra de bâtir ce que l'on appelle aujourd'hui l'acquis statistique communautaire⁹. En effet, auparavant, dans le cadre du traité instituant la CECA, la Haute Autorité pouvait collecter directement certaines statistiques auprès des entreprises sidérurgiques et charbonnières à la seule condition de l'accord des comités sectoriels compétents du

CECA. Désormais, dans le cadre du traité de Rome, les conditions sont plus restrictives : la collecte directe d'informations statistiques dans les États membres doit faire l'objet de règlements adoptés par le Conseil de la CEE.

*En 1958,
plusieurs services
communs aux trois
Communautés sont
créés, dont un service
des statistiques.*

En 1958, la décision est prise de créer plusieurs services communs aux trois Communautés (CECA, CEE et Euratom) : un service juridique, un service de presse et information et... un service des statistiques. Ce dernier est d'abord placé sous l'autorité de la division statistique

de la Haute Autorité, organe exécutif de la CECA. En 1959, il prend le nom d'**Office statistique des Communautés européennes (OSCE)**, plus connu aujourd'hui sous le nom d'Eurostat. Il est alors placé sous l'autorité d'un conseil d'administration composé d'un membre de l'exécutif de chacune des trois Communautés... jusqu'à la fusion en 1967 des trois exécutifs en une seule institution : la Commission des Communautés européennes¹⁰.

... pour mettre en place des statistiques harmonisées dans des domaines de plus en plus variés

Dans les années 1960, les priorités de l'OSCE sont de rassembler dans les différents pays les données disponibles, de rendre comparables les principes, les définitions et les méthodes, de combler les lacunes en matière de production statistique et, enfin,

⁸ Voir les références juridiques en fin d'article.

⁹ De manière générale, l'acquis communautaire est l'ensemble des droits et obligations communs qui constituent le corpus législatif de l'UE et sont incorporés dans les systèmes juridiques de ses États membres. Concernant plus spécifiquement les statistiques, l'acquis au 11 avril 2025 est rassemblé dans Eurostat (2025).

¹⁰ Elle prendra finalement le nom de Commission européenne en 2007.

d'améliorer la qualité des données statistiques (De Michelis et Chantraine, 2003). En 1967, le premier directeur général d'Eurostat, Rolf Wagenführ¹¹, évoque dans un ouvrage l'idée d'une loi statistique spéciale autorisant la Commission des Communautés européennes à collecter directement des informations individuelles (Wagenführ, 1967). Il considère en effet que cela serait le seul moyen de coordonner et d'harmoniser efficacement les données transmises par chacun des États membres. La première loi communautaire pour organiser la production statistique dans son ensemble ne verra cependant le jour que trente ans plus tard (voir infra).

En France, jusqu'au début des années 1960, l'Insee n'est pas très « attiré » par l'activité statistique internationale, même si les premières sollicitations européennes émergent (Insee, 1996). L'élan communautaire concerne surtout les comptes nationaux. L'OSCE crée dans ce domaine un groupe de travail auquel participe l'expert français André Vanoli (Muller, 2024). Ce dernier est chargé de faire des propositions pour établir un cadre communautaire de comptabilité nationale. De manière générale, concernant les statistiques européennes, **l'orientation privilégiée par l'OSCE et les six INS est de parvenir, à différentes échéances, à une convergence des méthodes de travail**, plutôt qu'à des structures et à une organisation de concertation commune.

Dans la seconde moitié des années 1960, l'OSCE est confronté à de nouvelles demandes statistiques générées par le développement des politiques européennes : la politique agricole commune (PAC), la coopération renforcée en matière d'économie et la préparation du premier élargissement. Ce dernier n'interviendra qu'en 1973, en raison d'une succession de crises qui retardent le processus d'intégration européenne (*encadré 1*). Dans ce contexte, l'approche envisagée par l'OSCE pour développer la statistique européenne demeure celle d'une harmonisation par les méthodes. Elle se heurte cependant aux positions alors défendues par les directeurs généraux des INS, confrontés également, à l'époque, à de fortes contraintes nationales. Le dialogue est difficile : si tous les pays sont d'accord sur le principe de renforcer l'harmonisation, chacun souhaite qu'elle se fasse sur la base de ses propres méthodes. Le rôle de conciliateur opéré par l'OSCE n'est pas encore reconnu...

L'approche par la convergence des méthodes de travail montre ses limites.

Cette approche par la convergence des méthodes de travail, plutôt que par des structures et une organisation de concertation commune, montre donc ses limites, surtout lors des périodes de crise de l'intégration européenne. Malgré les tensions politiques, De Michelis et Chantraine (2003) relatent que les relations entre les INS demeurent « très bonnes », autorisant « une harmonisation en douceur des projets statistiques communs aux six pays, passant par les décisions des experts et des directeurs

généraux [des INS] ». C'est donc **une harmonisation essentiellement informelle** qui s'opère entre les six pays fondateurs de la future Union européenne. Elle ne s'appuie sur aucune disposition légale et l'architecture balbutiante du futur système statistique européen ne repose alors que sur les comités sectoriels et sur la réunion semestrielle des directeurs généraux des INS.

¹¹ Rolf Wagenführ est alors à la retraite. Il a été remplacé en 1966 par Raymond Dumas, ancien cadre de l'Insee.

► Encadré 1. Périodes de crise de l'intégration européenne

Initié en 1951, le processus d'intégration européenne a traversé plusieurs périodes de crise. Chacune a eu des spécificités et des impacts différents, mais elles ont toutes contribué à façonner l'Union européenne dans sa forme actuelle. Les principales périodes de crise identifiées sont les suivantes :

1954-1957 : Cette période est marquée par le rejet par le Parlement français de la Communauté européenne de défense (CED), ce qui met un coup d'arrêt aux ambitions fédéralistes de l'époque.

1963-1969 : Cette période est caractérisée par les deux rejets de la candidature britannique à la CEE par le président français Charles de Gaulle (en 1963 et en 1967) et par la « crise de la chaise vide » (1965-1966), lors de laquelle la France a boycotté les réunions de la CEE en raison de désaccords sur le budget et sur la politique agricole commune (PAC).

1973-1985 : Cette période a été marquée par des tensions budgétaires et des crises liées à l'adhésion de la Grande-Bretagne (en 1973), ainsi que par des réformes institutionnelles qui n'ont pas toujours abouti à des résultats concrets.

1992-1995 : La ratification difficile du traité de Maastricht* ouvre une nouvelle période de crise, avec des débats intenses sur la souveraineté nationale et l'intégration européenne.

Depuis 2005 : La crise contemporaine de l'UE débute avec le rejet du traité constitutionnel européen par la France et les Pays-Bas. Cette période est marquée par des défis multiples, notamment : la crise de la zone euro (déclenchée par la crise financière mondiale de 2008), le Brexit (sortie du Royaume-Uni de la zone euro effective à partir du 1^{er} janvier 2021), et des tensions politiques internes au sein des États membres.

* Voir les références juridiques en fin d'article. Voir également le lien : <https://www.touteurope.eu/fonctionnement-de-l-ue/le-trait%C3%A9-de-maastricht-1992/>.

Les avancées sont cependant importantes dans le domaine des statistiques agricoles (**encadré 2**). La gestion de la PAC, mise en place en 1963, nécessite des statistiques très détaillées, à jour et surtout comparables. Sous l'impulsion du comité permanent de la statistique agricole, les statistiques harmonisées dans ce domaine se développent fortement¹². Plus largement, De Michelis et Chantraine (2003) soulignent que « seuls les domaines à forte composante politique comme l'agriculture (PAC) et le commerce extérieur (Union douanière et négociations commerciales du GATT) arrivent à faire progresser l'harmonisation statistique »¹³.

L'esprit singulier de la communauté des statisticiens aide à surmonter les crises politiques de l'Europe

Le développement de la statistique européenne est très fortement imbriqué à la construction de l'Europe et aux vicissitudes de cette dernière, qui soit impulsent un élan à l'élaboration d'indicateurs harmonisés (comme lors de la mise en place de la PAC), soit retardent sa mise en œuvre (comme lors des crises liées à l'adhésion de la Grande-Bretagne). Cependant, une des spécificités de la statistique européenne, qui perdure jusqu'à aujourd'hui, est **la communauté de partage qui lie l'ensemble de la profession statistique au niveau européen**. Davantage que le cadre réglementaire, c'est le partage de valeurs communes autour d'une statistique impartiale, fondée sur des approches scientifiques et – osons le mot – une même éthique professionnelle, qui établit la solidité du SSE (en construction ou actuel). Au point que celui-ci peut parfois se ressentir comme « une grande famille ». Il est en tout cas perçu comme tel par les acteurs extérieurs.

12 Au cours des années 1960, plusieurs textes juridiques relatifs à la statistique agricole sont approuvés par le Conseil des Communautés européennes afin d'améliorer l'harmonisation de certaines enquêtes (la production porcine, le lait et les produits laitiers, les œufs à couver et les poussins, le cheptel bovin, l'abattage des bovins de boucherie).

13 Le General Agreement on Tariffs and Trade (GATT, en français : accord général sur les tarifs douaniers et le commerce) a été signé le 30 octobre 1947 par 23 pays, pour harmoniser les politiques douanières des parties signataires.

► Encadré 2. L'émergence de la statistique agricole : un double défi pour l'OSCE

La gestion de la PAC, mise en place en 1963, nécessite de disposer de nombreuses statistiques harmonisées : sur la structure des exploitations agricoles, sur la production effective, sur les prix agricoles et sur le revenu des agriculteurs.

- Le premier défi pour l'OSCE résulte du défaut d'harmonisation des statistiques existantes. À la fois les calendriers des enquêtes et les méthodes employées divergent d'un État membre à l'autre. Ces derniers sont par ailleurs réticents à modifier leur calendrier d'enquêtes et/ou leurs définitions, car cela généreraient des coûts qui peseraient sur les budgets nationaux.

- Le second défi pour l'OSCE est de servir d'intermédiaire entre les gestionnaires de la PAC, qui ont besoin de données comparables et sans aucun délai, et les INS, qui doivent s'adapter à ces nouvelles exigences. Le défi est d'autant plus difficile à relever que l'OSCE souhaite préserver sa compétence exclusive en matière de production de statistiques au sein de la Commission des Communautés européennes. Or, la direction générale de l'agriculture n'hésite pas à financer directement des opérations de collecte de données dans les États membres, jugeant trop lente la réaction du système statistique européen en construction.

Dans ce contexte, l'outil législatif européen devient un catalyseur pour accélérer la comparabilité des statistiques agricoles. Ainsi, pour répondre aux attentes des gestionnaires de la PAC, l'OSCE a

d'abord recours à des « directives » du Conseil des Communautés européennes, textes qui fixent des objectifs aux États membres tout en leur laissant une marge dans la mise en œuvre. Ce fut le cas en 1971 sur le lait et les produits laitiers, puis en 1973 sur le cheptel bovin. L'OSCE recourt ensuite à des « règlements » du Conseil, textes directement applicables sans transposition. Un règlement est ainsi promulgué en 1976 sur les céréales.

Parallèlement, la direction générale de l'agriculture initie en 1968 le réseau d'information comptable agricole (RICA), encore actif aujourd'hui. Il s'agit d'une enquête communautaire qui suit l'évolution des résultats économiques des exploitations agricoles dans les États membres. Le RICA est alors confié aux ministères nationaux de l'Agriculture.

Dans les années 1960, l'OSCE privilégie le recours aux enquêtes plutôt qu'aux données administratives, disponibles trop tardivement. La mise en place des enquêtes est accompagnée de financements substantiels pour les ministères de l'Agriculture, ce qui permet de consolider leurs services statistiques. À la fin des années 1970, l'OSCE, désormais appelé plus communément Eurostat, change de stratégie. Jusqu'à présent, il finançait de manière systématique des enquêtes spécifiques dans tous les États membres, y compris dans les pays où ces enquêtes préexistaient. Désormais, il décide de financer davantage des plans de restructuration des systèmes statistiques dans les pays les plus en retard.

Ce partage de valeurs communes ne signifie pas, néanmoins, l'absence de disparités au sein des modèles statistiques européens. On pense ici, en particulier, à la différence encore prégnante – quoique de moins en moins pertinente – entre les systèmes nationaux fondés historiquement sur l'utilisation de registres administratifs et ceux fondés principalement sur des enquêtes statistiques. Ce partage n'est pas davantage une garantie contre les atteintes ou les manquements à ces valeurs communes, qu'ils viennent de l'extérieur ou parfois même de l'intérieur. Enfin, ce socle commun et le processus d'harmonisation progressive ne sont pas réduits à maintenir un plus petit dénominateur commun : la communauté statistique européenne a tout au long de son histoire démontré sa volonté de suivre des standards toujours plus élevés quant à la qualité des statistiques produites et à l'exigence d'un environnement qui protège son indépendance professionnelle.

Préserver la confidentialité des données : un enjeu essentiel dans le cadre de l'émergence du SSE

En 1973, le Danemark, l'Irlande et le Royaume-Uni rejoignent les Communautés européennes, qui comptent à présent neuf États membres. **L'OSCE est désormais surnommé Eurostat.** Un des sujets de longue date auquel Eurostat s'attelle à l'époque est celui du secret statistique. C'est en effet la clef de voûte du système

“ Un des sujets de longue date auquel Eurostat s'attelle est celui du secret statistique.

d'information et d'accès aux données à partir desquelles les statistiques sont élaborées. Dans cette perspective, les systèmes nationaux de statistiques, et plus largement le système européen, doivent être capables de protéger la confidentialité des données qui leur sont confiées (Redor, 2023). De Michelis et Chantraine (2003) racontent qu'au début de la construction du SSE, le problème était de savoir si l'OSCE pouvait donner aux INS les informations de

base collectées directement auprès de l'industrie sidérurgique. En effet, avec le traité de Paris, l'OSCE pouvait organiser directement des enquêtes statistiques auprès des entreprises des secteurs du charbon et de l'acier des pays membres, sans passer par les INS. Cependant, à partir de 1958, avec le traité de Rome instituant la CEE, l'organisation s'est renversée. Le problème est désormais de savoir si les INS peuvent transmettre des données confidentielles à l'OSCE.

Le sujet revient de manière récurrente, chaque fois que l'OSCE/Eurostat doit produire un agrégat sur l'ensemble de la Communauté et que certains États membres refusent de transmettre leurs données au prétexte de la confidentialité. Aussi, en 1976, Eurostat confie à un expert français, Gérard Ader, une étude sur les problèmes du secret statistique. L'expert préconise d'alléger les règles pour ce qui concerne la transmission à Eurostat (Ader, 1976)¹⁴. Le rapport est présenté à la conférence des DGINS de mai 1976. La discussion révèle que certains États membres craignent qu'Eurostat se retrouve obligé de transmettre ces données aux autres services de la Commission. Un groupe de travail est alors constitué pour examiner le rapport Ader et apporter des solutions à ce problème de la confidentialité et de la préservation du secret statistique.

Le programme statistique, outil de dialogue entre les INS, préfigure l'architecture future du SSE

À partir du milieu des années 1970, l'élargissement aux nouveaux États membres et les mauvaises relations d'Eurostat avec les services de la Commission affaiblissent l'esprit communautaire au sein des statisticiens européens (De Michelis et Chantraine, 2003). Pour y remédier, le commissaire en charge d'Eurostat, Ralph Dahrendorf, adresse à la Commission un document dans lequel il propose le cadre général d'un « programme statistique » que la Commission doit s'engager à faire appliquer. L'objectif est de mettre en place un programme pluriannuel, en commençant par la période 1976-1978. Ce programme, dont le principe est au fondement du système statistique européen, devient un véritable instrument de dialogue et de décision. Les discussions menées dans le cadre des comités sectoriels étaient en effet trop étroites et techniques pour créer la cohésion d'ensemble nécessaire au développement des statistiques européennes. Le programme statistique est quant à lui négocié entre, d'une part, les directeurs des INS qui disposent d'une vue d'ensemble des travaux à réaliser et, d'autre part, l'exécutif européen qui centralise les demandes des utilisateurs pour le suivi des politiques européennes. Pour renforcer la coopération nécessaire à l'harmonisation des statistiques, la Commission envisage également dès 1971 la création de deux comités :

¹⁴ Un article publié dans la revue Économie et Statistique reprend les thèmes principaux de ce rapport (Ader, 1978).

Le commissaire en charge d'Eurostat, Ralph Dahrendorf, propose à la Commission le cadre général d'un « programme statistique » que la Commission doit s'engager à faire appliquer.

le comité du programme statistique (CPS), pour gérer à un plus haut niveau les arbitrages sur le programme statistique, et le comité consultatif européen de l'information statistique dans les domaines économique et social (CEIES), pour mieux prendre en compte les besoins des utilisateurs et les coûts des producteurs. Nous reviendrons plus bas sur ces deux comités : il faudra encore attendre près de vingt ans avant qu'ils ne soient établis et n'échafaudent la future structure du SSE.

Au début des années 1980, l'Europe semble s'essouffler sous l'effet des crises de l'intégration européenne (**encadré 1**) et de l'affaiblissement de la croissance économique, dans un contexte international marqué par les deux chocs pétroliers (1973 et 1979) et la forte fluctuation des monnaies. Mais la nouvelle Commission Jacques Delors (1985-1988) va lui donner un nouveau souffle, avec notamment l'Acte unique européen¹⁵ et des réformes institutionnelles qui auront aussi des conséquences favorables à terme pour le développement du SSE¹⁶. Dans un premier temps, cependant, la suppression des contrôles aux frontières intracommunautaires entraînera considérablement les travaux des statisticiens européens, notamment pour établir les indicateurs du commerce extérieur, de la balance des paiements et les comptes nationaux des États membres.

► 1986-2003 : l'architecture et l'assise juridique du SSE se mettent en place

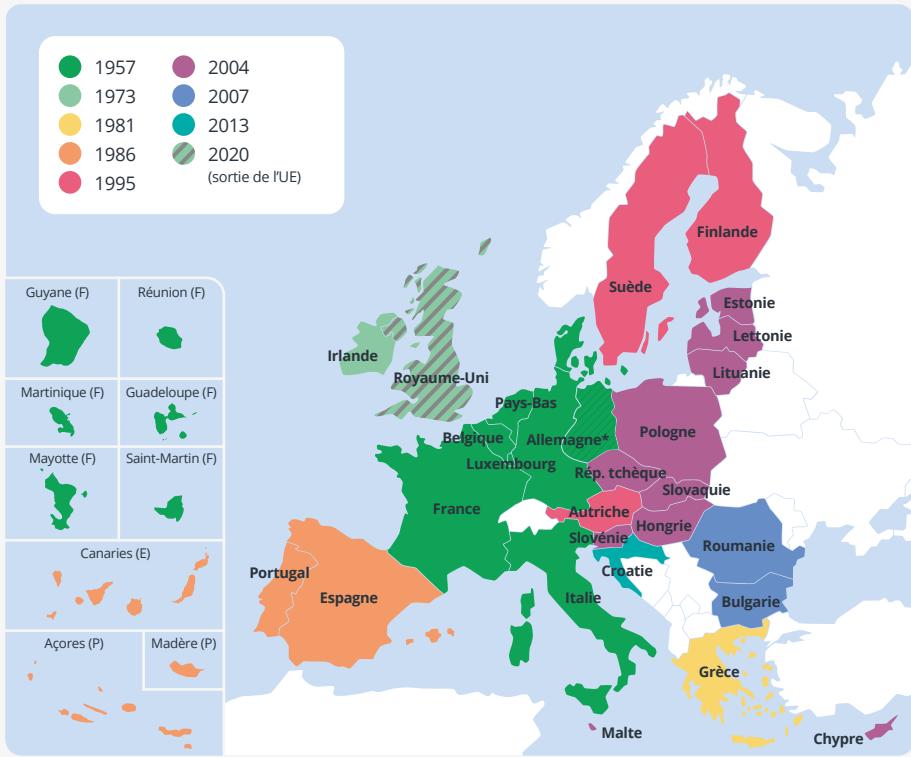
En 1986, les Communautés européennes comptent douze membres, après l'adhésion de la Grèce en 1981 et de l'Espagne et du Portugal en 1986 (**figure 2**). Dans ce contexte d'élargissement, et pour faire face aux exigences du futur marché unique, la statistique doit s'organiser de façon plus formelle¹⁷.

¹⁵ Voir les références juridiques en fin d'article. L'Acte unique européen vise à rendre effectives les libertés de circulation accordées en droit par le traité instituant la CEE sous le nom de « marché commun ». Il prévoit de mener à terme, avant le 1^{er} janvier 1993, la réalisation de ce marché commun, rebaptisé « marché unique » : un « espace sans frontières intérieures dans lequel la libre circulation des marchandises, des personnes, des services et des capitaux est assurée » (article 7 du traité instituant la CEE).

¹⁶ Le Conseil des ministres de l'UE décide désormais à la majorité qualifiée dans plusieurs domaines stratégiques pour l'achèvement du marché unique : tarif douanier extérieur, libre prestation de services, libre circulation des capitaux, transports maritimes et aériens, harmonisation des législations. De nouveaux domaines de compétences, comme la recherche et l'environnement, obéissent également à ce principe, ce qui facilite la prise de décision entre les États membres. Les questions fiscales et sociales, à l'exception de la santé et de la sécurité des travailleurs, sont en revanche toujours régies par la règle de l'unanimité.

¹⁷ Alastalo (2023) considère également que la préparation du marché unique marque un tournant dans l'histoire du SSE, dans laquelle elle distingue deux périodes : « Durant la première, qui va du début des années 1950 aux années 1990, l'office œuvre à harmoniser les concepts et les méthodes statistiques afin que les États puissent fournir des données commensurables. Durant la seconde, le développement d'indicateurs et la stabilisation législative du système statistique constituent la nouvelle priorité. » Franchet (2014) évoque quant à lui qu'à partir de 1987, « anticipant la croissance exponentielle de la demande de statistiques d'une Europe en intégration politique accélérée, une « statistique européenne » s'est construite d'une façon énergique dans un projet coopératif en réseau ».

► Figure 2 - Élargissement de l'Union européenne : les étapes



* Réunification allemande en 1990.

Note : Les lettres entre parenthèses derrière les départements et territoires indiquent les pays auxquels ils sont rattachés : E pour Espagne, F pour France, ou P pour Portugal (P).

Source : toutleurope.eu.

L'élargissement des Communautés européennes pousse à mieux structurer les relations entre les États membres

Au cours des années 1989 à 1991, la Commission crée successivement cinq grands comités : le comité du programme statistique (CPS), le comité européen du secret statistique, le comité directeur de la statistique (CDIS), le comité consultatif européen de l'information statistique dans les domaines économique et social (CEIES) et le comité des statistiques monétaires, financières et de la balance des paiements (CMFB) (**figure 3**). Ces cinq comités vont chacun, avec leur spécificité, faire intervenir les différents partenaires dans le processus de programmation des travaux statistiques :

• Le comité du programme statistique (CPS) :

De manière générale, pour gérer les politiques communautaires, les institutions européennes prennent une orientation plus juridique en mettant en place le système

En 1989, la Commission propose de mettre en place un comité de gestion du programme statistique impliquant les États membres.

dit de « comitologie »¹⁸. À douze, il est en effet devenu plus difficile de travailler sous forme de *gentleman's agreement*, où chaque pays reste libre de s'engager ou non (De Michelis et Chantraine, 2003). De même, pour la construction des statistiques européennes, la conférence des DGINS a montré ses limites dans le processus de programmation des travaux. De ce fait, en 1989, la Commission propose de mettre en place un comité de gestion du programme statistique impliquant les États membres. La décision n° 89/382 du Conseil instituant un comité du programme

statistique des Communautés européennes¹⁹ crée ainsi le CPS. Cette décision du Conseil marque en même temps une insertion de la statistique dans l'architecture juridique de la construction européenne. La conférence des DGINS devient alors davantage un séminaire de réflexion sur les orientations stratégiques de la statistique européenne.

- **Le comité européen du secret statistique :**

Ce comité est créé par l'article 7 du règlement n° 1588/90 du Conseil du 11 juin 1990 relatif à la transmission à l'Office statistique des Communautés européennes d'informations statistiques couvertes par le secret²⁰. Il est composé d'un représentant de chaque État membre et est présidé par un représentant de la Commission (le directeur général de l'OSCE ou une personne désignée par lui). Il veille au respect des mesures mises en œuvre par l'OSCE pour garantir la confidentialité des données échangées avec les INS. Il s'assure notamment que l'OSCE prend toutes les mesures administratives, techniques et opérationnelles nécessaires à cette fin. Le règlement vise en particulier à rassurer les INS, qui n'auront plus de raison d'invoquer des dispositions nationales relatives au secret statistique pour refuser la transmission de données confidentielles à l'OSCE.

- **Le comité directeur de la statistique (CDIS) :**

Interne à la Commission, le CDIS réunit les hauts responsables des directions générales pour examiner les développements du programme statistique. L'objectif est d'établir les besoins des directions générales en matière de statistiques et de fixer les priorités concernant la réalisation des projets. Il se substitue à la procédure de consultation interservices sur la programmation annuelle des travaux d'Eurostat. Ce comité ne fonctionnera toutefois efficacement que pendant deux à trois ans. Il s'étiolera peu à peu pour se transformer en un forum d'échanges.

- **Le comité consultatif européen de l'information statistique dans les domaines économique et social (CEIES) :**

Dans un contexte de diffusion de plus en plus élargie des statistiques européennes, Eurostat nourrissait l'ambition de se mettre au service de la société européenne dans son ensemble. C'est ainsi que naît le CEIES, qui deviendra ultérieurement le comité consultatif européen de la statistique (ESAC pour *European Statistical Advisory Committee*)²¹,

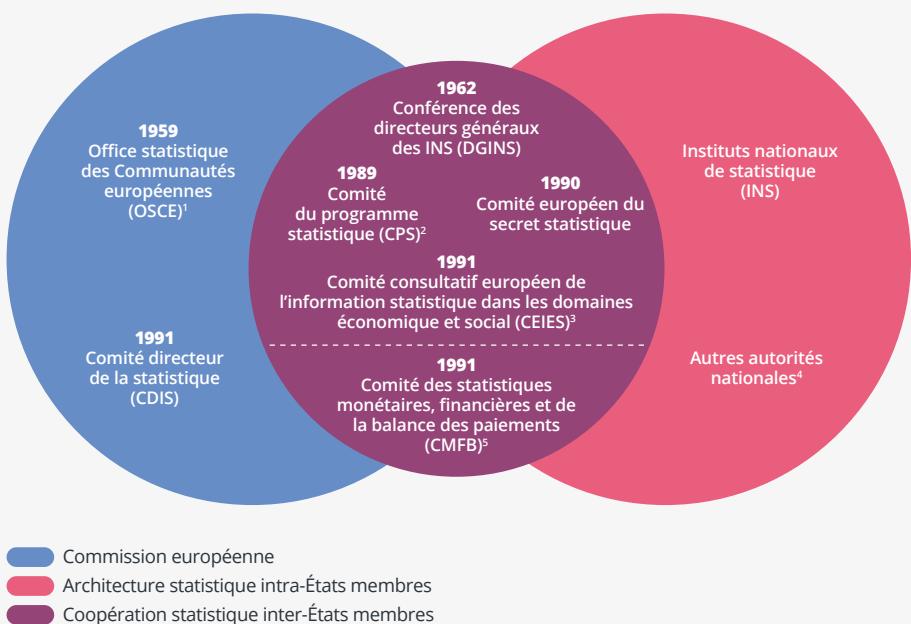
¹⁸ Voir le lien : https://commission.europa.eu/law/law-making-process/adopting-eu-law/implementing-and-delegated-acts/comitology_fr.

¹⁹ Voir les références juridiques en fin d'article.

²⁰ Voir les références juridiques en fin d'article.

²¹ Voir le lien : <https://ec.europa.eu/eurostat/web/european-statistical-system/governance-bodies/esac>.

► **Figure 3 - Prémices de l'architecture du système statistique européen au début des années 1990**



1. Plus connu sous le nom d'Eurostat depuis 1973.

2. Actuellement Comité du système statistique européen (CSSE).

3. Actuellement *European Statistical Advisory Committee* (ESAC).

4. ONAs (*Other National Authorities*). En France, dix services statistiques ministériels en font partie.

5. Ce comité réunit non seulement les représentants des INS mais aussi ceux des banques centrales.

équivalent du Conseil national de l'information statistique (Cnis) en France (Anxionnaz et Maurel, 2021). Ce comité consultatif voulu par Jacques Delors a pour mission d'assister le Conseil et la Commission dans la coordination des objectifs fixés en matière de politique de l'information statistique communautaire. Il réunit des représentants des utilisateurs et des producteurs et émet des avis prenant en compte à la fois les besoins des premiers et les coûts supportés par les seconds. Il va contribuer à démarquer Eurostat des autres directions de la Commission. Dans un premier temps, les avis rendus par ce comité seront mal perçus par les juristes du Conseil qui s'estimaient être les seuls légitimes à pouvoir conseiller la Commission (De Michelis et Chantraine, 2003).

• **Le comité des statistiques monétaires, financières et de la balance des paiements (CMFB) :**

Dans un contexte de relance de l'idée d'une Union économique et monétaire (voir infra), il convient de produire des indicateurs économiques et financiers fiables pour assurer la convergence des politiques économiques. Dans un grand nombre de pays, ces indicateurs sont produits par les banques centrales. Or, Eurostat a beaucoup investi dans la coordination entre le compte extérieur des comptes nationaux et le calcul de la balance des paiements. Il choisit alors de créer le CMFB pour réunir statisticiens des INS et des banques centrales.

De façon générale, l'approfondissement des politiques existantes des Communautés européennes et l'extension des politiques à de nouveaux domaines se traduisent par des besoins toujours plus grands en statistiques européennes... auxquels le SSE en construction doit répondre.

Le traité de Maastricht engage l'Europe vers une monnaie unique et place les statisticiens sur la brèche

Le 7 février 1992, les douze États membres de la CEE signent le traité de Maastricht qui crée l'Union européenne. Il organise aussi la création au 1^{er} janvier 1999 de l'Union économique et monétaire (UEM) entre les pays membres de l'UE qui satisferont aux critères de convergence en matière de prix, de finances publiques, de taux d'intérêt et de taux de change. Le suivi des indicateurs sous-jacents à ces critères nécessite le développement de statistiques encore plus harmonisées. Les statisticiens européens et leurs indicateurs sont alors soumis à rude épreuve.

Afin d'éviter les batailles de chiffres, le système européen de comptabilité (SEC) devient obligatoire dans chaque pays²². En effet, Eurostat s'était aperçu que la majorité des pays membres n'avait pas cette référence pour leurs comptes nationaux, le SEC n'étant utilisé que pour la transmission des données à Eurostat. Ce fonctionnement introduisait de nombreuses distorsions dans l'interprétation et l'application des règles comptables et des définitions, lesquelles donnaient lieu à d'importantes différences de résultats entre les pays membres. Cela n'était plus acceptable dès lors que la comptabilité nationale devenait la référence dans la mise en œuvre de politiques communautaires de plus en plus nombreuses. En érigéant le taux d'inflation au rang des critères de convergence les plus stricts, le traité de Maastricht allait aussi rendre impératif une harmonisation plus grande du calcul de l'inflation, les indices nationaux étant trop liés à des négociations nationales salariales ou contractuelles. Eurostat propose alors la création de l'**indice des prix à la consommation harmonisé (IPCH)**²³.

Le programme statistique, catalyseur de la négociation entre les INS, trouve enfin une légitimité juridique

Le programme statistique pour la période 1989-1992 relève pour la première fois d'une recommandation du Conseil. Il devient à ce titre un document juridique et non plus un document pour information, bien souvent perçu comme un catalogue de bonnes intentions. Des ressources budgétaires y sont par ailleurs associées : pour la première fois, les travaux statistiques sont financés au titre du budget des études de la Commission. Ainsi, les projets statistiques sont désormais basés sur des actes juridiques, ce qui renforce leur légitimité au regard du développement de la construction européenne et permet également d'obtenir des financements. Jusqu'en 1990, à l'exception des domaines

²² Cette obligation est l'objet du règlement (CE) n° 2223/96 relatif au système européen des comptes nationaux et régionaux dans la Communauté (voir les références juridiques en fin d'article). Le Conseil approuva en juin 1996 ce règlement, dont une première base avait été proposée en 1995 par Eurostat, avant que quelques discussions techniques supplémentaires améliorent le contenu du SEC. Le nouveau SEC est mis en œuvre à partir de 1999 par les États membres de l'UE.

²³ Voir le lien : <https://www.insee.fr/fr/metadonnees/source/indicateur/p1654/description>.

relevant de la politique agricole et des échanges extérieurs, les travaux statistiques ne reposaient sur aucun fondement juridique²⁴. Cette pratique est devenue depuis lors la règle pour les statistiques européennes.

La loi statistique de 1997 : un outil essentiel pour structurer le SSE

L'explosion des demandes d'informations statistiques, dont en particulier l'utilisation de plus en plus grande de celles-ci à des fins de surveillance des politiques communautaires (de Peretti et Touchelay, 2024), exigeait une organisation plus formelle de la statistique communautaire. Le 17 février 1997, après plusieurs années de négociations dans les groupes de travail d'Eurostat et du Conseil, ce dernier adopte le règlement (CE) n° 322/97 relatif à la statistique communautaire²⁵. Dans le jargon des statisticiens, on l'appelle « la loi statistique ». Cette dernière constitue un puissant levier pour le système statistique européen : elle précise le processus de programmation, les principes fondamentaux d'élaboration des statistiques, notamment d'impartialité et d'indépendance, et les principes de diffusion et de garantie de la confidentialité.

Avant son adoption, le texte a été examiné et discuté en profondeur dans tous les milieux politiques intervenant dans le processus décisionnel : Commission, Conseil, Parlement européen, CEIES, Institut monétaire européen (IME)²⁶ et les autorités nationales des États membres. Désormais, la place de la statistique dans la construction communautaire est reconnue publiquement et juridiquement. En 1997, la Commission désigne Eurostat comme la seule « autorité communautaire » chargée de la production statistique.

La statistique élevée au rang des politiques européennes relevant de la codécision !

Nouvelle étape de la construction européenne, le traité d'Amsterdam²⁷, signé en 1997 et entré en vigueur en 1999, consolide les dispositions du traité de Maastricht. Il introduit de nouvelles politiques relevant de l'UE, telles que la politique sociale, et renforce les

pouvoirs du Parlement européen : désormais, dans le cadre de la procédure de codécision, le Parlement européen et le Conseil deviennent colégislateurs pratiquement sur un pied d'égalité.

Avec le traité d'Amsterdam, la statistique devient un sujet à part entière de la politique européenne et relève désormais de la codécision. Le traité contient un article dédié spécifiquement à la statistique communautaire, divisé en deux paragraphes :

Avec le traité d'Amsterdam, la statistique devient un sujet à part entière de la politique européenne et relève désormais de la codécision.

²⁴ Deux actes juridiques figuraient dans ce nouveau programme : une directive pour le calcul du produit national brut (PNB) pour la gestion des ressources propres de la Communauté et le règlement pour l'établissement de la nomenclature statistique des activités économiques dans la Communauté européenne (NACE).

²⁵ Voir les références juridiques en fin d'article.

²⁶ Crée par le traité de Maastricht, l'IME constituait une étape intermédiaire en vue de la création de la Banque centrale européenne (BCE).

²⁷ Voir les références juridiques en fin d'article.

- Le premier paragraphe dispose qu'il revient au Parlement européen et au Conseil d'arrêter des mesures en vue de l'établissement de statistiques lorsque cela est nécessaire à l'accomplissement des activités de l'UE et conformément à la procédure législative ordinaire. Cette nouvelle base juridique générique se substitue à toutes les bases juridiques spécifiques et sectorielles utilisées jusque-là. Elle représente une avancée majeure pour la statistique européenne. Elle participe du mouvement général qui vise à élargir le champ de la codécision et à renforcer les pouvoirs du Parlement européen. Elle imprime une nouvelle logique d'intervention qui voit le Parlement, au côté du Conseil, décider de la législation statistique européenne. Elle est à l'origine du développement massif d'un corpus législatif européen dans le domaine statistique, totalement inédit et sans équivalent au niveau national.

Cette logique de juridiction de la matière statistique au niveau européen rencontre également le besoin de sécurité juridique de certains États membres, qui poussent à établir leurs droits et obligations dans le droit de l'UE, afin de pouvoir s'en prévaloir pour obtenir, au niveau national, les moyens nécessaires à leur réalisation. C'est le cas notamment de l'Allemagne. Ce mouvement entraîne aussi l'intervention croissante du juge européen dans ce nouvel ordre juridique statistique européen, à la demande d'États membres soucieux de faire respecter leurs droits face à la Commission européenne ou à Eurostat.

- Le second paragraphe établit les principes sur lesquels doit se fonder la production statistique. Il est également capital dans l'histoire du SSE. Pour la première fois, les principes d'impartialité, de fiabilité, d'objectivité, d'indépendance scientifique, ainsi que d'efficacité au regard du coût et de la confidentialité des informations statistiques, sont directement reconnus au niveau des traités de l'UE. Ils s'imposent ainsi aux États membres sans que la législation nationale puisse y faire obstacle.

► Les grands jalons du système statistique européen sont posés

En 2003, les grands jalons du futur système statistique européen sont donc posés, entre la création d'Eurostat, l'organisation pour la planification des travaux, l'installation des grands comités et l'adoption de la loi statistique de 1997. Ces premiers jalons sont donc le fruit d'une opération de longue haleine : un demi-siècle ! L'histoire a été complexe, la construction difficile, d'autant plus que le périmètre de la statistique européenne n'a cessé d'évoluer, au gré de l'élargissement à d'autres pays et à d'autres champs politiques. Les défis ont cependant été relevés. Les valeurs communes que partagent les statisticiens ont sans aucun doute facilité les avancées. Les impulsions politiques qui ont permis de passer de la CECA à l'UE ont été relayées au niveau des statistiques.

Fort des avancées réalisées au cours des cinquante premières années, la communauté statistique se prépare à répondre aux nouveaux besoins que l'approfondissement continual des politiques européennes entraîne sans cesse. Le « grand élan » va cependant être coupé net en 2004 avec la révélation de la falsification des comptes publics grecs (Ernoult, 2012). Les fortes révisions du déficit et de la dette publics de ce pays vont entacher la crédibilité générale des statistiques européennes et l'intégrité de la communauté statistique. Ce sera un nouveau défi à relever, qui conduira vers d'autres évolutions. Mais c'est une autre histoire, pour un autre numéro du Courrier des statistiques...

► Fondements juridiques

- Traité instituant la Communauté européenne du charbon et de l'acier (CECA), dit Traité de Paris. In : *site de l'Union européenne*. [en ligne]. 18 avril 1951. [Consulté le 31 octobre 2025]. Disponible à l'adresse :
<https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=LEGISSUM%3Axy0022>.
- Traité instituant la Communauté économique européenne (CEE), dit Traité de Rome. In : *site de l'Union européenne*. [en ligne]. 25 mars 1957. [Consulté le 31 octobre 2025]. Disponible à l'adresse :
<https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=LEGISSUM:xy0023>.
- Traité instituant la Communauté européenne de l'énergie atomique, dit Traité Euratom. In : *site de l'Union européenne*. [en ligne]. 25 mars 1957. [Consulté le 31 octobre 2025]. Disponible à l'adresse :
<https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:12012A/TXT>.
- Acte unique européen. In : *site de l'Union européenne*. [en ligne]. 17 et 28 février 1986. [Consulté le 31 octobre 2025]. Disponible à l'adresse :
<https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:11986U>.
- Traité de Maastricht sur l'Union européenne. In : *site de l'Union européenne*. [en ligne]. 7 février 1992. [Consulté le 31 octobre 2025]. Disponible à l'adresse :
<https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=LEGISSUM:xy0026>.
- Traité d'Amsterdam modifiant le traité sur l'Union européenne. In : *site de l'Union européenne*. [en ligne]. 2 octobre 1997. [Consulté le 31 octobre 2025]. Disponible à l'adresse : <https://eur-lex.europa.eu/FR/legal-content/summary/the-amsterdam-treaty.html>.
- Décision (CEE, Euratom) n° 89/382 du Conseil du 19 juin 1989 instituant un comité du programme statistique des Communautés européennes. In : *site de l'Union européenne*. [en ligne]. [Consulté le 13 octobre 2025]. Disponible à l'adresse : <https://eur-lex.europa.eu/legal-content/FR/TXT/?qid=1746440353970&uri=CELEX%3A31989D0382>.
- Règlement (Euratom, CEE) n° 1588/90 du Conseil du 11 juin 1990 relatif à la transmission à l'Office statistique des Communautés européennes d'informations statistiques couvertes par le secret. In : *site de l'Union européenne*. [en ligne]. [Consulté le 13 octobre 2025]. Disponible à l'adresse : <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A31990R1588&qid=1749649474352>.
- Règlement (CE) n° 2223/96 du conseil du 25 juin 1996 relatif au système européen des comptes nationaux et régionaux dans la Communauté. In : *site de l'Union européenne*. [en ligne]. [Consulté le 13 octobre 2025]. Disponible à l'adresse :
<https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:31996R2223>.
- Règlement (CE) n° 322/97 du Conseil du 17 février 1997 relatif à la statistique communautaire. In : *site de l'Union européenne*. [en ligne]. [Consulté le 13 octobre 2025]. Disponible à l'adresse : <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A31997R0322&qid=1760342561620>.

- Règlement (CE) n° 223/2009 du Parlement européen et du Conseil du 11 mars 2009 relatif aux statistiques européennes. In : *site de l'Union européenne*. [en ligne]. [Consulté le 13 octobre 2025]. Disponible à l'adresse : <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A32009R0223&qid=1760342713619>.

► Bibliographie

- ADER, Gérard, 1976. *Le secret et les statistiques d'entreprises dans la Communauté européenne*. Eurostat.
- ADER, Gérard, 1978. Secret et statistiques d'entreprises dans la communauté européenne. In : *Économie et statistique*. [en ligne]. Mai 1978. Insee. N° 100, pp. 85-87. [Consulté le 31 octobre 2025]. Disponible à l'adresse :
https://www.persee.fr/doc/estat_0336-1454_1978_num_100_1_3062.
- ALASTALO, Marja, 2023. Histoire d'Eurostat. In : *site de l'Encyclopédie d'histoire numérique de l'Europe*. [en ligne]. 19 avril 2023. [Consulté le 4 novembre 2025]. Disponible à l'adresse : <https://ehne.fr/fr/node/21803>.
- ANXIONNAZ, Isabelle et MAUREL, Françoise, 2021. Le Conseil national de l'information statistique – La qualité des statistiques publiques passe aussi par la concertation. In : *Courrier des statistiques*. [en ligne]. 8 juillet 2021. Insee. N° N6, pp. 123-142. [Consulté le 31 octobre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/5398693?sommaire=5398695>.
- DE MICHELIS, Alberto et CHANTRAYE, Alain, 2003. « Mémoires d'Eurostat – Cinquante ans au service de l'Europe ». In : *site d'Eurostat*. [en ligne]. 15 mai 2003. Eurostat. [Consulté le 10 octobre 2025]. Disponible à l'adresse :
<https://ec.europa.eu/eurostat/web/products-statistical-books/-/KS-49-02-183>.
- DE PERETTI, Gaël et TOUCHELAY, Béatrice, 2024. Statistiques publiques et débat démocratique : de nouvelles attentes et de nouveaux enjeux (1988-2016). In : *Courrier des statistiques*. [en ligne]. 8 juillet 2024. Insee. N° N11, pp. 11-30. [Consulté le 10 octobre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/8203034?sommaire=8203072>.
- ERNOULT, Thomas, 2012. Maquillage des comptes publics grecs : statistiques à la dérive ? In : *Regards croisés sur l'économie*. [en ligne]. 20 juin 2012. La Découverte. N° 11, pp. 113-115. [Consulté le 10 octobre 2025]. Disponible à l'adresse :
<https://shs.cairn.info/revue-regards-croises-sur-l-economie-2012-1-page-113?lang=fr>.
- EUROSTAT, 2025. Statistical Requirements Compendium – 2025 edition. In : *site d'Eurostat*. [en ligne]. 11 avril 2025. Eurostat. [Consulté le 31 octobre 2025]. Disponible à l'adresse :
<https://ec.europa.eu/eurostat/fr/web/products-manuals-and-guidelines/w/ks-01-25-009>.
- FRANCHET, Yves, 2014. Développement de la statistique de l'Union européenne : une action coopérative multinationale en réseau. In : *Statistique et société*. [en ligne]. Janvier 2014. Vol. 2, N° 1, pp. 15-19. [Consulté le 4 novembre 2025]. Disponible à l'adresse :
https://www.persee.fr/doc/staso_2269-0271_2014_num_2_1_889.
- INSEE, 1996. Cinquante ans d'Insee ou la conquête du chiffre (1946-1996). In : *site de la bibliothèque numérique de la statistique publique*. [en ligne]. [Consulté le 10 octobre 2025]. Disponible à l'adresse :
<https://www.bnsp.insee.fr/ark:/12148/bpt6k9124885q/f1.item.r=mireille%20moutardier>.

- LESUR, Romain, 2025. Sources de données privées : panorama et perspectives. In : *Courrier des statistiques*. [en ligne]. 23 juin 2025. Insee. N° N13, pp. 73-94. [Consulté le 31 octobre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/8546943?sommaire=8546949>.
- MULLER, Pierre, 2024. André Vanoli, au cœur des évolutions de la comptabilité nationale, française comme mondiale, tout au long des 50 dernières années. In : *Variances*. [en ligne]. 2 février 2023. [Consulté le 4 novembre 2025]. Disponible à l'adresse :
<https://variances.eu/?p=7228>.
- REDOR, Patrick, 2023. Confidentialité des données statistiques : un enjeu majeur pour le service statistique public. In : *Courrier des statistiques*. [en ligne]. 30 juin 2023. Insee. N° N9, pp. 46-63. [Consulté le 10 octobre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/7635823?sommaire=7635842>.
- WAGENFÜHR, Rolf, 1967. *La statistica in Europa*. Ferro.

Des données de papier

Une histoire des échanges de publications statistiques



Camille Beaurepaire* et Thierry Boucher**

L'échange international de données entre institutions statistiques est aujourd'hui une pratique banale, qui s'inscrit dans une longue histoire. À une époque où les données étaient de papier, quels étaient les enjeux et modalités de cet échange ? C'est en montrant les convergences et différences entre métiers des bibliothèques et des statistiques que cet article entend répondre à cette question.

Les publications statistiques étrangères constituent aux XIX^e et XX^e siècles une part très importante des fonds des bibliothèques de statistique. Elles servaient aux statisticiens pour produire des statistiques internationales comparées, ou pour s'inspirer de méthodes étrangères – mais étaient aussi utiles à un public externe. L'acquisition de ces documents supposait une infrastructure robuste d'échanges postaux, qui a donné lieu à de nombreux ajustements au XIX^e siècle. Une fois ces documents acquis, le travail des bibliothécaires consistait à les intégrer au sein d'un plan de classement – ces métadonnées jouant de longue date un rôle crucial dans l'univers des bibliothèques.

Que reste-t-il aujourd'hui de ces pratiques d'échanges soutenues entre bureaux statistiques ? Le fonds ancien étranger de la bibliothèque de l'Insee Alain Desrosières en constitue un témoignage très utile pour la recherche en sciences sociales. Le traitement et la valorisation de ce fond contribuent à faire rayonner la statistique publique par son patrimoine.

 The international exchange of data between statistical institutions is now commonplace. It has a long history. At a time when data was paper-based, what were the issues and procedures involved in this exchange? This article attempts to answer this question by showing the similarities and differences between the professions of librarian and statistician.

In the 19th and 20th centuries, foreign statistical publications formed a very large part of the collections of statistical libraries. They were used by statisticians to produce international comparative statistics or to gain inspiration from foreign methods – but they were also useful to external audiences. Acquiring these documents required a robust infrastructure of postal exchanges, which led to numerous adjustments in the 19th century. Once these documents were acquired, librarians had to integrate them into a classification scheme – metadata has long played a crucial role in the world of libraries.

What remains today of this ongoing exchange between statistical offices? The patrimonial foreign collections of the Alain Desrosières INSEE library are a very useful testimony for social science research. The processing and valorization of these collections contribute to make official statistics shine through their heritage.

* Chef de la section « Valorisation et services aux publics », DDAR, Insee.
camille.beaurepaire@insee.fr

** Chef de la division Documentation, DDAR, Insee.
thierry.boucher@insee.fr

Bibliothécaires et statisticiens sont aujourd’hui deux professions se revendiquant des « données ». La cause est entendue pour la statistique, elle l’est peut-être moins en ce qui concerne l’univers des bibliothèques¹. Pourtant, le terme « données » y est très utilisé depuis plusieurs décennies, dans le cadre de la transformation numérique des pratiques documentaires. Il y revêt plusieurs sens, renvoyant non seulement à l’information contenue *dans* les ouvrages, mais aussi aux informations *sur* ces derniers. Selon l’arrêté du 22 décembre 1981 relatif à l’enrichissement du vocabulaire de l’informatique², on appelle « donnée » toute « représentation d’une information sous une forme conventionnelle destinée à faciliter son traitement ». De ce point de vue, les éléments d’une notice bibliographique (auteur, titre, date et maison d’édition, etc.) sont bien des données. C’est en ce sens que l’on parle couramment des données d’un catalogue de bibliothèque.

Au regard de leur histoire longue, les savoirs des bibliothèques et des statistiques ont des origines bien distinctes (Desrosières, 1993 ; Barbier, 2021). Nous allons voir cependant que ces deux fils de tradition savante ont été tressés ensemble à partir de la fin du XIX^e siècle. À cette époque apparaît un nouveau courant intellectuel : celui de la documentation

(Fayet-Scribe, 2000 ; Gardey, 2008). Ce courant transforme les pratiques en bibliothèque, avec l’invention d’un outil novateur : le catalogue sur fiches. En son sein, chaque ouvrage est désormais représenté par une unique fiche cartonnée, sur laquelle sont reportées les informations bibliographiques. Ce catalogue sur fiches, véritable microcosme des collections physiques, peut être facilement manipulé pour accéder à l’information recherchée.

Cette invention modifie à son tour l’univers des statisticiens. Quand le statisticien américain Billings souffle dans les années 1880 à son collègue Hollerith l’idée de transformer le bulletin de

recensement en carte perforée³, c’est du système des fiches en bibliothèque qu’il s’inspire (Heide, 2009, p.23). L’introduction de ce nouveau médium des données bouleverse en quelques décennies le paysage des statistiques officielles, aboutissant à partir des années 1940 à l’administration mécanographique⁴ de grands registres.

De fait, l’univers du répertoire statistique (Rivière, 2022) n’est génétiquement pas si éloigné de celui du catalogue bibliographique. De la documentation à la donnée, en passant par l’information (Tricot, 2008), statisticiens et bibliothécaires ont tracé des sillons en partie parallèles depuis cette époque. Cette similarité est notamment illustrée par

¹ Parmi les institutions patrimoniales, les bibliothèques conservent tous les documents graphiques et audiovisuels édités (notamment les publications imprimées, comme les livres et les périodiques), mais aussi des manuscrits antérieurs à l’invention de l’imprimerie et des archives privées. Les musées conservent des objets d’intérêt historique, artistique, scientifique ou technique. Les archives conservent l’ensemble des documents (y compris les données), quels que soient leur forme et leur support, produits ou reçus par des services ou organismes publics ou privés dans l’exercice de leur activité.

² Voir les références juridiques en fin d’article.

³ La carte perforée permettait de coder, stocker et récupérer des données en mode binaire (0 : absence de trou, 1 : présence de trou).

⁴ La mécanographie désigne les techniques mécaniques et électromécaniques de traitement de l’information mobilisant par exemple des cartes perforées.

les bibliothèques consacrées à la statistique, comme la bibliothèque de l'Insee Alain Desrosières (**encadré 1**). Les ouvrages, principalement des publications statistiques, y sont des données à deux titres :

Tout au long du XIX^e siècle et pendant la première moitié du XX^e siècle, la bibliothèque est au cœur de l'univers de la donnée.

- Du point de vue des statisticiens, il s'agit d'un produit fini (output), prenant souvent à la fin du XIX^e siècle la forme de grands tableaux de chiffres (et s'apparentant ainsi à une forme primitive de base de données, non informatique).
- Du point de vue des bibliothécaires, il s'agit de l'intrant (input) sur lequel vont s'exercer les pratiques professionnelles, et à partir duquel seront produites de nouvelles données (bibliographiques).

Tout au long du XIX^e siècle et pendant la première moitié du XX^e siècle, la bibliothèque est au cœur de l'univers de la donnée. La publication papier est alors le moyen le plus commode de communiquer la donnée statistique, un rôle aujourd'hui dévolu aux bases informatiques. Cet article retrace le parcours, ou voyage (Leonelli et Tempini, 2020), de ces données de papier particulières qu'ont été à l'époque les publications statistiques. Comment circulaient-elles, comment étaient-elles triées, quels étaient leurs usages ? Pour répondre à ces questions, nous nous intéresserons plus particulièrement aux publications statistiques étrangères, dont la bibliothèque Alain Desrosières possède un fonds ancien d'environ 75 000 documents. Le récit du

► Encadré 1. La bibliothèque Alain Desrosières

La [bibliothèque de l'Insee Alain Desrosières*](https://bibliotheque.insee.net) est spécialisée dans les domaines de l'économie, de la statistique et des sciences sociales (Groudev et de Saboulin, 2015 ; Arduin et Collignon, 2017 ; Bach, 2017). Elle a été créée en 1946 au même moment que l'Insee. Ses missions consistent à fournir aux agents de l'Insee la documentation dont ils ont besoin, à conserver l'ensemble des publications de l'institut et à donner accès à ses collections à un public élargi. Elle a hérité notamment à sa création des collections antérieures de la bibliothèque de la Statistique générale de la France (SGF)**.

En 2018, la Direction générale de l'Insee a quitté le bâtiment qu'elle occupait depuis 1974 à Malakoff pour s'installer dans un bâtiment neuf, le White, à Montrouge. La bibliothèque a de même déménagé et une nouvelle salle de lecture a été officiellement inaugurée le 13 mars 2019. Cela fut l'occasion de rendre hommage à [Alain Desrosières***](#) en donnant son nom à la bibliothèque.

La bibliothèque en chiffres :

- 80 000 notices de monographies ;
- 15 000 notices de titres de périodiques ;
- 55 000 documents sur la [Bibliothèque numérique de la statistique publique \(BNSP\)****](#) ;
- 9 réserves où sont conservés les ouvrages à la Direction générale de l'Insee ;
- 7 kilomètres linéaires dont 1,5 au [Centre technique du livre de l'enseignement supérieur \(CTLes\)*****](#).

Au sein des collections de la bibliothèque, le fonds ancien étranger représente environ 75 000 documents (près d'un kilomètre linéaire), publiés entre le XVIII^e siècle et 1946. Il couvre plus de 90 pays, et inclut une vingtaine de langues différentes, dont plusieurs en alphabets non latins (cyrillique, arabe, grec, chinois, japonais...). Il s'agit pour l'essentiel de documents publiés par les bureaux statistiques des pays concernés, mais sont également conservés des documents d'autres provenances (chambres de commerce, travaux universitaires, etc.).

* <https://bibliotheque.insee.net>.

** Service statistique rattaché successivement à différentes administrations françaises de 1840 à 1940.

*** <https://bibliotheque.insee.net/Default/alain-desrosieres.aspx>.

Voir aussi : <https://www.youtube.com/watch?v=1og8cTR8wEU>.

**** <https://www.bnsp.insee.fr/bnsp>.

***** <https://www.ctles.fr/>.

parcours de ces publications est aussi l'occasion de souligner certaines similitudes des traitements de données réalisés par les bibliothécaires, ou au contraire certaines spécificités, comparés à ceux opérés par les statisticiens.

► Des publications statistiques étrangères : pour quoi faire ?

De 1840 à 1940, la Statistique générale de la France (SGF), un des ancêtres de l'Insee, a investi une quantité non négligeable d'énergie dans la constitution d'une bibliothèque statistique très fournie. Le fonds ancien, clos peu après la création de l'Insee, entre 1946 et 1951, représente aujourd'hui plus d'un kilomètre linéaire de documents, dont la majorité sont des publications étrangères. En 1913, 2 des 72 membres du personnel de la SGF travaillaient à plein temps pour la bibliothèque (Statistique générale de la France, 1913). Cet investissement témoigne de l'intérêt que les statisticiens publics portaient à leur bibliothèque, et aux publications statistiques étrangères en particulier. Mais à quoi leur servaient donc ces publications ?

Créer les statistiques internationales

“ Une première utilité de ces publications est de permettre la comparaison de chiffres et de servir la coordination statistique entre différents pays.”

Une première utilité de ces publications est de permettre la comparaison de chiffres et de servir la coordination statistique entre différents pays. Cette coordination trouve ses racines dans les Congrès internationaux de statistique, dont le premier a lieu en 1853 à Bruxelles, et qui sont organisés régulièrement jusqu'en 1878. À partir de 1885, ce sont les sessions du nouvel **Institut international de statistique**⁵ (*International Statistical Institute*, ISI) qui permettent la réunion des statisticiens publics de différents pays. Il faut attendre la création en 1919

de la Société des Nations, précurseur de l'Organisation des Nations unies (ONU), et de l'**Organisation internationale du travail**⁶ pour qu'existent des institutions ayant mandat de publier des statistiques internationales harmonisées (à la manière aujourd'hui de l'ONU ou d'Eurostat, l'autorité statistique de l'Union européenne⁷).

Avant l'émergence de ces institutions, la SGF est missionnée à plusieurs reprises par les statisticiens congressistes pour réaliser des publications de statistique internationale comparative : en 1873 sur l'agriculture, puis en 1905 et en 1910 sur ce que l'on appelle à l'époque le mouvement de la population, c'est-à-dire une composante de la démographie⁸. Sur cette période, la coordination entre bureaux statistiques de différents pays est assurée directement par l'un d'entre eux, mandaté par les autres au cours des

⁵ Site web de l'ISI : <https://isi-web.org/>.

⁶ Site web de l'OIT (*International Labour Organization*, ILO) : <https://www.ilo.org/fr>.

⁷ Voir l'article de Bischoff et Piffeteau sur l'histoire d'Eurostat dans ce même numéro.

⁸ La statistique de la population puise à l'époque à deux sources : le recensement de la population, quinquennal et exhaustif, et le mouvement de la population, calculé annuellement au début du XX^e siècle à partir des naissances et décès enregistrés à l'état civil.

sessions des Congrès ou de l’Institut international de statistique. Or, pour réaliser ces travaux ambitieux de comparaison internationale, il est nécessaire de disposer d’une solide documentation statistique ! Les ouvrages étrangers de la bibliothèque servent naturellement de base à ces travaux, en complément desquels sont mobilisés des rapports des directeurs de bureaux étrangers, sollicités pour l’occasion.

L’essor des travaux de conjoncture économique au début du XX^e siècle s’appuie également sur les publications statistiques étrangères. Ainsi, le *Bulletin de la Statistique générale de la France*⁹, publié trimestriellement à partir de 1911, inclut régulièrement des analyses tirant parti des dernières statistiques économiques parues dans d’autres pays – soit dans la rubrique des « statistiques générales », soit dans le cadre d’études plus approfondies.

Diffuser et inspirer les méthodes statistiques

Une seconde utilité des publications étrangères, qu’il s’agisse de publications officielles ou de revues scientifiques, est de servir de support à l’inspiration et à l’harmonisation internationale des méthodes statistiques. C’est en effet grâce à elles que les statisticiens s’informent des choix méthodologiques retenus à l’étranger, qui font ensuite l’objet de débats passionnés lors des rencontres internationales (Brian, 1989). Les publications de statistique étrangère servent ici d’opérateur de transfert international des techniques ou méthodes.

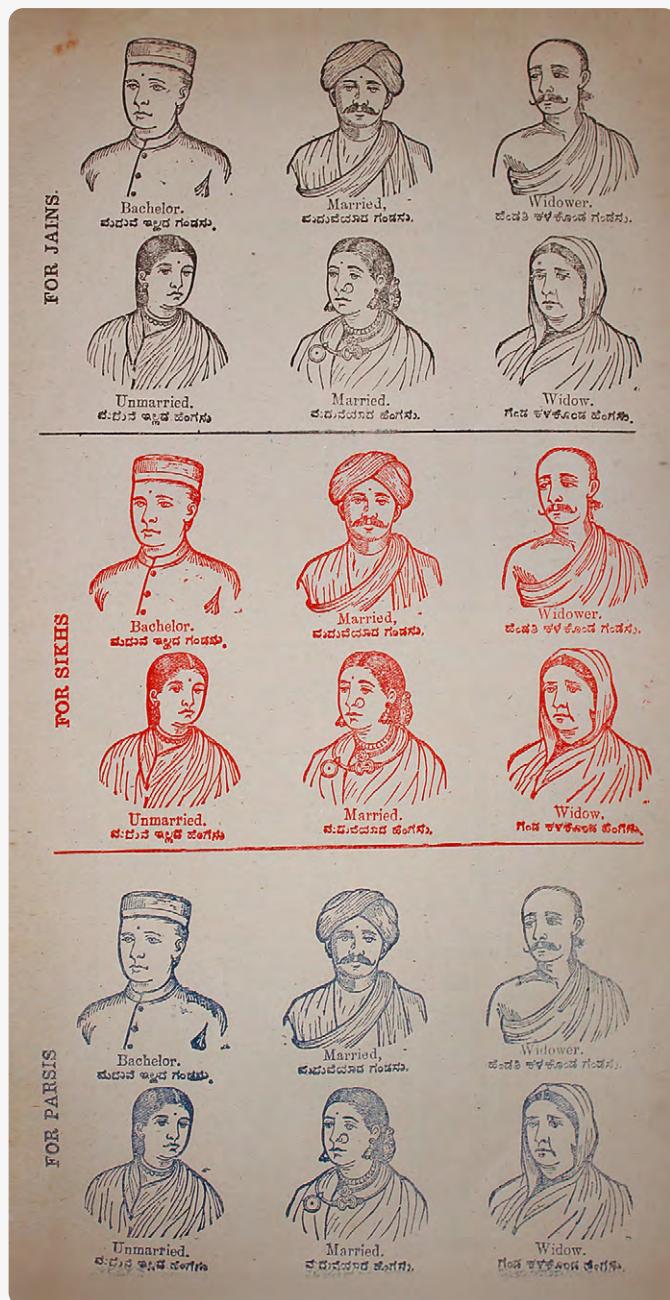
Par exemple, pour le recensement de la population bavaroise de 1871, le statisticien Georg von Mayr invente une technique visant à faciliter, au moment du dépouillement des bulletins, le travail manuel de comptage de la population par grande caractéristique (Von Oertzen, 2017). Appelée « *Zählblättchen[system]* » en allemand, puis « *slip system* » en anglais, cette technique repose sur l’utilisation de formes et de couleurs différentes pour les bulletins selon les caractéristiques des enquêtés. Elle est présentée en 1883, de manière résumée, dans un article du *Journal of the Statistical Society of London* (Hooper, 1883). Pour le recensement des Indes britanniques de 1901, le statisticien britannique Riesley décide de s’en inspirer. Il demande alors à Georg von Mayr un rapport plus détaillé sur ce qui a été fait pour le recensement en Bavière. Les statisticiens de l’*Indian civil service*¹⁰ mettent ensuite en place la méthode en faisant varier la couleur des bulletins selon la religion, et leur forme selon le sexe et le statut marital (*figures 1 et 2*).

La circulation des publications permet donc celle des méthodes statistiques. En mettant à disposition ces supports de communication que sont les documents papier, les bibliothèques sont alors au cœur des discussions internationales de méthodologie statistique.

⁹ <https://gallica.bnf.fr/ark:/12148/cb34437393k/date>.

¹⁰ Haute fonction publique de l’Empire britannique en Inde entre 1858 et 1947.

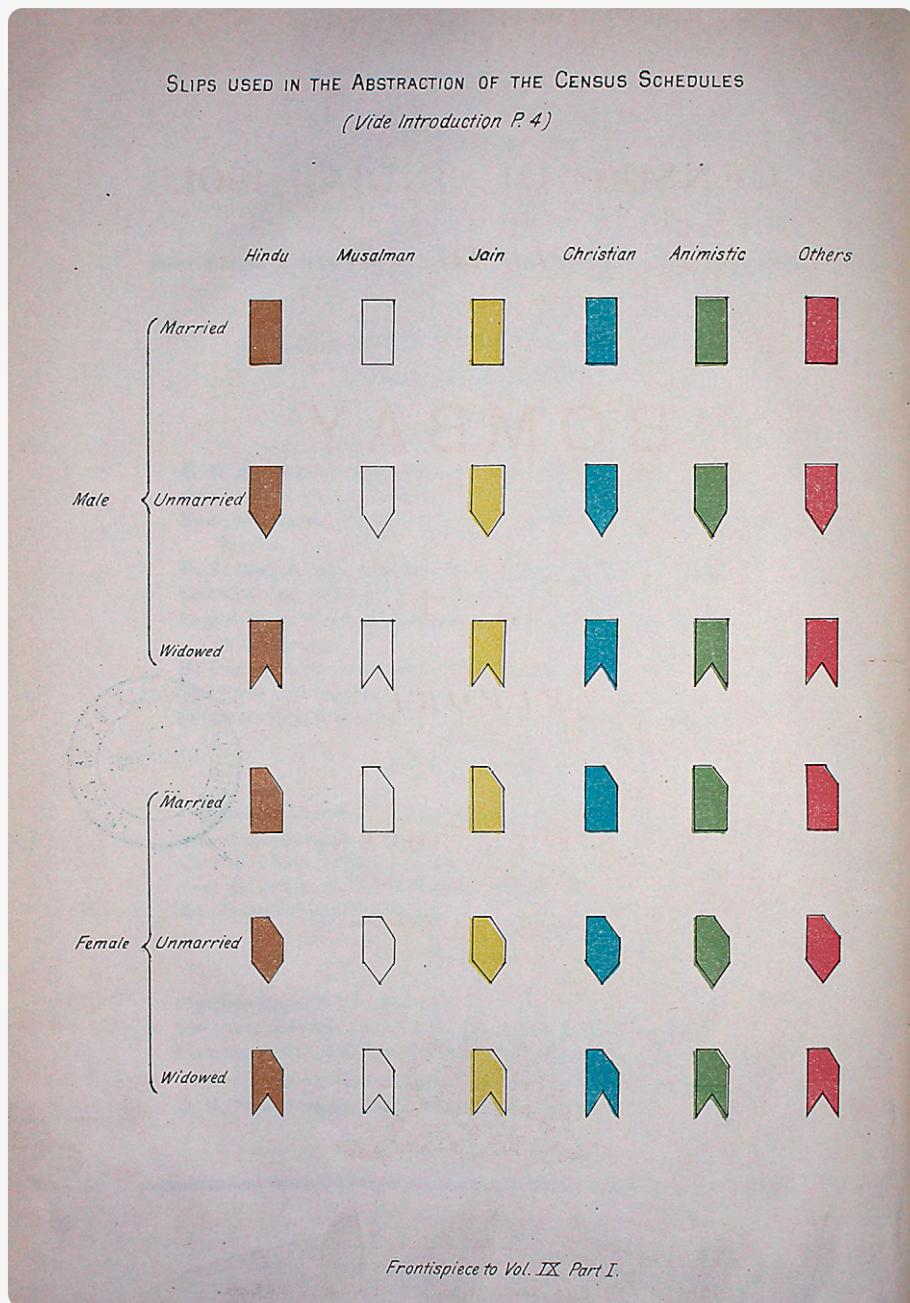
► Figure 1 - Slip system* à Mysore pour le recensement des Indes britanniques de 1901



* Technique visant à faciliter le travail manuel de dépouillement des bulletins individuels du recensement de la population, en vue d'obtenir rapidement un comptage par grande caractéristique de la population.
 Lecture : Dans l'Etat du Mysore, un portrait était reproduit sur chaque bulletin du recensement, dont le dessin variait selon le statut marital et le sexe de la personne enquêtée et la couleur selon sa religion.

Source : India. Census Commissioner (1903).

► Figure 2 - Slip system* à Bombay pour le recensement des Indes britanniques de 1901



* Technique visant à faciliter le travail manuel de dépouillement des bulletins individuels du recensement de la population, en vue d'obtenir rapidement un comptage par grande caractéristique de la population.
Lecture : Dans l'Etat de Bombay, la forme de la fiche associée au bulletin de recensement variait selon le sexe et le statut marital de la personne enquêtée et la couleur variait selon sa religion.
Source : India. Census Commissioner (1902).

Participer au rayonnement de la statistique publique

Que ce soit pour construire des statistiques internationales comparées ou pour discuter de l'harmonisation internationale des méthodes, les collections étrangères de la bibliothèque de la SGF ont rencontré leur public. Dans les années 1920 et au début des années 1930, 1 500 lecteurs extérieurs se rendent chaque année, en moyenne, à la bibliothèque. La suppression du poste de bibliothécaire en 1934 entraîne d'ailleurs de vives protestations¹¹.

Les usages des publications statistiques étrangères sont donc bien établis. Mais ils supposent une infrastructure robuste d'acquisition de ces documents. Comment étaient-ils récupérés par la bibliothèque de la SGF ?

► Échanger entre pays des publications statistiques

**Dès le XIX^e siècle,
les bureaux statistiques
échangent une
masse considérable
de données, sous la
forme de publications
statistiques.**

Dès le XIX^e siècle, les bureaux statistiques échangent une masse considérable de données, sous la forme de publications statistiques, qui alimentent alors leurs bibliothèques respectives. L'échange de publications entre bibliothèques existait déjà auparavant, mais il prend désormais davantage d'ampleur (Lilja, 2006). Un ouvrage de la SGF (Statistique générale de la France, 1913) en donne une illustration :

« [L]a bibliothèque de la Statistique générale contenait [au 1^{er} janvier 1913] 18 000 volumes [...].

Elle s'accroît chaque année d'environ 1 500 volumes, la plupart transmis en échange des publications françaises, par les offices de statistique étrangers ».

Le dernier registre d'entrée existant de la bibliothèque, pour la période 1937-1951, témoigne de la diversité des arrivages (*figure 3*). Le 11 janvier 1938 arrivent seize publications statistiques de pays différents : les statistiques d'assistance publique des communes finlandaises (Helsinki), les statistiques baleinières internationales (Oslo), la situation des affaires en Argentine (Buenos Aires), une étude sur les compensations salariales (New York), les statistiques britanniques de naufrages et décès en mer (Londres), et la liste continue...

Au milieu du XX^e siècle, la bibliothèque de la SGF est donc insérée dans un réseau international dense en matière d'échange des publications. Comment ce réseau s'est-il mis en place, et comment se sont organisés ces échanges ?

Un sujet de discussion récurrent entre statisticiens internationaux

Les comptes rendus des Congrès internationaux de statistique, entre 1853 et 1878, et des sessions de l'Institut international de statistique, à partir de 1885, permettent de comprendre comment ces échanges se sont mis en place et organisés. En l'absence à

¹¹ Ces informations sont issues des rapports au conseil de la SGF, publiés au journal officiel de la République française entre 1908 et 1936.

► Figure 3 - Registre d'entrée de la bibliothèque de la Statistique générale de la France (début de l'année 1938)

1938					
12.952	7/1	Mr. III+359/	Rome	C. Otoloughi	Éléments de démographie économique, fait par le Comité de statistique économique (1938) adulte 10 livres
12.953	*	Mr. 124/	Paris	Albert Stalzen	L'Or et le diamant (1938) 25
12.954	11/1	Mr. 15+55	Helsinki	Mr. C. S.	Assistance publique des communes 1935
12.955	*	Mr. 56/	Oulu	Committee for Whaling Statistics	International Whaling Statistics 1936/37.
12.956	*	Mr. III+123	Buenos Aires	Ernest Tongue & Co	Situation des affaires en Argentine 1937 oct. no 26
12.957	*	Mr. 59/	New York	Nat'l Inst. of Conference Board	Profit-Sharing and other Supplementary Configuration Plans Governing Shipping Casualties and Deaths 1936.
12.958	*	Mr. 77/	London	Board of Trade	The Health of the School Child 1936.
12.959	*	Mr. 136/	*	Chief Medical Officer of the Board	Annual report of the Registrar-General 1936.
12.960	*	Mr. XXIII+153	Dublin	Dept. of Local Government Public Health	Report of the Director of Labour, Chief Inspector of Factories, Sheriff of Commerce 1936 937/1937.
12.961	*	Mr. 36/	Bruxelles	Dept. of Labour	Annuaire Statistique des Salariés 1936
12.962	*	Mr. XXIV+6	Algiers	pour général	Bulletin trimestriel 1937 n°3.
12.963	*	Mr. 48/	Stockholm	Bus. Affaires Sociales	Le Travail et l'Institut 1937 n°3.
12.964	*	Mr. 71/	Belgrade	Banque nationale du royaume	Censo industrial Distrito federal 1936.
12.965	*	Mr. III+158	Prague	Jur. d'Hygiène publique	1. Recens agricole mondial, Bull. 225 Résumé 225 résumé
12.966	*	Mr. 140/	Caracas	Direcc. de Estadística	Comptabiliste agricole: Recueil de Statist. 1933-34 et 1933-34. 1933-34.
12.967	*	Mr.	Rome	Jur. Inst. d'Agriculture	Commerce int'l de marchandises agricoles 1936
12.968	*	Mr. XXIV+165	*	"	Annuaire des statistiques du travail 1937
12.969	*	Mr. XX+169/	*	"	Recueil des traités 1937. vol. CLXXV - CLXXXI
12.970	12/1	Mr. III+239/	Gênes	M. I. T.	Balances des paiements 1936
12.971	*	Mr.	*	S. D. N.	Compte de poste pour la réception de la fabrication de l'agriculture et statistiques du commerce international 1936.
12.972	*	Mr. 245/	*	*	Recueil annuel de la situation commerciale et industrielle 1936.
12.973	*	Mr. 3/	*	*	Budget ex. 1938
12.974	*	Mr. 384/	*	*	La situation commerciale et industrielle 1936.
12.975	13/1	Mr. 179/	Marseille	Ch. de CCE administrative	Rapport P.T.T. 1935.
12.976	*	Mr.	Paris	Min. Finance	La production industrielle en France
12.977	*	Mr. 156/	La Haye	Ch. de CG du Hébre	Compte définitif des dépenses ex. 1934.
12.978	*	Mr. 166/	Paris	Min. P.T.T.	203 Studies in Personnel Policy. Multiple-Shift Operation 1937/41
12.979	*	Mr. 180 à 187/	*	A Saumur (B.I.G.F. oct/dez 1937)	Stat. Accidents de Roulage et de Circulation 1936
12.980	*	Mr.	*	Min. Travail	- Véhicules à moteurs 1937
12.981	17/1	Mr. 1/	New York	National Industrial Conference Board	Sec. d'assurances, 1936.
12.982	*	Mr. 52/	Edmonton	M. C. S.	Women in the U.S. 1937
12.983	*	Mr. XX+30/	*	*	The Exporters 1936.
12.984	*	Mr. IV+109/	Oslo	*	Indian Customs Tariff 14 th Issue 1938
12.985	*	Mr. 139/	Washington	Women's Bureau	Annual report 1936.
12.986	*	Mr. 21/	Natalia	Dept. of Economic Affairs	Addendum 1936 to the Index Numbers of Indian Prices 1930-31
12.987	18/1	Mr. 14/	Bellu	Indian Trade Fair	Die Haushaltsgenossenschaften Volks-Bauh. u. Betriebsabteilung (several issues)
12.988	*	Mr. 32/	*	Chief Inspector of Mines	Enquête 1933 Part XXIV - XXV 1933-1935 # I Radi 474
12.989	*	Mr. 22/	Calcutta	Dept. of Col. Intelligence & Stats.	Annuaire Statistique 1937
12.990	*	Mr. 23/	Madras	Statist. Reichsamt	Assistance publique 1935
12.991	*	Mr.	Canberra	Min. of Com. & Stat.	Imposition 1937
12.992	*	Mr. XX+394/	Helsingfors	M. C. S.	Recueil Stat. 1939 - XII
12.993	*	Mr. II+92/	Stockholm	*	
12.994	*	Mr. 21/	*	*	
12.995	*	Mr. XXIV+303/	*	*	

l'époque d'institutions internationales assurant la coordination statistique, ces rencontres permettent aux statisticiens occidentaux¹² d'échanger sur les différents aspects de leurs travaux (Randeraad, 2020). Dans la deuxième moitié du XIX^e siècle, le sujet des échanges de publications statistiques y constitue un serpent de mer, régulièrement discuté par les directeurs de bureaux statistiques : comment organiser rationnellement cet échange ?

“ **Il faut attendre 1895 pour que l'Institut international de statistique crée un « comité pour l'échange régulier des publications statistiques entre les différents pays ».**”

entre les différents pays ». L'existence de ce comité est mentionnée dans les comptes rendus des rencontres internationales jusqu'en 1903.

À chacune de ces discussions, les statisticiens butent sur des obstacles similaires, de nature économique et géopolitique. En effet, envoyer mensuellement des centaines d'ouvrages par la poste coûte très cher pour des administrations aux budgets et effectifs limités. Pour les pays n'ayant pas de frontière commune, ce coût est par ailleurs renchéri par l'absence de franchise postale de transit. Par exemple, si le bureau statistique suédois veut envoyer des publications à son homologue français, les ouvrages transiteront probablement par le Danemark et l'Allemagne. Ce ne sont donc pas moins de quatre services postaux qui doivent être rémunérés.

Certains pays acceptent d'accorder une franchise postale aux ouvrages de statistique, en plus d'échanges à titre gracieux, mais il n'y a pas d'harmonisation absolue en la matière. Ainsi, au début du XX^e siècle, alors que l'emprise géopolitique de la guerre franco-prussienne de 1870-1871 reste vive, certains statisticiens déplorent que le bureau allemand refuse d'envoyer gratuitement ses publications, qui sont par ailleurs vendues en librairie.

L'organisation matérielle des échanges de publications

Comment ces échanges ont-ils été matériellement organisés ? Il est difficile de répondre à cette question, faute de conservation des archives de la bibliothèque. Des traces indirectes des échanges peuvent cependant ponctuellement être retrouvées dans les ouvrages eux-mêmes.

¹² Les participants des Congrès puis des sessions de l'Institut proviennent initialement principalement des pays d'Europe de l'Ouest, rejoints ensuite par les États-Unis et l'Europe de l'Est. À la fin du XIX^e siècle, des représentants de pays non occidentaux (Amérique du Sud, Turquie, Égypte...) assistent plus régulièrement aux séances.

¹³ C'est le philanthrope français Alexandre Vattemare (1796-1864) qui y prend la parole pour plaider la mise en place d'un « système régulier d'échanges entre tous les États, des documents statistiques publiés par les gouvernements et les particuliers ». Même s'il n'est pas statisticien (mais... ventiloque !), il travaille alors depuis vingt-cinq ans à l'édification d'un système d'échange des publications officielles (Richards, 1944), reposant sur une grande bibliothèque générale rassemblant l'ensemble des documents, puis les redistribuant vers des sociétés locales.

Certaines acquisitions ont été occasionnelles : envois divers spontanés¹⁴, réceptions d'ouvrages lors de Congrès internationaux ou d'expositions universelles, ou encore envois *ad personam*¹⁵.

D'autres échanges ont été organisés de façon plus systématique, sur la base de listes de diffusion et d'attestations de réception. On retrouve ainsi ponctuellement des cartes préremplies attestant réception, n'attendant qu'un coup de tampon pour être renvoyées à l'expéditeur de la publication. Ces cartes renvoyées permettent en retour d'alimenter des listes de diffusion, dont le défaut met à mal l'organisation de ces échanges. En témoigne la désorganisation qu'a connue la bibliothèque statistique hongroise après un incendie lors du siège de Budapest de 1944-1945, comme l'évoque une lettre de l'office hongrois :

« [J]e m'empresserai de vous faire expédier les numéros de nos publications parus depuis janvier 1940 dès que j'aurai votre réponse me signalant les derniers tomes ou livraisons de nos publications faisant l'objet de l'échange qui sont en votre possession, parce que la liste de distributions fut consumée par les flammes de l'incendie qui s'est déclaré dans notre bibliothèque pendant le Siège de Budapest ».

Une fois les documents acquis grâce aux échanges, il reste encore à savoir comment les trier au sein des rayonnages de la bibliothèque pour permettre leur accès au public.

Une fois les documents acquis grâce aux échanges, il reste encore à savoir comment les trier au sein des rayonnages de la bibliothèque pour permettre leur accès au public. Si l'échange de données informatiques donne aujourd'hui lieu à de nombreuses discussions sur la normalisation des métadonnées¹⁶ (Dondon et Lamarche, 2023), ces dernières jouent depuis longtemps un rôle central en bibliothèque.

► Classer des données : un rôle central en bibliothèque depuis la plus haute Antiquité

Il ne suffit pas d'accumuler les documents, encore faut-il pourvoir les retrouver. La tâche n'est pas toujours simple pour un particulier disposant d'une belle collection d'ouvrages imprimés. Lorsqu'il s'agit de plusieurs kilomètres de rayonnages, elle devient impossible sans tout un ensemble de règles et de techniques de tri, de rangement et de signalement. Apparues très tôt dans l'histoire de l'humanité, ces dernières s'adaptent et se perfectionnent en permanence avec l'évolution des supports, des mentalités, des besoins et des technologies (Barbier, 2021 ; Viry, 2013 ; Salvan, 1962).

Dès la plus haute Antiquité, les premières grandes bibliothèques apparaissent dans les temples et les palais, réunissant des milliers de tablettes d'argile en Mésopotamie et de rouleaux de papyrus en Égypte. Autant de documents à stocker et à retrouver. À cette

¹⁴ Comme ce quotidien nationaliste hongrois qui adresse en 1930 au directeur de la SGF un exemplaire de l'ouvrage « Justice pour la Hongrie ! » (dont la couverture représente la Hongrie à la fois démembrée et crucifiée par le traité de Trianon de 1920).

¹⁵ Comme cet ouvrage de la statistique du mouvement de la population espagnole de 1863, dédicacé par le vice-président du bureau statistique pour un envoi personnel au « docteur Bertillon ».

¹⁶ Une métad donnée est une donnée qui définit et décrit une autre donnée ([encadré 2](#)).

époque, différents moyens sont utilisés pour faciliter leur accès : regroupement par genre ou par sujet (littérature, religion, science, histoire, géographie, droit), étiquettes et colophons¹⁷, parfois catalogues et numérotations. Les métadonnées existent donc de longue date en bibliothèque (**encadré 2**).

Le besoin de classer le monde est un aspect essentiel de l'esprit humain. Chez les peuples premiers, une pensée classificatoire extrêmement développée appréhende le monde en ses moindres détails, et lui confère du sens tout en reflétant les structures sociales. Avec le développement de la pensée rationnelle et de l'écriture, les philosophes de l'Antiquité classique entreprennent des classifications du savoir : les classifications du monde laissent place aux classifications de la connaissance du monde. Parallèlement, la bibliothèque d'Alexandrie, de par son ambition et son ampleur, a dû perfectionner et développer les techniques de gestion documentaire. Elle marque une avancée notable dans le signalement des collections. Dès lors, le classement des ouvrages, physiquement dans les rayonnages comme intellectuellement dans les catalogues, évolue en regard des classifications de la connaissance, mais aussi des usages auxquels ils répondent.

► Encadré 2. Qu'est-ce qu'une métadonnée en bibliothèque ?

Selon l'Organisation internationale de normalisation (*International Organization for Standardization*, ISO), une métadonnée est une « donnée qui définit et décrit une autre donnée » (norme ISO/IEC 11179-3:2023). L'organisation américaine *National Information Standards Organization* (NISO) la définit quant à elle comme « de l'information structurée qui décrit, explique, localise la ressource et en facilite la recherche, l'usage et la gestion » (Riley, 2017).

Les métadonnées peuvent être de plusieurs types. Dans le cas des bibliothèques, elles peuvent être notamment :

- descriptives (pour identifier un document) : titre, auteur, date, éditeur, etc. ;
- techniques : format, taille ou poids, date de création, compression, etc. ;
- structurelles (sur l'organisation d'un document) : pour permettre par exemple de lier les différents fichiers qui le composent (textes, images, etc.) afin de le reconstituer ;
- administratives : localisation du document, gestion des droits d'accès et d'usage, conditions de conservation, etc.

On remarque que, dans la mesure où elles permettent d'identifier, de décrire et de

comprendre le contenu d'une ressource, les métadonnées descriptives et les données bibliographiques se confondent. La Bibliothèque nationale de France (BnF) indique d'ailleurs que « les métadonnées ne décrivent pas nécessairement des documents électroniques ». Notons encore qu'une métadonnée peut être externe à la ressource qu'elle décrit (c'est le cas d'une notice de catalogue) ou interne (par exemple les balises XML d'un document électronique). En passant des catalogues papier aux catalogues informatiques, la mise en abyme se poursuit : la notice bibliographique comporte à son tour des métadonnées ou, faudrait-il dire, des méta-métadonnées. Il en va de même pour les bases de données numériques, qui peuvent être décrites par une notice dans un catalogue informatique, laquelle s'accompagne de son lot de métadonnées.

Précisons enfin que, dans le domaine documentaire, les normes (normes Afnor* de catalogage, listes d'autorités** pour les noms propres ou les mots-matières***) et les formats (MARC**** pour les grilles de saisie des informations bibliographiques), qui servent à renseigner les métadonnées, ne font pas partie des métadonnées, contrairement à la définition des métadonnées statistiques (Bonnans, 2019).

* Association française de normalisation.

** Une liste d'autorités est un vocabulaire contrôlé, contenant des termes autorisés (et des termes interdits) pour identifier de manière harmonisée, et donc sans ambiguïté, des organisations, lieux géographiques, œuvres, objets ou concepts.

*** Mot ou groupe de mots qui caractérise le contenu d'un document (sujet traité) et en facilite le repérage.

**** Acronyme de MACHine-Readable Cataloging (catalogage lisible par une machine) : format d'échange informatique de données bibliographiques.

¹⁷ Note ajoutée à la fin d'un manuscrit par le copiste, indiquant son nom, celui de son commanditaire, la date et le lieu de la copie, le nom de l'auteur, le titre de l'œuvre, etc.

Au XVIII^e siècle, la bibliographie acquiert une dimension épistémologique et un statut scientifique

À partir du XV^e siècle, avec le développement de l'imprimerie, qui concourt à la diffusion des idées, et le début de la révolution scientifique (travaux de Copernic au XVI^e siècle, de Newton au XVII^e siècle...), les systèmes philosophiques d'organisation du savoir se reconfigurent. Parallèlement, parfois de concert, les systèmes de classification documentaire en font autant.

Au tournant des XVII^e et XVIII^e siècles, le contexte intellectuel est marqué par l'apparition des publications scientifiques et le développement des grandes bibliothèques publiques. En France et dans plusieurs pays du Nord-Ouest de l'Europe, les catalogues imprimés des bibliothèques connaissent alors un essor remarquable (Neveu, 2010, 2013, 2014 ; Blom, 2020).

La réflexion sur les relations entre, d'une part, la classification des livres et, d'autre part, celle des savoirs – c'est-à-dire entre bibliothéconomie¹⁸ et épistémologie¹⁹ – se poursuit au XVIII^e siècle, à l'époque des Lumières, chez les philosophes et chez les bibliographes. À leurs yeux, la diversité des systèmes de classement des bibliothèques – nécessairement imparfaits – témoigne de leur caractère arbitraire. Les aspects pratiques auxquels les classifications des catalogues se soumettent leur semblent contradictoires avec l'ambition scientifique. Ils restent convaincus qu'il est possible de parvenir à une classification satisfaisante, c'est-à-dire scientifique.

Les problèmes pratiques et théoriques posés par les catalogues donnent lieu à des réflexions qui sont parfois placées en tête de ces derniers et qui témoignent d'une interrogation épistémologique beaucoup plus fondamentale : à travers la question pratique du rangement des livres et celle théorique de l'accès au savoir se pose aussi la question de la relation entre le parcours physique dans la bibliothèque et le parcours intellectuel dans la connaissance. La pertinence d'une classification des connaissances permet d'organiser les livres dans un ordre adéquat afin de favoriser l'acquisition du savoir et, inversement, la classification des ouvrages permet de comprendre la généalogie des sciences.

**Au XVIII^e siècle,
la bibliographie acquiert
un statut scientifique
et se voit reconnaître
comme une discipline
auxiliaire des sciences.**

Au XVIII^e siècle, la notion de bibliographie recouvre non seulement la classification documentaire, mais aussi, dans un sens très contemporain, la connaissance des œuvres et de leurs éditions. Elle acquiert un statut scientifique et se voit reconnaître comme une discipline auxiliaire des sciences. Pour cela, elle a dû adopter les principes mis en œuvre dans la science de l'époque, c'est-à-dire faire preuve d'une rigueur dans ses méthodes, mais aussi inclure une dimension spéculative qui en

théorise les principes. Avec son ordre méthodique, à plusieurs niveaux d'arborescence, selon des critères logiques abstraits, le catalogue devient une œuvre scientifique. Cette idée selon laquelle la bibliographie constitue une science qui sert d'introduction à la

¹⁸ Science et technique de l'organisation et de la gestion d'une bibliothèque (source : Robert).

¹⁹ Discipline qui prend la connaissance scientifique pour objet (source : Larousse).

connaissance triomphe à la fin du XIX^e siècle lorsque Melvil Dewey la place en tête de sa classification (voir infra), puis au début du XX^e siècle avec la reconnaissance des sciences de l'information.

Au XIX^e siècle, les classifications documentaires s'inspirent des taxonomies du vivant

La réflexion théorique ne se limite pas aux spéculations sur la classification générale de la connaissance et des sciences. Elle porte aussi de manière très concrète sur un type particulier de classification apparu au XVIII^e siècle et initié par le naturaliste suédois Carl von Linné (1707-1778), relatif cette fois aux êtres vivants : les taxonomies, que la rigueur et la valeur heuristique érigent en exemple. La révolution provoquée par Linné met cependant du temps à marquer les classifications documentaires. En effet, elle repose sur trois règles incompatibles avec les besoins des bibliothèques (Parrochia, 2017) :

1. Tout item doit être classé. Ce principe d'exhaustivité ne pose pas de difficulté. En revanche, il suppose un catalogue complet et achevé dans sa structure. Or, le principe même des classifications documentaires prévoit une extension toujours possible – théoriquement infinie – afin de suivre l'évolution incessante de la connaissance.

2. Il ne peut pas y avoir de classe vide. Or, la classification documentaire présente des classes vides, soit parce que l'évolution du savoir conduit à des suppressions, soit parce que l'apparition de connaissances nouvelles est envisagée.

3. Aucun item ne peut être classé dans plus d'une classe. Ce point est fondamental dans les sciences de la nature, car son non-respect remettrait en cause les critères de la classification, donc sa validité. Or, le contenu d'un ouvrage est souvent complexe, de sorte qu'il est impossible de lui assigner un seul descriptif sans limiter considérablement les chances de le trouver. Il est donc non seulement possible, mais fréquemment nécessaire, de lui associer plusieurs descripteurs ou indices, voire, lorsqu'on dispose de plusieurs exemplaires, de les ranger à des endroits différents. L'enjeu est que tous les lecteurs potentiels puissent facilement trouver l'ouvrage à partir de leurs points de vue particuliers²⁰.

C'est au XIX^e siècle, aux États-Unis, que s'opère la synthèse entre les classifications documentaires et la taxonomie (Eaton, 1959). Jusque-là, la réflexion est avant tout marquée par les débats philosophiques. Puis, la recherche scientifique et la production imprimée connaissent une croissance nouvelle. Parallèlement, le besoin de connaissance et d'information scientifiques s'accroît rapidement avec le développement industriel et social. Les opérations de stockage et de signalement des documents prennent une autre dimension. Le XIX^e siècle est d'abord une période d'expérimentation pour les classifications documentaires : en l'absence de modèle standard, chaque responsable de bibliothèque construit son propre système selon ses propres besoins.

²⁰ Tout comme les classifications documentaires, les nomenclatures statistiques comportent des redondances. Toutefois, là où la pratique bibliothéconomique permet de placer le même document en deux ou trois endroits différents afin de s'assurer qu'il sera plus facilement trouvé par le lecteur, le respect de la règle taxonomique de l'univocité impose le choix d'un emplacement unique dans la nomenclature. La préface de la nomenclature des activités économiques de l'Insee en 1949 témoigne de ce dilemme. Ce choix n'est jamais neutre ou objectif : il répond à un besoin, sociologiquement déterminé. La finalité interfère avec l'exigence scientifique de description objective (Guibert et al., 1971 ; Desrosières et Thévenot, 1979).

Les États-Unis présentent deux particularités qui vont permettre d'aboutir aux grandes classifications documentaires modernes. D'une part, dans ce pays jeune, la proportion d'ouvrages récents est plus importante qu'en Europe, où les collections héritées de l'Ancien Régime s'insèrent bien dans les classifications traditionnelles. D'autre part, le libre accès y est très répandu, ce qui invite à porter sur le rangement des documents un regard plus ouvert aux besoins et aux pratiques des lecteurs. La question reste de placer ensemble les ouvrages relatifs à un même sujet. Il importe aussi que l'on puisse passer d'un sujet à l'autre selon une progression logique, reposant sur l'organisation du savoir, laquelle reflète l'organisation du monde.

En créant une indexation thématique selon un code décimal, le bibliothécaire américain Melvil Dewey (1851-1931) conçoit un système qui permet de placer (et retrouver) les documents dans le même ordre, à la fois dans les catalogues et dans les rayons, dans n'importe quelle bibliothèque (Béthery, 2013). Sa classification (Classification décimale de Dewey ou CDD), publiée en 1876 et régulièrement actualisée depuis, est claire, objective, hiérarchisée, souple, autant de qualités qui contribuent à son succès jusqu'à nos jours. Elle permet notamment de classer les publications statistiques (**encadré 3**).

Dewey élabora 12 éditions successives de sa classification avant son décès en 1931. La version actuellement en vigueur est la 23^e, parue en 2011. La CDD se montre particulièrement bien adaptée au libre accès et à la lecture publique. En 1905, Henri La Fontaine et Paul Otlet publient, avec l'accord de Dewey, leur Classification décimale universelle (CDU). Celle-ci est inspirée de la CDD mais mieux adaptée aux bibliothèques d'étude et de recherche. En France, toutefois, l'accès du public aux collections reste indirect, aussi bien dans les bibliothèques municipales qu'universitaires. De ce fait, le catalogue imprimé demeure la règle. Si les fichiers sont connus pour le travail interne depuis le XVIII^e siècle au moins, ils ne sont proposés au public qu'à partir du début du XX^e siècle (Neveu, 2015 ; Sarrazin, 2015). Dans les bibliothèques ministérielles, comme celle de la SGF (dont nous étudierons le plan de classement spécifique dans la partie suivante), les solutions de classification sont quant à elles déterminées par la nature des fonds et l'usage qui en est fait.

► Encadré 3. Classer des ouvrages statistiques avec la CDD

La Classification décimale de Dewey (CDD), utilisée dans de nombreuses bibliothèques pour classer les documents, permet de distinguer trois sens de la « statistique » :

- Le premier renvoie à la statistique comme branche des mathématiques. Le code des documents est alors « 519.5 ».
- Le second renvoie à la statistique économique et sociale (celle effectuée par les organismes officiels de la statistique publique, mais aussi dans la recherche en sciences sociales). Le code des documents est alors « 310 », suivi d'une éventuelle subdivision géographique.

• Enfin, la CDD prévoit une subdivision de forme permettant d'indiquer, pour toute thématique de la CDD, les statistiques s'y rapportant. Il suffit alors d'ajouter « 021 » à l'indice du sujet concerné (en respectant les règles de construction), par exemple « 551.550 21 » pour les statistiques des ouragans, « 364.944 021 » pour celles des crimes et délits en France.

Ces différents sens sont parfois proches en pratique : pour le classement d'un ouvrage statistique donné, la CDD laisse donc une certaine marge d'appréciation à chaque bibliothèque se l'appropriant.

► Classer des données : le cas des bibliothèques de statistique

Revenons-en à présent au circuit des documents étrangers à la bibliothèque de la SGF. Une fois que les publications statistiques étrangères ont été acquises, elles sont inscrites au registre d'inventaire, tamponnées, cataloguées sur une fiche cartonnée. Puis, elles sont équipées d'une étiquette indiquant leur cote, c'est-à-dire une suite de signes correspondant à un lieu précis de rangement dans les rayonnages.

Pour réaliser cette dernière opération, il est nécessaire de disposer d'un plan de cotation, qui permet d'organiser spatialement les données. Ce plan de cotation revient à organiser le savoir statistique, de façon hiérarchisée et exhaustive. Rétrospectivement, il s'agit d'une fenêtre permettant d'appréhender la vision du monde des statisticiens, non pas sur l'économie et la société (comme certains ont pu l'étudier au prisme des nomenclatures statistiques : Guibert et al., 1971 ; Desrosières et Thévenot, 1988), mais sur leur propre activité statistique.

Dans l'état d'esprit positiviste²¹ de la fin du XIX^e siècle, des discussions sont organisées à l'échelle internationale entre statisticiens publics (aux Congrès et à l'Institut international de statistique), afin d'arrêter un plan de classement universel du savoir statistique. L'harmonisation souhaitée n'a cependant pas lieu. L'un des partisans les plus véhéments de la conception d'un plan de classement harmonisé, Troïnitsky (directeur du bureau de la statistique russe), présente en 1895 à l'Institut la solution retenue en Russie pour tenir compte de besoins nationaux. Tout en tâchant de respecter au mieux la proposition du groupe de travail international mandaté par l'Institut, des ajustements ont ainsi été réalisés pour mieux distinguer la « statistique du bétail » et celle de la « propriété foncière », afin de correspondre à la spécialisation agricole de la statistique officielle russe tsariste des *zemtsvas*²² (Blum et Mespoulet, 2003).

Le plan de classement à la bibliothèque de la Statistique Générale de la France

Quelle solution a alors été retenue à la bibliothèque de la SGF ?

S'inspirant des discussions internationales, le plan de classement adopté imbrique trois dimensions : une dimension géographique et politique, une dimension thématique et une dimension chronologique. Par exemple, un ouvrage peut avoir pour cote « Ap – 6 – 1928 ». Le premier bloc renvoie à l'entité géographique : la majuscule pour le pays (« A » pour Allemagne), la minuscule pour une éventuelle subdivision régionale ou municipale (« p » pour Prusse). Le deuxième bloc renvoie au plan de classement thématique (« 6 » pour commerce extérieur et intérieur). Enfin, le dernier bloc précise l'année concernée par la publication. Ainsi, l'ensemble des publications du commerce extérieur et intérieur prussien sont rangées d'un seul tenant, par ordre chronologique.

²¹ Le positivisme renvoie au système d'Auguste Comte, qui considère que toutes les activités philosophiques et scientifiques ne doivent s'effectuer que dans le seul cadre de l'analyse des faits réels vérifiés par l'expérience et que l'esprit humain peut formuler les lois et les rapports qui s'établissent entre les phénomènes et ne peut aller au-delà (source : Larousse).

²² Au singulier, *zemtso* : assemblée territoriale assurant l'administration locale dans les gouvernements de la Russie d'Europe (1864-1917) (source : Larousse).

Le classement thématique s'applique uniformément à chaque pays, permettant l'identification rapide de publications étrangères traitant du même sujet. Dix catégories principales sont distinguées, elles-mêmes subdivisées en plusieurs sous-classes gigognes :

- 0 – annuaires statistiques ;
- 1 – recensements de la population ;
- 2 – mouvement de la population (dont statistique sanitaire) ;
- 3 – éducation et justice ;
- 4 – travail ;
- 5 – économie productive (agriculture, industrie) ;
- 6 – commerce extérieur et intérieur (dont trafic ferroviaire, trafic fluvial ou maritime) ;
- 7 – successions, donations, finance et assurance ;
- 8 – budget de l'État, armée ;
- 9 – colonies.

Un plan qui témoigne des conceptions d'alors de la statistique

Le logement a tendance à être pensé au prisme du travail, c'est-à-dire de la question sociale ouvrière. L'éducation et la justice sont rangées ensemble, comme choses du « fait moral »...

Le classement de la bibliothèque de la SGF est intéressant rétrospectivement en ce qu'il révèle le fossé qui nous sépare aujourd'hui des conceptions de la statistique qui prévalaient parmi les statisticiens publics. Le recensement de la population est pensé de façon distincte du mouvement de la population, car ce dernier se fonde sur l'exploitation de l'état civil (et implique une autre administration, celle du ministère de l'Intérieur). En tant que sous-catégorie de la catégorie 4, le logement a tendance à être pensé au prisme du travail, c'est-à-dire de la question sociale ouvrière. L'éducation et la justice sont rangées ensemble, comme choses du « fait moral », à l'inverse des budgets généraux de l'État et des statistiques militaires.

La statistique sanitaire est subordonnée au mouvement de la population : la maladie est pensée comme une extension du domaine de la mortalité. La statistique coloniale est radicalement distinguée du reste de la nomenclature, reflétant la division administrative du travail statistique à ce sujet.

L'intérêt de ce plan de classement thématique réside aussi dans l'uniformité avec laquelle il est appliqué pour l'ensemble des pays : ces catégories permettent de comparer la place respective des statistiques de telle ou telle nature dans les collections aujourd'hui conservées.

L'usage du plan de classement de la SGF a pris fin lors de la création de l'Insee en 1946. Celle-ci a entraîné une réorganisation importante des collections de la bibliothèque. Alors qu'une nouvelle ère s'ouvre pour la statistique publique française, le plan de classement évolue en même temps que les mentalités et moyens des statisticiens. Le nouveau plan fait la part belle aux « études économiques » composant désormais le nom de l'institut. Les outils des bibliothécaires, comme ceux des statisticiens, sont ainsi amenés à évoluer. Nous allons voir dans la dernière partie ce qu'il reste aujourd'hui des outils d'hier.

► **Le fonds ancien étranger aujourd'hui : un témoignage précieux pour la recherche en sciences sociales**

Si les publications ont pu servir aux statisticiens français à comparer des données internationales ou à s'inspirer des méthodes étrangères, le public désormais concerné est plutôt celui des universitaires.

Le fonds ancien étranger de la bibliothèque, constitué de dizaines de milliers d'ouvrages échangés sur plus d'un siècle, a été clos en 1946 – au moment de la réorganisation de la bibliothèque et de la refonte du plan de classement. À quoi servent désormais ces documents ?

Si les publications ont pu servir aux statisticiens français à comparer des données internationales ou à s'inspirer des méthodes étrangères, le public désormais concerné est plutôt celui des universitaires. En tant que sources

documentaires, ces publications renseignent sur de nombreux aspects quantifiés de la société et de l'économie des nombreux pays du fonds, ainsi bien sûr que sur l'histoire des méthodes statistiques. Les disciplines potentiellement intéressées par les collections sont de ce fait nombreuses : histoire bien sûr, mais aussi démographie, économie, sociologie...

Du fait de la tradition d'échanges, aucun de ces documents n'est unique au monde : les conserver en France présente cependant deux intérêts majeurs. D'une part, conserver un accès large à la documentation sur place des chercheurs français (un enjeu croissant à l'heure de la décarbonation de la recherche). D'autre part, tirer parti de la présence des différentes collections à un même endroit pour mieux les comparer, ainsi que nous y invite le plan de classement de la bibliothèque.

La bibliothèque de l'Insee Alain Desrosières s'engage aujourd'hui résolument dans la valorisation de ce patrimoine, pour faire rayonner la statistique publique auprès des publics universitaires. Débuté en 2022, un chantier pluriannuel de tri, conservation et catalogage du fonds ancien étranger permet de signaler l'existence des collections –

désormais dans le catalogue interuniversitaire du Sudoc²³. Ces ouvrages sont consultés en salle de lecture. Différentes actions de valorisation sont aussi construites à destination du monde universitaire : séminaire d'histoire de la statistique (« stat'histoire publique »)²⁴, accueil de colloques, partenariats d'enseignement...

La conservation de ce patrimoine présente aussi un intérêt pour les pays étrangers concernés par ces collections. Si chaque pays conserve des exemplaires de ses publications statistiques, aucun lieu de conservation n'est à l'abri de sinistres : dégâts des eaux (comme celui de 2002 à Prague), incendies (comme celui de 1944-1945 à Budapest), etc. Conserver les publications en plusieurs lieux réduit ainsi le risque de perte définitive. Un cas particulier mérite d'être signalé : l'histoire de la statistique a accompagné celle de la colonisation française et certains pays anciennement colonisés n'ont pas pu conserver sur leur territoire d'exemplaire des publications statistiques produites en contexte colonial. Une campagne de numérisation a ainsi été entreprise en 2018, en partenariat avec l'institut national de statistique tunisien, pour mettre à disposition du public en ligne plus d'une centaine d'ouvrages statistiques de la période coloniale. De quelque nature qu'ils soient, les changements de régime politique peuvent être la cause de pertes patrimoniales. L'histoire des échanges continue de s'écrire au présent.

²³ Le Sudoc – Système Universitaire de Documentation – constitue le dispositif national de catalogage partagé des ressources documentaires à la disposition de l'ensemble des bibliothèques universitaires et de recherche : <https://abes.fr/reseau-sudoc/le-reseau/etablissements-membres/>.

²⁴ <https://bibliotheque.insee.net/Default/cycle-de-seminaires-statistique-publique.aspx>.

► Fondements juridiques

- Arrêté du 22 décembre 1981 relatif à l'enrichissement du vocabulaire de l'informatique.
In : *site du Bulletin des bibliothèques de France*. [en ligne]. [Consulté le 28 août 2025].
Disponible à l'adresse :
<https://bbf.enssib.fr/consulter/bbf-1982-06-0355-009>.

► Bibliographie

- ARDOUIN, Véronique et COLLIGNON, Laure, 2017. La bibliothèque de l'Insee : une institution unique en son genre. In : *Bulletin des bibliothèques de France*. [en ligne]. 12 juillet 2017. ENSSIB. N° 12, pp. 80-89. [Consulté le 25 juillet 2025]. Disponible à l'adresse : https://bbf.enssib.fr/matieres-a-penser/la-bibliothèque-de-l-insee_67644.
- BACH, Angélique, 2017. La bibliothèque de l'Institut national de la statistique et des études économiques (INSEE). In : *Histoire @ Politique*. [en ligne]. 1^{er} juin 2017. Centre d'histoire de Sciences Po. N° 32. [Consulté le 25 juillet 2025]. Disponible à l'adresse : <https://doi.org/10.4000/histoirepolitique.10416>.
- BARBIER, Frédéric, 2021. *Histoire des bibliothèques. D'Alexandrie aux bibliothèques virtuelles*. 2^e édition. Février 2021. Armand Colin. Mnemosya. ISBN 978-2-2006-3012-6.
- BÉTHERY, Annie, 2013. *Guide de la classification décimale de Dewey. Tables abrégées de la XXIII^e édition intégrale en langue anglaise*. Le Cercle de la librairie. Bibliothèques. ISBN 978-2-7654-1352-3.
- BLOM, Helwi, 2020. Philosophie ou commerce ? : L'évolution des systèmes de classement bibliographique dans les catalogues de bibliothèques privées publiés en France au XVIII^e siècle. In : *Les bibliothèques et l'économie des connaissances 1450-1850 : Colloque international 9-13 avril 2019 Sárospatak (Hongrie)*. [en ligne]. 17 juin 2020. Magyar Tudományos Akadémia Könyvtár és Információs Központ. Pp. 203-234. [Consulté le 25 juillet 2025]. Disponible à l'adresse : <https://doi.org/10.36820/SAROSPATAK.2020.11>.
- BLUM, Alain et MESPOULET, Martine, 2003. *L'anarchie bureaucratique. Statistique et pouvoir sous Staline*. 23 janvier 2003. La Découverte. L'espace de l'histoire. ISBN 978-2-7071-3903-0.
- BONNANS, Dominique, 2019. RMéS, le référentiel de métadonnées statistiques de l'Insee. In : *Courrier des statistiques*. [en ligne]. 27 juin 2019. Insee. N° N2, pp. 46-57. [Consulté le 25 juillet 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/information/4168396?sommaire=4168411>.
- BRIAN, Éric, 1989. Statistique administrative et internationalisme statistique pendant la seconde moitié du XIX^e siècle. In : *Histoire & mesure*. [en ligne]. Varia. Volume 4, n° 3-4., pp. 201-224. [Consulté le 25 juillet 2025]. Disponible à l'adresse : <https://doi.org/10.3406/hism.1989.1357>.
- DESROSIÈRES, Alain, 1993. *La politique des grands nombres. Histoire de la raison statistique*. La Découverte. Poche. ISBN 978-2-7071-2253-7.
- DESROSIÈRES, Alain et THÉVENOT, Laurent, 1979. Les mots et les chiffres : les nomenclatures socioprofessionnelles. In : *Économie et Statistique*. [en ligne]. Avril 1979. Insee. N° 110, pp. 49-65. [Consulté le 25 juillet 2025]. Disponible à l'adresse : https://www.persee.fr/doc/estat_0336-1454_1979_num_110_1_426.

- DESROSIÈRES, Alain et THÉVENOT, Laurent, 1988. *Les catégories socioprofessionnelles*. La Découverte. Repères. ISBN 978-2-7071-1758-8.
- DONDON, Alexis et LAMARCHE, Pierre, 2023. Quels formats pour quelles données ? In : *Courrier des statistiques*. [en ligne]. 30 juin 2023. Insee. N° N9, pp. 86-103. [Consulté le 25 juillet 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/information/7635827?sommaire=7635842>.
- EATON, Thelma, 1959. The Development of Classification in America. In : *The role of classification in the modern American library*. [en ligne]. Graduate School of Library Science. University of Illinois at Urbana-Champaign. Allerton Park Institute Proceedings, n° 6, pp. 8-30. [Consulté le 30 juillet 2025]. Disponible à l'adresse : <https://www.ideals.illinois.edu/items/1547>.
- FAYET-SCRIBE, Sylvie, 2000. *Histoire de la documentation en France. Culture, science et technologie de l'information, 1895-1937*. 13 décembre 2000. CNRS Éditions. CNRS Histoire. ISBN 978-2-271-05790-7.
- GARDEY, Delphine, 2008. *Écrire, calculer, classer. Comment une révolution de papier a transformé les sociétés contemporaines (1800-1940)*. 10 janvier 2008. La Découverte. Textes à l'appui / Anthropologie des sciences et des techniques. ISBN 978-2-707-15367-8.
- GROUDIEV, Stéphanie et DE SABOULIN, Michel, 2015. La bibliothèque et les archives de l'Insee. In : *Statistique et Société*. [en ligne]. Octobre 2015. Société Française de Statistique. Volume 3, n° 2, pp. 21-28. [Consulté le 27 août 2025]. Disponible à l'adresse : https://www.persee.fr/doc/staso_2269-0271_2015_num_3_2_961.
- GUIBERT, Bernard, LAGANIER, Jean et VOLLE, Michel, 1971. Essai sur les nomenclatures industrielles. In : *Économie et Statistique*. [en ligne]. Février 1971. Insee. N° 20, pp. 23-36. [Consulté le 27 août 2025]. Disponible à l'adresse : https://www.persee.fr/doc/estat_0336-1454_1971_num_20_1_6122.
- HEIDE, Lars, 2009. *Punched-Card Systems and the Early Information Explosion, 1880-1945*. [en ligne]. 27 avril 2009. Baltimore: Johns Hopkins University Press. [Consulté le 27 août 2025]. Disponible à l'adresse : <https://muse.jhu.edu/book/3454>.
- HOOPER, Wynnard, 1883. The Theory and Practice of Statistics. In : *Journal of the Statistical Society of London*. [en ligne]. Septembre 1883. Oxford University Press. Volume 46, n° 3, pp. 461-516. [Consulté le 27 août 2025]. Disponible à l'adresse : <https://doi.org/10.2307/2979303>.
- INDIA. CENSUS COMMISSIONER, 1902. Bombay. Part I. Report. In : *Census of India, 1901*. Volume IX. Bombay, Government Central Press.
- INDIA. CENSUS COMMISSIONER, 1903. Mysore. Part 1, Report. In : *Census of India, 1901*. Volume XXIV, Bangalore, Office of the Superintendent of Government.
- LEONELLI, Sabina et TEMPINI, Niccolò, 2020. *Data Journeys in the Sciences*. [en ligne]. 30 juin 2020. Springer. [Consulté le 27 août 2025]. Disponible à l'adresse : <https://link.springer.com/book/10.1007/978-3-030-37177-7>.

- LILJA, Eeva Johanna, 2006. History of the International Exchange of Publications. In : *Handbook on the International Exchange of Publications*. K.G. Saur. Pp. 49-68. ISBN 978-3-598-11752-7.
- NEVEU, Valérie, 2010. L'inscription de la classification bibliographique dans le champ des sciences (fin XVII^e-début XVIII^e s.). In : *Séminaire ALMA 2009-2010 : les raisons classificatoires. Séance du 4 novembre 2010*. [en ligne]. Novembre 2010. Centre de recherches historiques de l'Ouest (CERHIO-UMR CNRS 6258). Site d'Angers. [Consulté le 27 août 2025]. Disponible à l'adresse : https://www.academia.edu/72290824/Linscription_de_la_classification_bibliographique_dans_le_champ_des_sciences_fin_XVIIe_d%C3%A9but_XVIIIe_s.
- NEVEU, Valérie, 2013. Classer les livres selon le « Système figuré des connaissances humaines » : émergence et déclin des systèmes bibliographiques d'inspiration baconienne (1752-1812). In : *Recherches sur Diderot et sur l'Encyclopédie*. [en ligne]. 10 décembre 2013. Société Diderot. Varia, n° 48, pp. 203-224. [Consulté le 27 août 2025]. Disponible à l'adresse : <https://journals.openedition.org/rde/5047.>
- NEVEU, Valérie, 2015. L'héritage classificatoire de l'Ancien Régime. Apogée et déclin de la classification des libraires de Paris dans les bibliothèques publiques (fin XVIII^e-début XX^e siècle) : le cas de la bibliothèque municipale de Rouen. In : *Classer les archives et les bibliothèques. Mise en ordre et raisons classificatoires*. [en ligne]. Presses universitaires de Rennes. Histoire. Pp. 59-85. [Consulté le 27 août 2025]. Disponible à l'adresse : <https://books.openedition.org/pur/88620.>
- PARROCHIA, Daniel, 2017. Bibliothèque et théorie des classifications. In : *Robert Damien, du lecteur à l'électeur : Bibliothèque, démocratie et autorité*. [en ligne]. Presses de l'Enssib. Papiers, pp. 18-30. [Consulté le 27 août 2025]. Disponible à l'adresse : <https://books.openedition.org/pressesenssib/7317.>
- RANDERAAD, Nico, 2020. *States and statistics in the nineteenth century. Europe by numbers*. Juin 2020. Manchester University Press. ISBN 978-0-719-08142-2.
- RICHARDS, Elizabeth May, 1944. Alexandre Vattemare and His System of International Exchanges. In : *Bulletin of the Medical Library Association*. [en ligne]. Octobre 1944. Volume 32, n° 4, pp. 413-448. [Consulté le 27 août 2025]. Disponible à l'adresse : [https://pmc.ncbi.nlm.nih.gov/articles/PMC194400/.](https://pmc.ncbi.nlm.nih.gov/articles/PMC194400/)
- RILEY, Jenn, 2017. *Understanding Metadata. What is Metadata, and What is it For?*. [en ligne]. National Information Standards Organization (NISO). [Consulté le 27 août 2025]. Disponible à l'adresse : https://digital.library.unt.edu/ark:/67531/metadc990983/m2/1/high_res_d/understanding_metadata.pdf.
- RIVIÈRE, Pascal, 2022. Qu'est-ce qu'un répertoire ? De multiples exigences pour un système complexe. In : *Courrier des statistiques*. [en ligne]. 29 novembre 2022. Insee. N° N8, pp. 52-71. [Consulté le 25 juillet 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/information/6665186?sommaire=6665196.>

- SALVAN, Paule, 1962. *Les classifications. Préparation au diplôme supérieur de bibliothécaire. Résumé du cours.* [en ligne]. Bibliothèque nationale. [Consulté le 27 août 2025]. Disponible à l'adresse : <https://www.enssib.fr/bibliothèque-numérique/documents/48774-les-classifications.pdf>.
- SARRAZIN, Véronique, 2015. Les classifications dans les bibliothèques universitaires françaises (1878-1962). In : *Classer les archives et les bibliothèques, mise en ordre et raisons classificatoires.* [en ligne]. Presses universitaires de Rennes. Histoire. Pp. 87-107. [Consulté le 27 août 2025]. Disponible à l'adresse : <https://books.openedition.org/pur/88623>.
- STATISTIQUE GÉNÉRALE DE LA FRANCE, 1913. *Historique et travaux de la fin du XVIII^e siècle au début du XX^e, avec 103 tableaux graphiques relatifs aux travaux les plus récents. Exposition universelle et internationale de Gand en 1913.* Imprimerie nationale.
- TRICLOT, Mathieu, 2008. *Le moment cybernétique. La constitution de la notion d'information.* 24 avril 2008. Champ Vallon. Milieux. ISBN : 978-2-87673-484-5.
- VIRY, Claude-Michel, 2013. *Guide historique des classifications de savoirs : Enseignement-Encyclopédies-Bibliothèques.* 1^{er} décembre 2013. L'Harmattan. Pour Comprendre. ISBN 978-2-336-33075-4.
- VON OERTZEN, Christine, 2017. Machineries of Data Power: Manual versus Mechanical Census Compilation in Nineteenth-Century Europe. In : *Data Histories.* [en ligne]. The University of Chicago. Osiris, volume 32, n° 1, pp. 129-150. [Consulté le 30 juillet 2025]. Disponible à l'adresse : <https://www.journals.uchicago.edu/doi/10.1086/693916>.

Fonction publique : des chiffres et des lettres au cœur du débat public



Gaël de Peretti*

Au début des années 1970, le service statistique ministériel (SSM) de la fonction publique voit le jour. Son activité est décrite dans le premier numéro du « premier » Courrier des statistiques, en 1977. Près de cinquante ans plus tard, son champ d'intervention s'est élargi : de la fonction publique de l'État, essentiellement, à l'ensemble de la fonction publique (de l'État, territoriale et hospitalière). Les thématiques suivies sont toujours plus nombreuses, en lien avec l'exploitation d'enquêtes de la statistique publique, et les concepts se sont précisés pour faciliter les comparaisons : d'une part entre les trois versants de la fonction publique, d'autre part avec le secteur privé. Les sujets traités par le SSM de la fonction publique sont éminemment politiques : ils sont encadrés par des instances de concertation et de discussion avec des utilisateurs attentifs, comme les organisations syndicales et les employeurs. Ces instances, tout comme le Conseil national de l'information statistique (Cnis), permettent d'orienter les travaux du SSM en tenant compte d'une demande sociale exprimée, voire revendiquée. Présentes depuis le début, certaines difficultés perdurent : celles inhérentes à l'utilisation de données de gestion des ressources humaines pour produire de l'information statistique. Les récents travaux en la matière sont source d'amélioration à venir.

 At the beginning of the 1970s was born the Ministerial Statistical Office (SSM) of the Civil Service. Its activities are described in the first issue of the "first" Courrier des statistiques, in 1977. Almost fifty years later, its scope of intervention has broadened: starting essentially from the state civil service to eventually encompass the whole civil service (state, territorial and hospital). More and more themes are being explored, in relation to the exploitation of public statistics surveys, and concepts are now more precise to facilitate comparisons: between the three constituents of the civil service on the one side, and with the private sector on the other. The topics handled by the SSM of the Civil Service are eminently political: they are supervised by consultation and discussion bodies with watchful users, like unions and employers. These bodies as well as the National Council on Statistical Information (CNIS) can guide the SSM's works while taking into consideration social revindications that may be expressed, or even claimed. Some difficulties, present from the start, remain: those inherent to the use of human resources management data for the production of statistical information. Recent works may offer future improvement in that regard.

* Sous-directeur, sous-direction des études, des statistiques et des systèmes d'information, DGAFF.
gael.de-peretti@finances.gouv.fr

Dans sa thèse « Trop de fonctionnaires ? Contribution à une histoire de l'État par ses effectifs (France, 1850-1950) », Émilien Ruiz retrace l'histoire des statistiques « construites » sur les fonctionnaires, et à travers elle, celle des politiques les concernant (Ruiz, 2013). Ce qu'il constate, c'est qu'à chaque période peuvent coexister des définitions différentes, et donc des chiffres différents. Pour arriver à tirer parti de ces chiffres qui, selon lui,

« offrent de précieuses informations concernant les rythmes et les structures de transformations de l'État », un travail de « reconstruction » est nécessaire (Maruani et Meron, 2012). Ruiz (2021) montre que cette persistance de définitions et chiffres concurrents se poursuit : en 2018, entre une vision restrictive, les 1,5 million de fonctionnaires de la fonction publique de l'État, et une vision extensive, les 7,8 millions d'agents du secteur « Administration publique », les écarts sont considérables.

De son côté, le **service statistique ministériel de la fonction publique (SSM FP)**¹ publie chaque année de nombreuses statistiques dans le cadre du rapport annuel sur l'état de la fonction publique

(RAEFP)² : sur les effectifs de la fonction publique, mais aussi sur les flux de personnels (recrutements, retraites...), les rémunérations, la formation, la durée et l'organisation du travail, etc. Cet article est l'occasion de retracer l'histoire du SSM FP au travers de différents prismes, dont en particulier l'étude de cette publication phare qu'est le RAEFP.

► Un ancrage au sein de la DGAFP dès l'origine

Le SSM FP a toujours fait partie de la direction générale de l'administration et de la fonction publique (DGAFP). Cette dernière a été créée en 1945 et, dès l'origine, il est précisé qu'elle est notamment chargée « d'établir ou de faire établir une documentation et des statistiques d'ensemble touchant la fonction publique »³. Au début des années 1970, cette activité apparaît clairement dans l'organigramme de la direction sous le nom de « cellule statistique » (Secrétariat général du Gouvernement, 2024) : le SSM FP est né. Et c'est bien cette cellule qui est nommée dans la liste des services statistiques des ministères qui figure à la fin des différents numéros du Courrier des statistiques, depuis le premier opus de cette collection (Insee, 1977).

Au fil des réorganisations de la DGAFP, le SSM FP évolue⁴. En 1990, la cellule statistique devient le bureau Statistiques, bureau à vocation transversale directement rattaché au chef de service⁵, puis, en 2007, le bureau des statistiques, des études et de l'évaluation,

¹ Voir le lien : <https://www.fonction-publique.gouv.fr/etudes-statistiques-et-publications/etudes-statistiques>.

² Voir par exemple le rapport relatif à l'année 2025 à l'adresse suivante : <https://www.fonction-publique.gouv.fr/toutes-les-publications/rapport-annuel-sur-l-etat-de-la-fonction-publique-edition-2025>.

³ Ordonnance n° 45-2283 du 9 octobre 1945 relative à la formation, au recrutement et au statut de certaines catégories de fonctionnaires, et instituant une direction de la fonction publique et un conseil permanent de l'administration civile, titre V, article 15, alinéa 2. Voir les références juridiques en fin d'article.

⁴ Voir en fin d'article les références juridiques des différents textes relatifs à l'organisation de la DGAFP.

⁵ Qui deviendra un directeur adjoint à la suite de la publication du décret n° 98-574 du 10 juillet 1998 portant création d'un emploi de directeur, adjoint au directeur général de l'administration et de la fonction publique. Voir les références juridiques en fin d'article.



L'ancre, depuis son origine, du SSM FP au sein de la DGAFF lui permet d'être au cœur de l'organisation où se conçoivent les politiques de la fonction publique, notamment celles de l'État.



au sein de la sous-direction de l'information et de la légitique⁶. En 2012, à la suite d'une nouvelle réorganisation, le SSM FP est renommé département des études et statistiques, au sein du service du pilotage et des politiques transversales. En 2017, il devient le département des études, des statistiques et des systèmes d'information et se structure alors en deux bureaux : « études et statistiques » et « applications et systèmes d'information pour la fonction publique ». En 2019, il se transforme en sous-direction des études, des statistiques et des systèmes d'information (SDessi). Enfin, en 2022, et c'est l'organisation actuelle, la SDessi se structure

en deux départements : celui des études et des statistiques et celui de la donnée et des systèmes d'information relatifs aux ressources humaines.

Depuis l'origine du SSM FP, son ancrage au sein de la DGAFF lui permet d'être au cœur de l'organisation où se conçoivent les politiques de la fonction publique, notamment celles de l'État. Il lui assure une proximité avec les acteurs interministériels et ministériels des politiques de ressources humaines, les organisations syndicales et, de façon plus distanciée, avec les deux autres versants de la fonction publique (hospitalière et territoriale).

► Les fichiers de la paie de l'État au cœur du système d'information

Un article est consacré au SSM FP dans le premier numéro du Courrier des statistiques. Dominique Quarré décrit les travaux lancés en 1973 pour mettre en place un système d'information sur les agents de l'État (Quarré, 1977). Au lancement du projet, dénommé « projet Fonction publique », l'Insee, la direction du budget et la DGAFF constituent un groupe pour le piloter conjointement. C'est à cette occasion que la DGAFF décide de créer sa cellule statistique, à la fois pour participer efficacement à ces travaux et pour réaliser d'autres études statistiques utiles aux besoins internes de la direction (Lhuillier, 1978).

L'objectif du projet Fonction publique est de produire des statistiques sur les effectifs et les rémunérations des agents de la fonction publique de l'État (FPE), mais aussi sur la masse salariale qu'ils représentent, à la fois sur le champ le plus complet et de façon la plus automatisée possible. Pour ce faire, l'idée est d'exploiter des fichiers informatiques issus des systèmes de paie des plus grosses administrations (effectuant chacune la paie d'un grand nombre d'agents de la FPE) et de compléter le champ par une enquête auprès des administrations plus petites. Les administrations concernées par la transmission de fichiers informatiques représentent alors environ 75 % des effectifs de la FPE : direction de la comptabilité publique, ministères de l'Intérieur, des PTT⁷, de la Coopération et des Affaires étrangères, etc.

⁶ La légitique est l'ensemble des règles, principes et méthodes utiles à la conception et à la rédaction des textes législatifs et réglementaires (source : Larousse).

⁷ Postes, Télégraphes et Téléphones. Ce ministère a disparu en 1991, après avoir changé plusieurs fois d'intitulé, le dernier étant « Postes et Télécommunications ».

À sa création, la cellule statistique de la DGAFF regroupe deux personnes. Elle s'étoffe rapidement et, en janvier 1978, elle compte treize personnes (Lhuillier, 1978). Elle travaille avec l'ensemble des directions du personnel des administrations, les services statistiques ministériels quand ils existent, ainsi que les administrations à vocation transverse comme les directions du budget et de la comptabilité publique. Surtout, elle a une relation privilégiée avec l'Insee dans le cadre du projet Fonction publique. En effet, l'Insee alloue des moyens de conception (Dominique Quarré, alors au sein de la division Emploi) et des ressources informatiques (quatre personnes du département de l'informatique). L'institut gère également la partie impression et diffusion des publications.

[O]n ne connaît jamais avec certitude la fiabilité de l'information recueillie et on n'est jamais certain de son exhaustivité, d'où la nécessité de passer cette information au crible afin de redresser les plus grossières erreurs [...].

Au-delà de ce projet, la cellule réalise de l'analyse de documents budgétaires et administratifs, en particulier sur la masse salariale et les recrutements par concours. Elle pilote également des enquêtes auprès des services de la FPE, plutôt sous forme de tableaux récapitulatifs, dont la gestion de collecte est déléguée aux administrations centrales. Les ambitions de ces enquêtes sont généralement limitées à une information minimale et cohérente entre l'ensemble des administrations. Mais, au final, « on ne connaît jamais avec certitude la fiabilité de l'information recueillie et on n'est jamais certain de son exhaustivité, d'où la nécessité de passer cette information au crible afin de redresser les plus grossières erreurs » (Lhuillier, 1978).

Dès cette époque, il est envisagé à terme de mobiliser les systèmes d'information sur les ressources humaines (SIRH) des différentes administrations de la FPE, sous réserve que ces derniers deviennent unifiés en termes de concepts, afin de remonter plus facilement une information fiable et cohérente entre les services. Cependant, certaines des difficultés rencontrées dans cet objectif aux prémissives du SSM FP perdurent toujours.

► La nécessité de prendre le temps de « qualifier » les données

Quinze ans plus tard, les effectifs du SSM FP ont peu évolué (12 personnes), comme on peut le constater dans le supplément du Courrier des statistiques d'octobre 1992 qui présente les effectifs et organigrammes des SSM (Insee, 1992b).

En juin 1994, dans le cadre de sa formation « Éducation, emplois, revenus », le Conseil national de l'information statistique (Cnis)⁸ adopte l'avis suivant : « Le Conseil souligne l'intérêt qu'une mission soit conduite dans le but de : dresser un inventaire de l'état des données statistiques existantes sur l'ensemble des trois fonctions publiques (territoriale, État, hospitalière) ; proposer un système d'observation coordonné sur les trois fonctions publiques en matière d'emploi, salaires, accidents du travail... ; recommander l'utilisation

⁸ La mission principale du Cnis est d'assurer la concertation entre les producteurs et les utilisateurs de la statistique publique. Anxionnaz et Maurel (2021) décrivent le rôle et le fonctionnement de cette instance.

de nomenclatures adaptées ». Cet avis conduit à la mise en place d'une mission réalisée en deux temps : d'abord « un état des lieux des dispositifs statistiques », puis « du constat aux recommandations ». Démarrée en 1995, la mission rend son rapport début 1997 (Neyret et Ponsot, 1997).

Entre-temps, Françoise Ballu présente le bureau des statistiques de la fonction publique dans le Courrier des statistiques de décembre 1995 (Ballu, 1995), bureau qu'elle dirige depuis juin 1992 (Insee, 1992a). Les activités couvertes restent identiques pour l'essentiel à celles décrites par Hugues Lhuillier en 1978. Elles se partagent principalement entre :

- **de nombreuses enquêtes** : concours et formation professionnelle dans la FPE, positions statutaires⁹ des agents des ministères, emplois de direction et d'inspection (sous l'angle de l'égalité entre femmes et hommes), handicap ;
- **l'exploitation de documents ou fichiers administratifs** : fascicules budgétaires, procès-verbaux des élections professionnelles aux commissions administratives paritaires, remontées de données administratives (sur les rémunérations, les crédits sociaux, la discipline, les grèves, etc.) ;
- **et, bien sûr, l'exploitation des fichiers de paie des agents de l'État consolidés par l'Insee**, lesquels permettent de produire des informations détaillées à la fois sur les effectifs et les rémunérations dans la FPE.

Il y a une réelle difficulté à concilier le traitement des réponses à la demande et les missions de statistique publique. En effet, même si les outils informatiques sont plus réactifs, le travail préalable pour qualifier les données est long.

À ces travaux récurrents s'ajoutent **d'importants travaux de réponse à la demande**, en lien avec des demandes internes à la DGAFP ou externes (cabinet ministériel, inspections générales, Cour des comptes, etc.).

Depuis 1977, les difficultés exprimées sur la qualité des données remontant des ministères persistent. Elles s'expliquent par différentes raisons : une gestion éclatée entre directions et établissements publics au sein des « départements ministériels », un développement hétérogène des systèmes d'information dédiés à la gestion des ressources humaines (RH), un turnover important au sein des structures en

charge de répondre au SSM FP, pas d'obligation de réponse comme cela peut exister pour les enquêtes auprès des entreprises, etc. Plus généralement, il y a une réelle difficulté à concilier le traitement des réponses à la demande, souvent urgentes, et les missions de statistique publique. En effet, même si les outils informatiques sont plus réactifs, le travail préalable pour qualifier¹⁰ les données est long (Martel, 1995).

⁹ Afin de connaître les effectifs des personnes en situation de détachement, de mise à disposition, de disponibilité, en position hors cadre ou en congé parental.

¹⁰ Selon Cotton et Haag (2023), avant d'intégrer une source dans un processus statistique, il convient de s'assurer qu'elle dispose des qualités suffisantes pour pouvoir être utilisée à des fins statistiques. Ainsi, il est nécessaire d'échanger avec le producteur de la donnée afin de vérifier que la source est : exploitable (les données contenues peuvent être restructurées pour mesurer des concepts statistiques) ; complète (aucune sous-couverture évidente qui empêcherait son exploitation) ; disponible dans un délai raisonnable ; documentée (présence de métadonnées).

► Des nomenclatures à constituer pour s'extraire des référentiels de gestion



Le système de classification est souvent très détaillé : il ne correspond pas nécessairement à la notion de métier exercé, ce qui est logique dans une Fonction publique de carrière et non d'emploi [...].



L'usage à des fins de statistique publique des données de gestion des ministères se heurte à une autre difficulté : **l'absence de nomenclature professionnelle**, c'est-à-dire d'une nomenclature qui permette une description socioéconomique des agents publics. Ce sujet est détaillé dans le dossier consacré à l'emploi public du Courrier des statistiques d'août 1996 (Insee, 1996). Ballu et Martel (1996) y décrivent les différentes nomenclatures de corps, grades et emplois dans la FPE. Il s'agit de nomenclatures

comptables, utilisées pour la paie des agents de l'État. Ainsi, « les titulaires (et stagiaires) relèvent d'un système de classification du type « corps, grade, échelon », qui entraîne un niveau indiciaire et un traitement indiciaire basé sur la valeur du point de la fonction publique. Il en va de même pour les non-titulaires. [...] Le système de classification est souvent très détaillé : il ne correspond pas nécessairement à la notion de métier exercé, ce qui est logique dans une Fonction publique de carrière et non d'emploi » (Neyret et Ponsot, 1997). Ceci ne facilite donc pas le travail de l'Insee et du SSM FP.

Une dernière difficulté qui persiste est celle d'arriver à **gérer simultanément deux types de préoccupations : d'un côté, une demande d'information extrêmement fine**, pour tenir compte de toutes les spécificités réglementaires de la population d'étude (qu'il s'agisse de la seule FPE ou des trois fonctions publiques), **de l'autre, une visée statistique**, avec le souci d'intégrer cette population d'étude dans le champ plus global de l'emploi (public ou privé), aussi bien au niveau national que local (Ponsot, 1996). Or, cette intégration suppose de disposer d'une nomenclature stable et cohérente entre les secteurs. Cette double préoccupation est d'autant plus difficile à concilier qu'à la fin des années 1990, le système d'information reste éclaté : à chaque versant de la fonction publique correspond un dispositif spécifique, indépendant de celui utilisé pour produire de l'information statistique sur l'emploi et les rémunérations du secteur privé.

Au-delà du sujet des nomenclatures, le rapport du Cnis préconise de mettre en cohérence l'information, mais aussi de développer son accessibilité et sa visibilité au travers de l'établissement régulier d'un rapport de synthèse sur les trois fonctions publiques. Il est également important de s'accorder sur le périmètre de la fonction publique.

► La question du champ de la fonction publique

Dans son article sur les frontières mouvantes de l'emploi public, Brenot-Ouldali (1996) rappelle les deux approches possibles de l'emploi public :

- celle restreinte à **la seule fonction publique**, reposant sur les critères croisés du statut juridique de l'employeur (**catégories juridiques¹¹** 71, 72 et 73) et du statut de recrutement de droit commun des agents (droit public) ;
- celle de la comptabilité nationale, englobant **l'ensemble des administrations publiques (APU)¹²**.

Ces deux périmètres ont une importante partie commune, mais ne se recouvrent pas totalement. Ainsi, en 2023, la fonction publique emploie 5,8 millions d'agents, dont 0,1 million ne font pas partie des APU (essentiellement des établissements publics marchands comme les offices publics d'habitat¹³), tandis que les APU emploient 6,5 millions d'agents, dont 0,8 million ne font pas partie de la fonction publique (essentiellement des organismes de protection sociale et des hôpitaux à but non lucratif). **Le SSM FP consacre toutes ses publications statistiques au champ de la fonction publique tel que défini ci-dessus.**

En 2001, dans le Courrier des statistiques de septembre consacré au système statistique public français (Insee, 2001), un encadré présente le partage des tâches entre l'Insee et les services statistiques ministériels. Pour la statistique démographique et sociale, il est précisé que le SSM FP « suit l'emploi public et les rémunérations des agents de l'État ». Les missions décrites dans ce numéro restent par ailleurs les mêmes que celles détaillées en 1995.

La revue souligne aussi la création en juillet 2000 de l'Observatoire de l'emploi public. Ce dernier est notamment chargé :

- de réaliser des études statistiques et prospectives sur l'emploi dans les trois fonctions publiques ;
- de présenter au Parlement un état statistique annuel des effectifs de la FPE ;
- de formuler des propositions pour la mise en place de systèmes d'information permettant d'harmoniser les données recueillies pour chacune des fonctions publiques.

Cette création est clairement une réponse aux recommandations du rapport du Cnis de 1997. Nous verrons par la suite ce qu'est devenu cet observatoire.

¹¹ Voir le lien : <https://www.insee.fr/fr/information/2028129>.

¹² Voir le lien : <https://www.insee.fr/fr/metadonnees/definition/c1244>.

¹³ Organismes publics qui construisent et gèrent les habitations à loyer modéré (HLM).

► De la fonction publique de l'État à l'ensemble de la fonction publique

En 2008, dans la rubrique « Il y a trente ans dans le Courrier des statistiques »¹⁴, en écho à l'article d'Hugues Lhuillier de 1978, le journal réalise une interview de Christine Gonzalez-Demichel, responsable à l'époque du SSM FP (Insee, 2008). Cette interview permet de prendre conscience de la forte évolution des missions et activités dans les années 2000, en lien avec les recommandations du rapport du Cnis, mais aussi avec l'évolution du contexte institutionnel : décentralisation accrue, mise en place de la loi organique du 1^{er} août 2001 relative aux lois de finances¹⁵, révision générale des politiques publiques¹⁶, etc.

En 2007, la DG AFP s'est à nouveau réorganisée. Les missions du SSM FP sont désormais les suivantes :

- élaborer et mettre en œuvre le programme d'études et de recherche sur la gestion des ressources humaines dans la fonction publique, dans sa dimension statistique ;
- élaborer et diffuser des données et des études statistiques sur la fonction publique ;
- concevoir et gérer le système d'information statistique sur la fonction publique ;
- contribuer à la mesure de l'impact, de l'efficience et de l'efficacité des politiques de gestion des ressources humaines par des actions d'évaluation ;
- établir et présenter des données et informations pour l'Observatoire de l'emploi public (OEP).

Dans ce cadre, le SSM élargit son champ de compétences à l'ensemble de la fonction publique (versus la seule FPE) et à la thématique des retraites. Il reprend aussi la production de l'[indice de traitement brut - grille indiciaire](#)¹⁷, préalablement calculé par l'Insee. Le SSM finance également des suréchantillons « fonction publique »¹⁸ dans des grandes enquêtes de la statistique publique. Enfin, il souhaite œuvrer pour intégrer l'emploi public dans l'emploi total et permettre ainsi des comparaisons avec le secteur privé.

Du côté des évolutions techniques, les collectes auprès des administrations sont dématérialisées, voire réalisées par Internet, et plusieurs projets devraient faciliter l'accès à des données homogènes, comme le [système d'information sur les agents du service public \(Siasp\)](#)¹⁹, la norme [DADS-U](#)²⁰ et l'Opérateur national de paye²¹. En particulier, le Siasp doit permettre d'améliorer la production de connaissance statistique sur l'emploi, les rémunérations et les caractéristiques des agents publics (âge, sexe et statut, par exemple).

14 Cette rubrique du « premier » Courrier des statistiques a été créée en décembre 2004. Pour mettre en avant son intérêt, la revue a alors réédité un article de Michel Volle intitulé « Pourquoi faire l'histoire de la statistique » (Volle, 1977).

15 Voir les références juridiques en fin d'article.

16 Réforme engagée en 2007, visant à réorganiser les missions de l'État pour améliorer l'efficacité des politiques publiques et réaliser des économies budgétaires.

17 Voir le lien : <https://www.insee.fr/fr/metadonnees/source/indicateur/p1648/description>.

18 Il s'agit d'augmenter suffisamment l'échantillon global d'agents de la fonction publique interrogés dans les enquêtes pour pouvoir mener des analyses ciblées sur ce champ.

19 Voir le lien : <https://www.insee.fr/fr/information/2044926>.

20 Voir le lien : <https://www.insee.fr/fr/metadonnees/source/serie/s1163>.

21 Ce projet de prestataire centralisé de la paye des agents de l'État a été abandonné en 2014 au profit de la création, en 2015, du Centre interministériel de services informatiques relatifs aux ressources humaines (CISIRH).

Dans cette première décennie des années 2000, le SSM FP prend de l'épaisseur et devient responsable des travaux et études statistiques menés auparavant par l'Observatoire de l'emploi public et, surtout, de la rédaction en chef du volume « Faits et chiffres » du rapport annuel sur l'état de la fonction publique. Ce rapport annuel, mixant à l'origine des éléments sur la politique RH de l'État et des statistiques (voir infra), est l'ouvrage phare de la DGAFP : en même temps qu'il dresse un état des lieux statistique de la fonction publique, il retrace l'activité de ses sous-directions métiers.

► Le rapport annuel sur l'état de la fonction publique, une publication statistique ?²²

Le RAEFP, et plus particulièrement sa partie « Faits et chiffres », structure fortement l'activité du SSM FP. Les « Faits et chiffres » constituent aujourd'hui l'entièreté du rapport, alors que dans l'édition 2022, il s'agissait de la deuxième partie, la première étant consacrée aux « Politiques et pratiques en ressources humaines ». Cependant,

le 29 juin 2023, l'Autorité de la statistique publique (ASP)²³ a auditionné le chef du SSM FP et a publié ensuite un délibéré, daté du 22 septembre 2023²⁴, dans lequel elle note :

Ce processus [...] ne respecte pas, en termes de support de diffusion et de délai d'accès anticipé, les principes du cadre de référence commun pour le processus de diffusion des publications statistiques du SSP [...].

« La principale question porte toutefois, du point de vue de l'Autorité, sur le fait que la production statistique du SSM est pour une très large part diffusée par l'intermédiaire du rapport annuel sur l'état de la fonction publique, dont le Conseil commun de la fonction publique (CCFP), instance nationale de dialogue social, reçoit communication et débat préalablement à sa transmission au Parlement. Ce processus [...] ne respecte pas, en termes de support de diffusion et de délai d'accès anticipé, les principes du cadre de référence commun pour le processus de diffusion des publications statistiques du SSP adopté en début d'année et le délibéré de l'ASP du 15 mars 2023²⁵, qui en recommande une mise en œuvre opérationnelle rapide ».

anticipé, les principes du cadre de référence commun pour le processus de diffusion des publications statistiques du SSP adopté en début d'année et le délibéré de l'ASP du 15 mars 2023²⁵, qui en recommande une mise en œuvre opérationnelle rapide ».

En décembre 2023, la partie « Politiques et pratiques en ressources humaines » du rapport est supprimée. **En novembre 2024, le SSM FP transforme complètement le rapport en publication statistique**, d'une part en supprimant l'éditorial du ministre de la Fonction publique et d'autre part en respectant le cadre de référence des publications statistiques, en particulier sur le délai d'accès privilégié. Désormais, le rapport est

²² Les parties qui suivent s'inspirent fortement de Peretti (2024a) et de Peretti (2024b).

²³ L'Autorité de la statistique publique (ASP) veille à l'indépendance professionnelle dans la conception, la production et la diffusion de statistiques publiques. Elle assure également une vigilance quant au respect des principes d'objectivité, d'impartialité, de pertinence et de qualité des données produites, en référence aux recommandations européennes en matière de bonnes pratiques statistiques. Voir le lien vers le site de l'ASP : <https://www.autorite-statistique-publique.fr/>.

²⁴ Voir le lien : <https://www.autorite-statistique-publique.fr/event/delibere-suite-a-laudition-du-chef-du-service-statistique-ministere-du-ministere-charge-de-la-fonction-publique/>.

²⁵ Voir le lien : <https://www.autorite-statistique-publique.fr/event/delibere-sur-le-cadre-de-reference-pour-le-processus-de-diffusion-des-publications-statistiques-du-service-statistique-public-ssp/>.

diffusé immédiatement à l'issue de sa présentation devant la formation spécialisée du CCFP sur les questions relatives aux évolutions de l'emploi public, à la politique des retraites dans la fonction publique et à la connaissance statistique de la situation, de la rémunération et des pensions des agents publics. C'est une véritable rupture par rapport à la nature éminemment politique du RAEFP à son origine. Ce changement de statut de la publication demandera pédagogie et fermeté pour que l'ensemble des parties prenantes (employeurs, organisations syndicales, cabinet ministériel, presse, etc.) en saisissent toutes les conséquences.

► À l'origine, un rapport éminemment politique

Revenons donc à l'origine pour mesurer le chemin parcouru. **Le rapport naît officiellement du décret n° 76-509 du 10 juin 1976 relatif au Conseil supérieur de la fonction publique²⁶**. Dans son article premier, il est écrit : « Il [le Conseil supérieur de la fonction publique] entend un rapport annuel sur l'état de la fonction publique ». De fait, en juin 1977, le premier rapport est présenté au Conseil et il est loin d'être une publication statistique. Il est avant tout un exposé des mesures et orientations prises en matière de politique générale de la fonction publique. Il contient cependant une annexe statistique portant sur la fonction publique de l'État. Celle-ci comprend en particulier des données sommaires sur les effectifs de la FPE par statut, catégorie, sexe et âge, ainsi que sur les dépenses de personnel, le travail à mi-temps et les arrêts maladie.

À l'inverse, le rapport de 1978 commence par la connaissance statistique dans la fonction publique²⁷, avant d'enchaîner sur l'ensemble des mesures de politique salariale, sociale, statutaire, de recrutement et de formation. La partie statistique est plus détaillée que celle du premier rapport et elle s'étoffe encore en 1979, puis en 1980. Notamment, le rapport de 1980 présente les sources alors utilisées : le recensement de la population, les fichiers de paie des agents de l'État, l'*enquête Emploi*²⁸ et des données budgétaires.

L'année 1981 marque un premier tournant : il s'agit maintenant d'un véritable ouvrage publié par La Documentation française (figure 1). Il comporte par ailleurs pour la première fois un éditorial du ministre de la Fonction publique et des réformes administratives, Anicet Le Pors. Cet éditorial, titré « Un rapport sans précédent », souligne les avancées réalisées : « [...], l'abondance et la richesse des informations et des idées contenues dans le rapport établi aujourd'hui sont sans précédent ; de même que la qualité de la présentation ».

Pour ce premier véritable ouvrage, le choix est fait de le nommer « La fonction publique en 1981 ». Dans les faits, l'essentiel des statistiques porte sur la FPE. Il y a bien une courte analyse de l'emploi public à partir de l'enquête Emploi au début du premier chapitre « L'état statistique de la fonction publique et sa place dans la nation ». Pour le reste,

²⁶ Voir les références juridiques en fin d'article.

²⁷ Cette partie est la reprise intégrale d'une note que le SSM FP considère comme la version zéro du RAEFP, du moins au sein de la Bibliothèque numérique de la statistique publique. Il s'agit d'une note titrée « Principales données statistiques sur la fonction publique » portant essentiellement sur l'année 1975, qui a été cataloguée comme « Rapport annuel 1975 » et ainsi datée en 1975 (Direction générale de l'administration et de la fonction publique, 1975). Il est vraisemblable que ce ne soit pas le cas. Il faudrait consulter les archives du Conseil supérieur de la fonction publique pour trancher ce sujet.

²⁸ Voir le lien : <https://www.insee.fr/fr/metadonnees/source/serie/s1223>.

cependant, tout porte sur la FPE : emploi détaillé par catégorie, sexe, âge, etc., dépenses de personnel, répartition indiciaire des agents et, pour la première fois, salaire moyen par groupe social.

► Publication statistique ou bilan social ?

Les tableaux statistiques [...] sont présentées sous la forme d'un bilan social inspiré de celui prévu par le code du travail pour les entreprises et organismes relevant du droit privé [...].

Cet état de fait – le centrage sur la FPE – sera confirmé en 1984 avec le changement de titre de l'ouvrage au profit de « La fonction publique de l'État en 1984 »²⁹, même si la fonction publique hospitalière est intégrée dans les analyses sur l'emploi. Désormais, l'emploi est quantifié en effectif au 31 décembre de l'année et les salaires sont des salaires nets moyens pour les personnes à temps complet. **À partir de l'édition 1984, l'essentiel des statistiques est présenté dans une partie « Le bilan social »** introduite de la façon suivante : « Les tableaux statistiques et les informations chiffrées relatifs au bilan d'activité de la fonction publique de l'État sont présentés sous la

forme d'un bilan social inspiré de celui prévu par le code du travail pour les entreprises et organismes relevant du droit privé, sous réserve des nécessaires adaptations qui s'imposent d'elles-mêmes ».

Ce bilan social, composé d'une centaine de pages et de plus de quatre-vingts tableaux, s'appuie sur plusieurs sources : des travaux réalisés par l'Insee publiés dans sa collection Économie et Statistique, les fichiers de paie de la FPE, une enquête de la DGAFP auprès des directions du personnel des différents ministères et de nombreuses remontées de données des ministères ou budgétaires. Le rapport aborde de très nombreux sujets : emploi, recrutement, retraite, handicap, rémunérations, temps de travail, formation, action sociale, etc. Il va garder cette structure sans trop évoluer jusqu'au milieu des années 1990, si ce n'est avec un niveau de détail parfois plus important, en particulier sur le sujet rémunération (indice, traitement, primes et indemnités, etc.).

La nouveauté la plus significative est l'apparition, à partir du rapport 1993/1994, du concept de **rémunération moyenne des personnes en place (RMPP)**. Il s'agit du salaire net moyen en « équivalent temps plein » (converti à un temps plein pendant toute l'année) des personnes présentes deux années consécutives chez le même employeur et avec la même quotité de temps de travail. Calculée sur une sous-population stable par construction, l'évolution de la RMPP est décomposée entre les effets respectifs :

- des mesures générales (telles que la revalorisation du point d'indice de la fonction publique) ;
- des mesures catégorielles (ciblées sur certains groupes d'agents) ;
- des progressions de carrière (avancements d'échelons et promotions de grades ou de corps).

²⁹ Ce titre sera conservé jusqu'en 2002, modulo trois éditions (1998/1999, 1999/2000, 2000/2001) pour lesquelles il est nommé « La fonction publique et la réforme de l'État ».

► Figure 1 - Évolution du RAEFP : 1981-2025





Lecture : Cet échantillon de couvertures du RAEFP illustre les grands changements qui ont affecté le rapport quant à son périmètre et à sa nature. Pour son édition 2025, il adopte l'**identité visuelle commune pour l'Insee et les services statistiques ministériels**.

► Des missions pour améliorer les statistiques sur l'emploi public

Dans le rapport 1995/1996, de nouveau, le champ des statistiques est étendu à l'ensemble de la fonction publique. Les versants État et territorial s'appuient sur des travaux de l'Insee et le versant hospitalier s'appuie sur des travaux du service statistique ministériel de la santé³⁰. Toujours dans ce même rapport sont mentionnées une mission de l'inspection générale de l'administration de l'Éducation nationale (IGAEN) et la mission du Cnis déjà évoquée. La première a vocation à documenter les problèmes de remontées d'informations statistiques en provenance des directions de personnel et de déterminer les causes des insuffisances actuelles.

Ces travaux se traduisent l'année suivante, dans le rapport annuel 1996/1997, par une présentation des préconisations du Cnis et un détail sur les sources existantes. Pour la première fois, également, **les entreprises publiques PTT et France Télécom sont sorties du champ**.

Parmi les préconisations du Cnis sont mises en avant :

- la nécessité d'articuler les nomenclatures entre les versants et avec la PCS³¹ (nomenclature de l'Insee sur les professions et catégories sociales), pour faciliter les comparaisons entre versants, mais aussi avec le secteur privé ;
- la nécessité de refondre la nomenclature des corps, grades et emplois ;
- la publication annuelle d'une synthèse sur l'emploi et les rémunérations, coordonnée par l'Insee.

Il faut attendre le rapport 1998/1999 pour voir une nouvelle évolution dans les statistiques de l'emploi public, avec la **prise en compte des contrats aidés**. L'année suivante, s'il n'y a pas de nouvelle modification du champ de l'emploi public, il faut noter la parution d'une étude intitulée « Séries longues sur l'emploi dans la fonction publique 1990-1998 ». Dans son avant-propos, celle-ci se place dans le cadre des recommandations du Cnis et plus précisément « d'une meilleure diffusion de l'information statistique portant sur l'ensemble de la fonction publique » (Brenot-Ouldali, 2000).

► L'Observatoire de l'emploi public : un nouveau rapport rapidement « ingéré »

À la fin des années 1990, une Délégation interministérielle de la réforme de l'État est créée. Elle succède au comité interministériel pour la réforme de l'État et au Commissariat à la réforme de l'État, et est chargée de produire un rapport annuel. Ce dernier est intégré dans l'édition 1998/1999 du rapport de la DGAEP, alors renommé « La fonction publique et la réforme de l'État » jusqu'à l'édition 2000/2001. Puis ces deux rapports sont dissociés, le rapport annuel 2001/2002 étant consacré aux seules statistiques, avec pour

³⁰ Nommé à l'époque SESI, pour service des statistiques, des études et des systèmes d'information, et remplacé en 1998 par la Drees, direction de la recherche, des études, de l'évaluation et des statistiques.

³¹ Voir le lien : <https://www.insee.fr/fr/information/6208292>.

la première fois de courtes analyses des nombreuses figures (tableaux et graphiques) en tête de chaque chapitre : budget, effectifs réels, égalité professionnelle, rémunérations, temps de travail, recrutement, retraites, relations professionnelles, politique sociale et conditions de travail, formation.

En parallèle, comme on l'a évoqué précédemment, un nouvel acteur apparaît : l'Observatoire de l'emploi public. Créé par le décret n° 2000-663 du 13 juillet 2000³², l'Observatoire a deux missions selon sa secrétaire générale Catherine Zaidman : « La première est d'assurer la transparence sur l'emploi public, la seconde est de favoriser la mise en place de la gestion prévisionnelle des emplois, des effectifs et des compétences (GPEEC) [...]. Il ne s'agit pas d'une structure de négociation mais d'une structure de concertation, [...]. Il établit tous les ans un rapport [qui] présente des données sur la fonction publique selon différents concepts (juridique et économique) et procède à des analyses sur l'emploi public » (Zaidman, 2007).

Comme on peut le constater, les rapports sur la fonction publique et sa réforme se multiplient. À cette époque, le SSM FP contribue à la fois au rapport annuel de la DGAFP et à celui de l'Observatoire de l'emploi public. Mais la situation va rapidement se clarifier. En 2003, le rapport de la DGAFP devient le « Rapport annuel sur la fonction publique » et est sous-titré « Faits et chiffres ». Les statistiques s'étoffent, portent plus souvent sur l'ensemble de la fonction publique, et des séries longues sont intégrées. Le rapport de l'Observatoire est finalement intégré dans l'édition 2006/2007 du rapport de la DGAFP, qui prend alors son titre actuel : Rapport annuel sur l'État de la fonction publique. Enfin, en 2007-2008, la partie « Faits et chiffres » adopte la structure d'aujourd'hui :

- des vues d'ensemble mises à jour annuellement et enrichies au gré des échanges avec les partenaires sociaux et les employeurs ;
- des fiches thématiques (appelées aujourd'hui « données complémentaires »), sous forme de tableaux et graphiques (désormais entièrement dématérialisés), avec la constitution dès que possible de séries longues ;
- des études, parfois regroupées dans un dossier, qui permettent de faire des zooms sur des sujets d'actualité ou innovants.

Finalement, l'Observatoire disparaît en 2012 au profit du Conseil commun de la fonction publique. Celui-ci reprend, parmi ses nombreuses missions, celles attribuées auparavant à l'Observatoire³³.

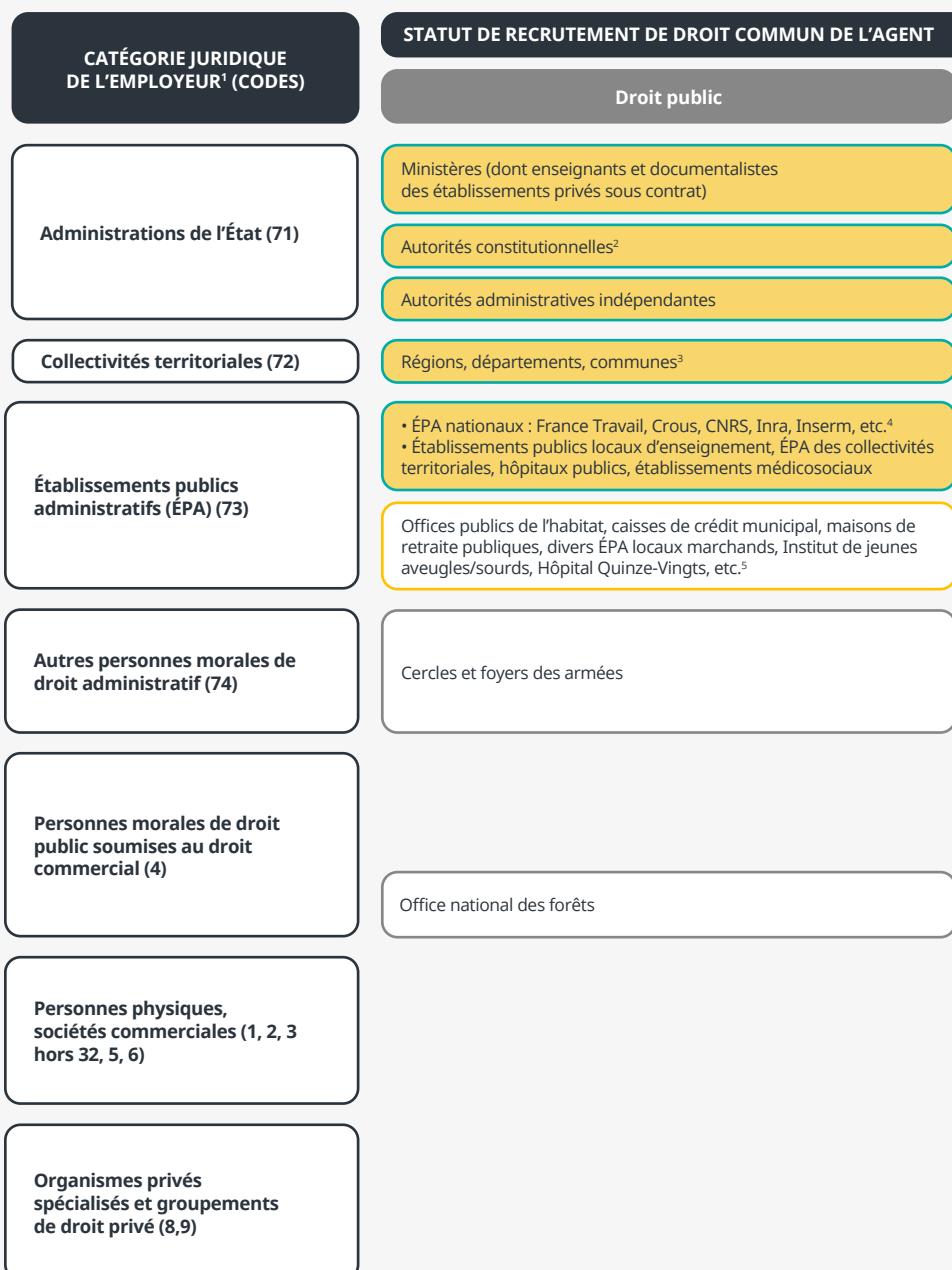
► Des statistiques qui s'appuient essentiellement sur le système d'information sur les agents des services publics

Le début des années 2010 va permettre de stabiliser le champ et les concepts des statistiques produites sur l'emploi et les rémunérations.

³² Voir les références juridiques en fin d'article.

³³ Voir les références juridiques en fin d'article.

► **Figure 2 - Les employeurs de la fonction publique dans l'économie nationale**



1. Voir le lien : <https://www.insee.fr/fr/information/2028129>.

2. Les autorités constitutionnelles sont la présidence de la République, le Conseil constitutionnel, le Conseil économique, social et environnemental, l'Assemblée nationale et le Sénat.

3. La métropole de Lyon est une collectivité territoriale à statut particulier. Néanmoins, elle est classée ici avec les autres métropoles au sein des établissements publics.

4. Crous : Centre régional des œuvres universitaires et scolaires, CNRS : Centre national de la recherche scientifique, INRA : Institut national de la recherche agronomique, Inserm : Institut national de la santé et de la recherche médicale.

STATUT DE RECRUTEMENT DE DROIT COMMUN DE L'AGENT

Droit public particulier
ou droit privé à statut

Droit privé

Organismes consulaires, associations syndicales autorisées, associations foncières urbaines ou de remembrement

Caisse nationale du régime général de Sécurité sociale

Groupements d'intérêt public non marchands, groupements de coopération sanitaire, agences départementales d'insertion

Institution de gestion sociale des armées, Institut franco-allemand Saint-Louis, etc.

Cultes d'Alsace-Lorraine

Agence de l'environnement et de la maîtrise de l'énergie, Charbon de France, Centre national d'études spatiales, Institut national de l'audiovisuel, Opéra national de Paris, Réunion des musées nationaux, offices de tourisme, etc.

Régie autonome des transports parisiens (RATP)

Grands ports maritimes, Union des groupements d'achats publics, agences françaises de développement, etc.

- La Poste
- Électricité de France (EDF)
- Aéroports de Paris
- Société nationale des chemins de fer français (SNCF)

Sociétés d'aménagement foncier et d'établissement rural (Safer)

Air France, Orange, France Télévision, Naval Group, GDF Suez, KNDS France⁵, Imprimerie nationale, etc.

Organismes de protection sociale, hôpitaux à but non lucratif, organismes divers d'administration centrale et locale

Mutuelles, comités d'entreprise, organismes professionnels, syndicats de propriétaires, associations et fondations dont le financement n'est pas majoritairement public

5. Y compris la Caisse des dépôts et consignations, assimilée à un ÉPA, mais relevant de la catégorie juridique 74.

6. Industrie de l'armement, autrefois GIAT Industries, puis Nexter.

Lecture : Le périmètre de la fonction publique recouvre les employeurs des catégories juridiques 71 à 73 pour lesquels le statut de recrutement de droit commun est le droit public. Il ne correspond pas strictement à celui des APU. Ainsi, les Offices publics de l'habitat, les Caisses de crédit municipal, etc., sont des ÉPA de la fonction publique, mais ne sont pas des APU.

Tout d'abord, le **périmètre** retenu pour la fonction publique va s'appuyer sur un croisement de la catégorie juridique de l'employeur et du statut de recrutement de droit commun de l'agent public (*figure 2*).

À partir du rapport 2012, le **système d'information sur les agents des services publics** (Siasp), que l'on a déjà brièvement évoqué, permet de produire des informations statistiques sur les trois versants de la fonction publique (*encadré*). Depuis 2022, le Siasp est alimenté principalement par les déclarations sociales nominatives des employeurs (Humbert-Bottin, 2018).

Les **effectifs** sont comptabilisés sur les postes dits « principaux » (au cas où un agent occupe plusieurs emplois), au 31 décembre de l'année décrite, sur le champ France hors Mayotte. L'ensemble des statuts sont suivis à l'exception des bénéficiaires de contrats aidés.

Pour les **rémunérations**, le concept retenu est celui des salaires en équivalent temps plein. Là encore, l'analyse est réalisée au niveau des postes principaux, mais le champ est plus restrictif que pour les emplois : France hors Mayotte, y compris bénéficiaires de contrats aidés, mais hors militaires, assistants maternels et familiaux, apprentis, internes et externes des hôpitaux publics.

Aujourd'hui, il est possible de construire des **séries longues** sur la période 2011-2023, même si l'arrivée de la DSN perturbe la fin des séries.

► Encadré. Le système d'information sur les agents des services publics

Le système d'information sur les agents des services publics (Siasp), actualisé chaque année, recense les données sur l'emploi des agents des trois fonctions publiques, dans une dimension nationale mais aussi territoriale. Le périmètre de la fonction publique renvoie aux personnes morales et organismes soumis au droit administratif dans lequel le recrutement se fait sur la base du droit public. Le tracé de chacun des trois versants s'appuie essentiellement sur la catégorie juridique de l'employeur.

Le Siasp est piloté par l'Insee en collaboration avec l'ensemble des SSM concernés.

Jusqu'au millésime 2021, le Siasp était principalement alimenté par :

- les fichiers mensuels de paie pour la majorité des agents relevant de la fonction publique de l'État,
- le recensement des agents de l'État pour les militaires des armées,
- les déclarations annuelles de données sociales pour les agents relevant des fonctions publiques territoriale et hospitalière, ainsi que les agents d'une partie de la fonction publique de l'État.

Depuis 2017, cependant, une part croissante des données du Siasp provient des déclarations sociales nominatives (DSN), en substitution des autres sources. Enfin, en 2022, conformément au décret n° 2016-611 du 18 mai 2016 fixant les dates limites pour la transmission obligatoire de la déclaration sociale nominative*, la quasi-totalité des employeurs de la fonction publique ont basculé leur système de déclaration en DSN. Les employeurs ont pu revoir à cette occasion leurs systèmes d'information et de déclaration, et donc modifier potentiellement leur comportement déclaratif.

Dans ce contexte de changement de sources d'information, l'Insee a engagé une refonte des traitements statistiques réalisés sur l'emploi et les rémunérations des agents de la fonction publique, et en a ajusté les concepts. Le principal changement conceptuel est la prise en compte de la multi-activité des agents dans les secteurs public et privé. Le poste principal au 31 décembre d'un agent est désormais déterminé parmi les postes occupés au 31 décembre dans le public et le privé. Le champ des études sur l'emploi porte sur les postes principaux au 31 décembre.

* Voir les références juridiques en fin d'article.

► Au-delà du RAEFP

L'autre source souvent utilisée par le SSM FP est l'enquête Emploi de l'Insee, en particulier sur le sujet du temps et de l'organisation du travail : durée annuelle de travail, temps partiel, sous-emploi, horaires atypiques, absence pour raison de santé, télétravail. Le SSM mobilise également de façon plus apériodique deux enquêtes de la Direction de l'animation de la recherche, des études et des statistiques (Dares)³⁴ : l'enquête Conditions de travail et l'enquête Surveillance médicale des expositions des salariés aux risques professionnels.

Outre le rapport annuel, le SSM FP diffuse ses travaux dans deux collections :

- **Stat Rapides**, pour les « marronniers » (publications récurrentes) sur l'emploi public, la formation, les concours de la FPE, les apprentis, la diffusion d'offres d'emploi sur la plateforme « Choisir le service public », la victimisation des agents publics dans le contexte professionnel, les écarts de rémunération entre les femmes et les hommes fonctionnaires des ministères et, enfin, l'indice trimestriel du traitement brut de la grille indiciaire ;
- **Point Stat**, pour les études originales comme celles sur les déterminants de la réussite à certains concours de la fonction publique³⁵, le mode de déplacement entre le domicile et le travail des agents publics, le télétravail, les contractuels de la fonction publique, le profil sociodémographique des agents publics, les souhaits concernant la durée de travail et le changement d'emploi, la mobilité départementale (pour ne citer que les publications de la période 2024-2025).

► Une concertation régulière et encadrée pour définir le programme de travail

Les travaux sont donc riches et variés et le SSM FP se présente souvent comme une « mini-Dares » (« mini » au regard de la taille de son département des études et des statistiques). Cependant, ce qui fait l'originalité de ce SSM, c'est le processus particulier de construction de son programme de travail.

Bien sûr, comme tous les partenaires de la statistique publique, il s'attache à mettre en œuvre le programme de moyen terme du Cnis. Mais il a d'autres lieux de concertation entre les producteurs et les utilisateurs. Nous avons vu, ainsi, qu'il échange avec les organisations syndicales et les employeurs au sein de la formation spécialisée du CCFP sur les questions relatives aux évolutions de l'emploi public, à la politique des retraites dans la fonction publique et à la connaissance statistique de la situation, de la rémunération et des pensions des agents publics.

³⁴ La Dares est le SSM au sein du ministère en charge du travail.

³⁵ Depuis 2023, le SSM FP réalise des enquêtes auprès des candidats à certains concours de la fonction publique. Une présentation du dispositif Base Concours et de l'enquête est accessible sur le site du SSM FP : <https://www.fonction-publique.gouv.fr/etudes-statistiques-et-publications/etudes-statistiques/enquetes-et-sources-statistiques>.

Ces échanges sont l'occasion de recueillir les critiques et demandes de ces utilisateurs attentifs, afin d'orienter les nouveaux champs d'investigation ou les sujets à creuser.

Cette instance se réunit deux fois par an : à l'automne, à l'occasion de la diffusion du rapport annuel, et au premier trimestre de chaque année. Ces échanges sont l'occasion de recueillir les critiques et demandes de ces utilisateurs attentifs, afin d'orienter les nouveaux champs d'investigation ou les sujets à creuser. Les ressources du SSM FP étant limitées, c'est l'occasion de faire exprimer les priorités de ces partenaires. Ainsi, parmi les propositions faites début

2025, les partenaires sociaux et employeurs ont privilégié le sujet de la santé. C'est aussi dans cette instance qu'a été discutée la transformation du rapport annuel en publication statistique, avec les conséquences sur l'accès privilégié et la diffusion de ce rapport.

► Une interaction riche et nécessaire avec des fournisseurs de données dispersés

L'autre département du SSM FP, le département de la donnée et des systèmes d'information RH (SIRH), a lui aussi des instances de concertation. En effet, il organise la collecte des données relatives à la formation et aux recrutements, mais aussi, dans un domaine en construction au sein de la FPE, l'ensemble des données relatives aux Bases de données sociales³⁶ des départements ministériels. Dans ce cadre, il travaille en partenariat avec les ministères, ainsi qu'avec la direction générale de l'offre de soins (DGOS) et le Centre national de gestion (CNG) pour la fonction publique hospitalière, et la direction générale des collectivités locales (DGCL) et la Fédération nationale des centres de gestion (FNCDG) pour la fonction publique territoriale, en particulier sur les sujets liés aux recrutements. Et, bien sûr, il échange avec les organisations syndicales du Conseil supérieur de la fonction publique de l'État pour la Base de données sociales.

Enfin, ce département assure l'animation de **la feuille de route des SIRH de l'État pour la période 2023-2027**. Celle-ci vise à orienter les travaux des missions ministérielles en charge des SIRH. L'ambition est de mutualiser les outils, de partager des référentiels (afin de faciliter à terme la circulation des données entre ministères), de limiter globalement les coûts des systèmes, de partager les bonnes pratiques et de faciliter le pilotage des ressources humaines. Il s'agit de faciliter la phase dite « d'acquisition » des données administratives (ici les données de gestion RH), et de l'encadrer pour limiter le coût des phases suivantes que sont leur « transformation » et leur « traitement statistique » (Koumarianos et Rivière, 2025)³⁷.

³⁶ Une présentation de ce dispositif est accessible sur le site du SSM FP : <https://www.fonction-publique.gouv.fr/etudes-statistiques-et-publications/la-base-de-donnees-sociales-bds>.

³⁷ Voir également l'article de Lamarche et Rivière sur le temps de production des statistiques publiques dans ce numéro.

Les récents travaux menés par l'Insee sur l'usage des données administratives permettront de guider l'action du SSM FP, que ce soit sur la manière de les intégrer dans la production statistique ou de coopérer avec les acteurs pour leur exploitation statistique (*ibid.*). À terme, la convergence des SIRH facilitera la production de statistiques. Elle ouvrira de nouveaux champs d'exploitation... qui permettront de mieux connaître la fonction publique !

► Fondements juridiques

- Loi n° 84-16 du 11 janvier 1984 portant dispositions statutaires relatives à la fonction publique de l'État. Article 93. In : *site de Légifrance*. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse :
<https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000501099>.
- Ordonnance n° 45-2283 du 9 octobre 1945 relative à la formation, au recrutement et au statut de certaines catégories de fonctionnaires et instituant une direction de la fonction publique et un conseil permanent de l'administration civile. Titre V, article 15, alinéa 2. In : *site de Légifrance*. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse : <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000521942>.
- Décret n° 59-210 du 3 février 1959 fixant les attributions de la direction générale de l'administration et de la fonction publique. In : *site de Légifrance*. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse :
<https://www.legifrance.gouv.fr/loda/id/JORFTEXT00000305727/2020-10-08/>.
- Décret n° 76-509 du 10 juin 1976 relatif au Conseil supérieur de la fonction publique. In : *site de Légirance*. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse :
https://www.legifrance.gouv.fr/jorf/id/JORFTEXT00000679175?init=true&page=1&query=76-509&searchField=ALL&tab_selection=all.
- Décret n° 95-1007 du 13 septembre 1995 relatif au comité interministériel pour la réforme de l'État et au Commissariat à la réforme de l'État. In : *site de Légifrance*. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse :
<https://www.legifrance.gouv.fr/loda/id/JORFTEXT00000353908/2021-01-10/>.
- Décret n° 98-573 du 8 juillet 1998 modifiant le décret n° 95-1007 du 13 septembre 1995 relatif au comité interministériel pour la réforme de l'État et au commissariat à la réforme de l'État. In : *site de Légifrance*. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse : <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT00000572099>.
- Décret n° 98-574 du 10 juillet 1998 portant création d'un emploi de directeur, adjoint au directeur général de l'administration et de la fonction publique. In : *site de Légifrance*. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse :
https://www.legifrance.gouv.fr/loda/id/JORFTEXT00000755008?init=true&page=1&query=98-574&searchField=ALL&tab_selection=all.
- Décret n° 2000-663 du 13 juillet 2000 portant création de l'Observatoire de l'emploi public. In : *site de Légifrance*. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse : <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT00000205916>.
- Décret n° 2012-148 du 30 janvier 2012 relatif au Conseil commun de la fonction publique. In : *site de Légifrance*. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse :
<https://www.legifrance.gouv.fr/loda/id/JORFTEXT000025242128>.

- Décret n° 2016-611 du 18 mai 2016 fixant les dates limites pour la transmission obligatoire de la déclaration sociale nominative. In : site de Légifrance. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse :
<https://www.legifrance.gouv.fr/loda/id/JORFTEXT000032542850/2016-09-20/>.
- Décret n° 2016-1804 du 22 décembre 2016 relatif à la direction générale de l'administration et de la fonction publique et à la politique de ressources humaines dans la fonction publique. In : site de Légifrance. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse : <https://www.legifrance.gouv.fr/jorf/id/JORFSCTA000033670487>.
- Décret n° 2022-1459 du 23 novembre 2022 modifiant le décret n° 2016-1804 du 22 décembre 2016 relatif à la direction générale de l'administration et de la fonction publique et à la politique de ressources humaines dans la fonction publique. In : site de Légifrance. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse :
https://www.legifrance.gouv.fr/loda/id/JORFTEXT000046591745?init=true&page=1&query=2022-1459&searchField=ALL&tab_selection=all.
- Arrêté du 16 janvier 2007 relatif à l'organisation de la direction générale de l'administration et de la fonction publique. In : site de Légifrance. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse :
<https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000000463564>.
- Arrêté du 10 avril 2012 relatif à l'organisation de la direction générale de l'administration et de la fonction publique. In : site de Légifrance. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse :
<https://www.legifrance.gouv.fr/loda/id/LEGITEXT000025678071/2012-04-13/>.
- Arrêté du 16 mars 2017 relatif à l'organisation de la direction générale de l'administration et de la fonction publique. In : site de Légifrance. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse :
<https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000034208354>.
- Arrêté du 24 novembre 2022 portant organisation de la direction générale de l'administration et de la fonction publique. In : site de Légifrance. [en ligne]. [Consulté le 29 septembre 2025]. Disponible à l'adresse :
<https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000046600264>.

► Bibliographie

- ANXIONNAZ, Isabelle et MAUREL, Françoise, 2021. Le Conseil national de l'information statistique – La qualité des statistiques passe aussi par la concertation. In : *Courrier des statistiques*. [en ligne]. 8 juillet 2021. Insee. N° N6, pp. 123-142. [Consulté le 17 octobre 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/information/5398693?sommaire=5398695>.
- BALLU, Françoise, 1995. Le bureau des statistiques de la fonction publique. In : *Courrier des statistiques*. [en ligne]. Décembre 1995. Insee. N° 75-76, pp. 33-37. [Consulté le 25 septembre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06z99gk.r=courrier%20des%20statistiques%201995?rk=42918;4#>.
- BALLU, Françoise et MARTEL, Xavier, 1996. Les nomenclatures de corps, grades et emplois dans la Fonction publique d'État. In : *Courrier des statistiques*. [en ligne]. Août 1996. Insee. N° 78, pp. 61-64. [Consulté le 25 septembre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06z998h.r=courrier%20des%20statistiques%201996?rk=64378;0#>.
- BRENOT-OULDALI, Annie, 1996. Les frontières mouvantes de l'emploi public. In : *Courrier des statistiques*. [en ligne]. Août 1996. Insee. N° 78, pp. 55-59. [Consulté le 25 septembre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06z998h.r=courrier%20des%20statistiques%201996?rk=64378;0#>.
- BRENOT-OULDALI, Annie, 2000. Séries longues sur l'emploi dans la fonction publique 1980-1998. In : *RésuStats*. [en ligne]. DGAFFP. [Consulté le 25 septembre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06z998h.r=courrier%20des%20statistiques%20ao%C3%BBt%201996?rk=42918;4#>.
- COTTON, Franck et HAAG, Olivier, 2023. L'intégration des données administratives dans un processus statistique – Industrialiser une phase essentielle. In : *Courrier des statistiques*. [en ligne]. 30 juin 2023. Insee. N° N9, pp. 104-125. [Consulté le 17 octobre 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/information/7635829?sommaire=7635842>.
- DE PERETTI, Gaël, 2024. Un éclairage statistique et réflexif sur les rémunérations dans la fonction publique au prisme du rapport annuel sur l'état de la fonction publique et de ses avatars, séquence « Rémunération et pouvoir d'achat : mise en perspective et pas de côté ». In : *Séminaire « Dialogues autour de la fonction publique : Histoire, sciences sociales et pratiques RH »*. 4 octobre 2024. Sciences Po Paris, Centre d'histoire.
- DE PERETTI, Gaël, 2024. L'emploi public au prisme du rapport annuel sur l'état de la fonction publique. In : *Semaine juridique administration et collectivités territoriales*. 18 novembre 2024. N° 46.
- DIRECTION GENERALE DE L'ADMINISTRATION ET DE LA FONCTION PUBLIQUE, 1975. *Rapport annuel 1975 : principales données statistiques sur la fonction publique*. [en ligne]. [Consulté le 21 octobre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p07j2wr0.r=DGAFP%20rapport%20annuel%201975?rk=21459;2>.

- HUMBERT-BOTTIN, Élisabeth, 2018. La déclaration sociale nominative – Nouvelle référence pour les échanges de données sociales des entreprises vers les administrations. In : *Courrier des statistiques*. [en ligne]. 6 décembre 2018. Insee. N° N1, pp. 25-34. [Consulté le 17 octobre 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/information/3647025?sommaire=3647035>.
- INSEE, 1977. Liste des services statistiques des ministères, des directions régionales et services régionaux de l'Insee et des services statistiques des départements d'outre-mer. In : *Courrier des statistiques*. [en ligne]. Janvier 1977. N° 1, p. 37. [Consulté le 25 septembre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06z980f.r=courrier%20des%20statistiques%201977?rk=21459;2#>.
- INSEE, 1992a. Nominations. In : *Courrier des statistiques*. [en ligne]. Juin 1992. N° 61-62, p. 104. [Consulté le 25 septembre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06z995g.r=courrier%20des%20statistiques%201992?rk=21459;2#>.
- INSEE, 1992b. Supplément Organigrammes des Services statistiques des ministères. In : *Courrier des statistiques*. [en ligne]. Novembre 1992. N° 63. [Consulté le 25 septembre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06z992f.r=courrier%20des%20statistiques%201992?rk=42918;4#>.
- INSEE, 1996. Dossier Emploi public. In : *Courrier des statistiques*. [en ligne]. Août 1996. N° 78, pp. 53-64. [Consulté le 26 septembre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06z998h.r=courrier%20des%20statistiques%20ao%C3%BBt%201996?rk=42918;4#>.
- INSEE, 2001. Le partage des tâches dans le système statistique public français. In : *Courrier des statistiques*. [en ligne]. Septembre 2001. N° 98-99, p. 7. [Consulté le 26 septembre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06xt7d5.r=courrier%20des%20statistiques%202001?rk=42918;4#>.
- INSEE, 2008. La statistique au ministère chargé de la fonction publique : interview de Christine Gonzalez-Demichel. In : *Courrier des statistiques*. [en ligne]. Novembre-décembre 2008. N° 125, pp. 96-98. [Consulté le 26 septembre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06xt47h.r=courrier%20des%20statistiques%202008?rk=64378;0#>.
- KOUMARIANOS, Heidi et RIVIÈRE, Pascal, 2025. Statistiques fondées sur des données administratives : esquisse d'un cadre général. In : *Document de travail*. [en ligne]. 23 juin 2025. Insee. N° M2025-03. [Consulté le 26 septembre 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/statistiques/8596376>.
- LHUILLIER, Hugues. 1978. La statistique au secrétariat d'Etat à la fonction publique. In : *Courrier des statistiques*. [en ligne]. Janvier 1978. Insee. N° 5, pp. 53-64. [Consulté le 26 septembre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06z97k0.r=courrier%20des%20statistiques%20janvier%201978?rk=85837;2>.

- MARTEL, Xavier. 1995. La micro-informatique au bonheur des statisticiens in Le bureau des statistiques de la fonction publique. In : *Courrier des statistiques*. [en ligne]. Décembre 1995. Insee. N° 75-76, encadré 2, p. 36. [Consulté le 26 septembre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06z99gk.r=courrier%20des%20statistiques%201995?rk=42918;4#>.
- MARUANI, Margaret et MERON, Monique, 2012. *Un siècle de travail des femmes en France, 1901-2011*. 18 octobre 2012. La Découverte. 978-2-70714-571-0.
- NEYRET, Guy et PONSOT, Jean-François, 1997. Le dispositif statistique sur les trois fonctions publiques. In : *Rapport de mission*. Cnis. N° 32, janvier 1997. ISBN 2-110-66514-9.
- PONSOT, Jean-François, 1996. Les trois fonctions publiques. In : *Courrier des statistiques*. [en ligne]. Août 1996. Insee. N° 78, pp. 53-54. [Consulté le 26 septembre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06z998h.r=courrier%20des%20statistiques%201996?rk=64378;0#>.
- QUARRÉ, Dominique, 1977. Système d'information sur les agents de l'Etat – opération 1977. In : *Courrier des statistiques*. [en ligne]. Janvier 1977. Insee. N° 1, p. 22. [Consulté le 26 septembre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06z980f.r=courrier%20des%20statistiques%201977?rk=21459;2#>.
- RUIZ, Émilien, 2013. Trop de fonctionnaires ? Contribution à une histoire de l'État par ses effectifs (France, 1850-1950). In : *portail HAL thèses*. [en ligne]. 19 septembre 2013. École des hautes études en sciences sociales (EHESS). [Consulté le 26 septembre 2025]. Disponible à l'adresse : <https://theses.hal.science/tel-00863780>.
- RUIZ, Émilien, 2021. *Trop de fonctionnaires ? Histoire d'une obsession française (XIX^e - XXI^e siècle)*. 29 septembre 2021. Fayard. L'épreuve de l'histoire. ISBN 978-2-21372-049-4.
- SECRÉTARIAT GÉNÉRAL DU GOUVERNEMENT (SGG), mission des archives. 2024. Archives des services du Premier Ministre, Administration et fonction publique, État thématique des versements conservés aux Archives nationales.
- VOLLE, Michel, 1977. Pourquoi faire l'Histoire de la Statistique ? In : *Courrier des statistiques*. [en ligne]. Janvier 1977. Insee. N° 1, pp. 31-32. [Consulté le 26 septembre 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06z980f.r=courrier%20des%20statistiques%201977?rk=21459;2#>.
- ZAIDMAN, Catherine, 2007. Emploi public en France : état des lieux et perspectives d'évolution. In : *Modernisation des services publics et management social en France et en Allemagne – Sous la direction de Leo Kißler, René Lasserre et Marie-Hélène Pautrat*. [en ligne]. Cirac. Pp. 65-76. [Consulté le 26 septembre 2025]. Disponible à l'adresse : <https://books.openedition.org/cirac/662>.

Prophyl : de la parcelle à la statistique, un projet pour simplifier la récolte de données agricoles



Pauline Cazaban*, François Chevalier** et Vincent Marcus***

Le projet Prophyl, porté par le service de la statistique et de la prospective (SSP) du ministère de l'Agriculture, de l'Agro-alimentaire et de la Souveraineté alimentaire, constitue une avancée prometteuse pour la modernisation de la statistique agricole. Pour les cultivateurs disposant d'un « logiciel de gestion parcellaire », l'idée est de mobiliser les informations qu'ils y saisissent sur les traitements phytosanitaires (pesticides) utilisés sur leurs cultures. Ainsi, il est aujourd'hui envisageable de remplacer partiellement la collecte classique de telles données, via des enquêteurs, par une collecte automatisée, moins contraignante pour les agriculteurs.

L'expérimentation conduite en 2024 auprès d'une centaine d'exploitations a permis de recueillir des informations sur près de 5 000 traitements phytosanitaires. Les premiers résultats sont cohérents avec ceux des enquêtes « pratiques culturales » classiques, tout en étant plus précis sur le calendrier et les modalités des traitements. Ce nouveau mode de collecte repose sur une chaîne technique sécurisée, garantissant la qualité des données et le respect de la confidentialité pour les exploitants.

Prophyl préfigure ainsi l'enquête Phyto 2026, qui intégrera ce dispositif à grande échelle. Au-delà du seul thème des produits phytosanitaires, il pourrait ouvrir la voie à une statistique agricole plus « connectée », mieux articulée avec les pratiques numériques des acteurs agricoles, et répondant aux besoins de pilotage des politiques publiques.

 The Prophyl project, led by the Statistics and Foresight Department (SSP) of the Ministry of Agriculture, Agri-food and Food Sovereignty (MASA), represents a promising advancement for the modernisation of agricultural statistics. Farmers equipped with a farm management information system can enter data in relation to plant protection treatments (pesticides) they have been using on their plantation. This data would allow statisticians to partially replace the traditional collection of similar data by an automated one, which would be less burdensome for farmers.

The 2024 pilot involving over a hundred farms enabled the collection of information on approximately 5,000 phytosanitary treatments. Early results show strong consistency with traditional “cultivation practices” surveys, while providing increased precision regarding the timing and modalities of treatments. This new collection method relies on a secure technical infrastructure that ensures data quality and the confidentiality of farmers.

Prophyl thus paves the way for the 2026 Phyto survey, which will integrate this system on a larger scale. Beyond the scope of phytosanitary products, it could open the door to more “connected” agricultural statistics, more coherent with field digital practices, and fulfilling the need for public policies steering.

* Cheffe du bureau des statistiques végétales, animales et environnementales, service de la statistique et de la prospective.
pauline.cazaban@agriculture.gouv.fr

** Sous-directeur des statistiques agricoles, forestières et agroalimentaires, service de la statistique et de la prospective.
francois.chevalier@agriculture.gouv.fr

*** Chef du service de la statistique et de la prospective.
vincent.marcus@agriculture.gouv.fr

Aujourd’hui, de plus en plus d’agriculteurs s’équipent de « logiciels de gestion parcellaire » pour piloter leur exploitation et simplifier leurs démarches administratives (voir infra). Ils y enregistrent notamment toutes les informations relatives aux traitements qu’ils ont réalisés pour protéger leurs cultures.

Et si ces données déjà saisies devenaient, sans effort supplémentaire pour les agriculteurs, une source fiable pour les statistiques nationales ? Ce qui relevait hier de l’utopie pourrait devenir une perspective sérieuse avec le projet Prophyl¹, porté par le service de la statistique et de la prospective (SSP) du ministère de l’Agriculture, de l’Agro-alimentaire et de la Souveraineté alimentaire.

La mesure de l’utilisation des **produits phytosanitaires**, plus communément appelés pesticides (herbicides, fongicides, insecticides, bactéricides, etc.)², est un enjeu central pour la santé publique, la protection de l’environnement et l’évaluation des politiques agricoles³. Le règlement européen SAIO (*Statistics on agricultural inputs and outputs*)⁴ impose à tous les États membres, pour 21 cultures identifiées⁵, de remonter des statistiques sur les quantités de substances actives⁶ utilisées et sur les surfaces traitées lors de la campagne agricole de 2026. Cette remontée deviendra annuelle à compter de la campagne de 2028. Dans ce contexte, une idée apparemment simple et séduisante a germé : plutôt que d’interroger les exploitants via des enquêtes classiques, pourquoi ne pas mobiliser les données déjà saisies dans leurs outils numériques lorsqu’elles existent ?

Initiée en 2019, la démarche inspirée par ce questionnement explore de nouvelles voies pour concilier modernisation statistique, allègement de la charge de réponse des agriculteurs et qualité des données.

► Une alternative crédible à l’enquête classique ?

Produire une information statistique fiable et solide sur l’usage des produits phytosanitaires constitue un exercice complexe et rigoureux. Jusqu’à présent, le SSP s’appuyait principalement sur les enquêtes « **pratiques culturales** » (Espinosa, 2023 ; Firquet et Montcoudiol, 2024). Ces dernières sont réalisées auprès des agriculteurs, dans le cadre d’entretiens en face-à-face, par les enquêteurs des services statistiques des Directions régionales de l’alimentation, de l’agriculture et de la forêt (Draaf). Elles visent à décrire les interventions effectuées par les agriculteurs sur leurs cultures (travail du sol, irrigation, semences, engrains et traitements), dont les traitements phytosanitaires. Elles couvrent successivement les quatre grandes filières de production végétale en France (grandes cultures, viticulture, arboriculture et cultures légumières) et sont réalisées à peu près une fois tous les cinq ans.

¹ « Prophyl » pour « PROduits PHYtosanitaires dans les Logiciels ».

² On parle aussi de « produits phytopharmaceutiques » (en abrégé « PPP »).

³ Voir le lien vers la page du site VizAgreste consacrée aux traitements phytosanitaires : <https://vizagreste.agriculture.gouv.fr/les-traitements-phytosanitaires.html>.

⁴ Règlement (UE) n° 2022/2379 relatif aux statistiques sur les intrants et les produits agricoles. Voir les références juridiques en fin d’article.

⁵ Blé tendre, blé dur, orge, maïs fourrage, maïs grain, colza, tournesol, pommes de terre, betteraves sucrières, olives, raisins de table, épeautre, oranges, choux, carottes, oignons, tomates, fraises.

⁶ Il s’agit des substances (molécules) qui confèrent au traitement ses propriétés préventives contre les organismes nuisibles.

“ Produire une information statistique fiable et solide sur l'usage des produits phytosanitaires constitue un exercice complexe et rigoureux.”

La qualité méthodologique de ces enquêtes est largement reconnue, mais elles présentent plusieurs limites : leur coût élevé (variant de 800 000 euros à plus de 2 millions d'euros selon la filière), leur lourdeur pour les agriculteurs (plus d'une heure d'interview), et une fréquence ainsi qu'une précision géographique (anciennes régions) qui ne sont plus complètement en phase avec les attentes. En effet, les besoins en information évoluent rapidement, portés par une pression sociétale croissante. Ainsi, les citoyens, les organisations non gouvernementales

et les élus appellent à plus de transparence sur l'usage des pesticides. Quant aux exploitants agricoles, ils expriment également le besoin de disposer de données plus fréquentes, leur permettant de se situer par rapport à leurs pairs et d'ajuster leurs pratiques. Ces attentes convergentes sont pleinement prises en compte par la nouvelle stratégie gouvernementale Écophyto 2030 ([encadré 1](#)).

Dans ce contexte, le SSP a engagé en 2019 une réflexion sur le recours à de nouvelles sources de données, de nature privée, potentiellement plus riches et mieux intégrées aux pratiques numériques désormais courantes dans le monde agricole. Par rapport aux autres données privées explorées jusqu'à présent par la statistique publique (Lesur, 2025), celles issues des logiciels de gestion parcellaire ont une particularité : ce ne sont pas uniquement des « traces numériques » dérivées de l'activité des individus (comme les données de téléphonie mobile ou les données de caisse), mais bien d'abord des traces numériques laissées par les individus pour leurs propres besoins (en l'occurrence, ici, le pilotage des cultures agricoles). Elles sont de ce fait, en quelque sorte, « doublement privées » : la statistique publique ne peut y accéder qu'en sollicitant directement les individus qui en sont à l'origine, et non en masse, par le biais d'un acteur économique centralisateur (comme un opérateur de téléphonie ou une enseigne de la grande distribution).

► Encadré 1. Les enjeux d'Écophyto 2030

La [stratégie gouvernementale Écophyto 2030*](#) s'inscrit dans la continuité des plans d'actions Écophyto I, II et II+. Elle poursuit l'objectif d'une agriculture durable par la réduction significative de l'usage des produits phytosanitaires, tout en garantissant la compétitivité agricole et la sécurité alimentaire. Son ambition est triple :

- préserver la santé publique et celle de l'environnement dans une logique « Une seule santé » ;
- soutenir les performances économique et environnementale des exploitations ;
- maintenir un haut niveau de protection des cultures par une adaptation des techniques utilisées.

La stratégie fixe ainsi des objectifs de réduction des risques et des usages de produits phytosanitaires cohérents avec les engagements de la France, aux niveaux européen et international, en matière de lutte contre le dérèglement climatique et de

préservation de la biodiversité. En même temps, elle donne à tous les agriculteurs les moyens de cette transition, par le développement de méthodes alternatives et le renforcement de leur accompagnement dans le changement de pratiques.

La stratégie comporte un volet « connaissance des usages » et souligne la nécessité de disposer d'indicateurs fiables et pertinents, permettant de mesurer régulièrement les progrès et le respect de la trajectoire et des objectifs. Elle mentionne explicitement le rôle des enquêtes statistiques sur ce sujet pour suivre les usages dans les différentes filières (mesure 5.3.3).

La mise en place d'un système d'observation robuste, reposant sur des données actualisées et comparables, constitue ainsi un pilier indispensable de la stratégie Écophyto 2030, assurant que les efforts déployés se traduisent par des résultats mesurables.

* Voir le lien : <https://agriculture.gouv.fr/strategie-ecophyto-2030>.

► Une innovation au croisement des technologies et des usages

Prophyl repose sur une idée à la fois simple et ambitieuse : de nombreux agriculteurs enregistrent déjà leurs traitements phytosanitaires dans leur **logiciel de gestion parcellaire**. Ces outils sont utilisés comme support d'enregistrement régulier d'une multitude de données sur le climat, le sol, et les pratiques agronomiques telles que la gestion de la fertilisation ou les traitements phytosanitaires. Ils permettent aux agriculteurs de disposer d'outils de traçabilité de leurs cultures, de mesurer des indicateurs agro-environnementaux, ou encore de simplifier leur démarche de demande d'aides au titre de la Politique agricole commune (PAC).

“ Prophyl repose sur une idée à la fois simple et ambitieuse : de nombreux agriculteurs enregistrent déjà leurs traitements phytosanitaires dans leur logiciel de gestion parcellaire.

Le recours aux logiciels de gestion parcellaire se diffuse largement dans le milieu agricole. Selon les résultats de l'enquête Pratiques culturelles en grandes cultures, 46 % des exploitations cultivant du blé tendre et 68 % de celles cultivant de la betterave sucrière utilisaient un logiciel de gestion parcellaire en 2021. Ce recours est voué à s'intensifier, dans un contexte où la réglementation européenne impose, à compter de janvier 2027, l'enregistrement par les agriculteurs des traitements phytosanitaires dans un format électronique, lisible par machine, en lieu et place des registres papier utilisés actuellement. Dès lors, pourquoi ne pas valoriser les données des logiciels de gestion parcellaire, saisies à des fins de gestion, pour en faire également une source pour la production statistique ?

Une telle approche est a priori gagnant-gagnant, et s'inscrit pleinement dans les principes du **Code de bonnes pratiques de la statistique européenne**⁷. Pour les exploitants, plus besoin de répondre à une enquête dédiée : les données qu'ils ont déjà saisies dans le cadre de la gestion de leur exploitation sont réutilisées, ce qui allège leur charge administrative. Pour les statisticiens, c'est l'accès à des informations plus fréquentes, plus fines, potentiellement exhaustives.

Cette idée séduisante de simplicité ne va pourtant pas de soi en pratique. Des démarches en ce sens ont déjà été entreprises dans d'autres pays européens :

- La Suède expérimente ce mode de collecte depuis une petite dizaine d'années pour ses enquêtes sur les pratiques de fertilisation (Rangel et al., 2019). L'agriculteur enquêté a la possibilité d'importer un fichier de données extrait de son logiciel de gestion parcellaire dans l'application web de saisie « classique » de réponse à l'enquête. Le taux de recours à cette fonctionnalité a été très modeste au début, se situant à 5 % des répondants en 2016 et 2017, et n'a que faiblement augmenté depuis.
- Les Pays-Bas ont conduit une expérimentation avec l'entreprise John Deere⁸ pour mobiliser les données collectées par les machines de leur marque lors des travaux

⁷ Voir le lien : <https://www.insee.fr/fr/information/4140105>.

⁸ Entreprise américaine spécialisée dans la fabrication de matériel agricole.

**Pour les exploitants,
plus besoin de répondre
à une enquête. Pour les
statisticiens, c'est l'accès
à des informations
plus fréquentes, plus
fines, potentiellement
exhaustives.**

agricoles (Snijkers et al., 2022 ; Snijkers et al., 2025). Elle s'est avérée peu concluante : pour des raisons techniques (difficulté de gestion de l'accord des exploitants), de qualité des données (beaucoup d'agriculteurs s'étonnent des données issues des machines et les corrigent) et finalement « statistiques » (les agriculteurs n'utilisant pas toujours exclusivement des machines John Deere sur leur exploitation).

Pour conduire une expérimentation inspirée des mêmes idées en France, il a fallu d'abord instaurer

un cadre de dialogue et de coopération étroite entre acteurs publics, privés et agricoles (**figure 1**). Ainsi, un groupe de travail a été constitué en 2021 sous l'égide du Conseil national de l'information statistique (Cnis)⁹, piloté par Mehdi Siné, directeur de l'Acta¹⁰, et par le SSP (Siné et al., 2025 ; Cravatte, 2025). Y participent :

- les éditeurs de logiciels, en charge de l'interface technique ;
- Agdatahub¹¹, pour la gestion du consentement des agriculteurs ;
- le service du numérique du ministère de l'Agriculture, de l'Agro-alimentaire et de la Souveraineté alimentaire, chargé de sécuriser les flux de données ;
- les chambres d'agriculture et les syndicats professionnels, assurant l'ancrage territorial et le relais auprès des acteurs ;
- les directions métiers du ministère (DGPE, DGAL), ainsi que les instituts scientifiques (INRAE, Acta...), pour l'expertise de fond¹².

Après les premiers échanges informels démarrés dès 2019 avec la profession agricole, ce cadre institutionnel a été indispensable pour instaurer de la confiance, convaincre tous les partenaires de l'intérêt de la démarche et des opportunités que cela pouvait représenter à terme, et lever les blocages et les réticences. En particulier, la question du statut de ces données et de la confidentialité dont elles pouvaient bénéficier a été un sujet sensible. L'investigation juridique qui s'en est suivie a montré que, s'agissant « d'émissions de substances dans l'environnement », le secret statistique ne pourrait être opposé aux dispositions du code de l'environnement¹³. En effet, ce dernier rend communicable à toute personne qui en fait la demande les informations sur ces substances détenues par une autorité publique¹⁴. Il a donc fallu apporter des garanties sur le fait que les données collectées à l'issue de l'expérimentation seraient rapidement détruites.

⁹ Le Cnis assure la concertation entre les producteurs et les utilisateurs de la statistique publique (voir Anxionnaz et Maurel, 2021).

¹⁰ L'Acta est la tête de réseau des instituts techniques agricoles. Voir le lien : <https://agriculture.gouv.fr/les-instituts-techniques-agricoles-acteurs-majeurs-de-linnovation-agricole>.

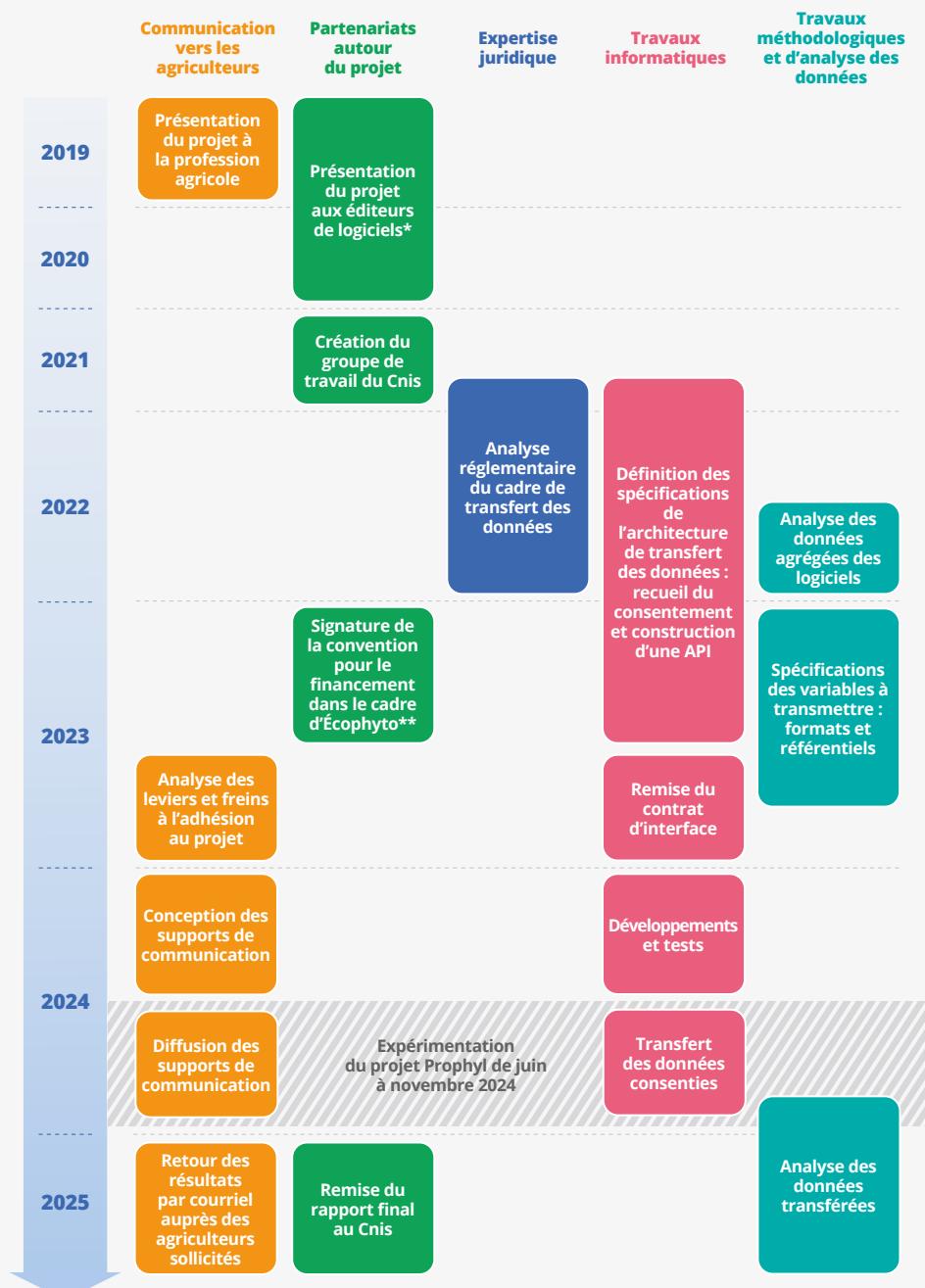
¹¹ Société ayant créé une plateforme de partage de données agricoles.

¹² DGPE : direction générale de la performance économique et environnementale des entreprises ; DGAL : direction générale de l'alimentation ; INRAE : Institut national de recherche pour l'agriculture, l'alimentation et l'environnement.

¹³ Voir les références juridiques en fin d'article.

¹⁴ Ce point est précisé dans le paragraphe II de l'article L. 124-5 du code de l'environnement. Il a été introduit dans le code au titre de la transposition dans le droit français de la directive européenne n° 2003/4/CE concernant l'accès du public à l'information en matière d'environnement. Voir les références juridiques en fin d'article.

► Figure 1 - Chronologie des étapes clés du projet Prophyl



Cnis : Conseil national de l'information statistique.

API : Application Programming Interface ou interface de programmation d'application.

* À la suite de cette présentation, trois éditeurs de logiciels ont manifesté leur intérêt : Chambres d'agriculture France, Isagri et Smag.

** Écophyto 2030 est une stratégie gouvernementale ayant pour objectif le développement d'une agriculture durable ([encadré 1](#)).

► Une première expérimentation encourageante en 2024 —

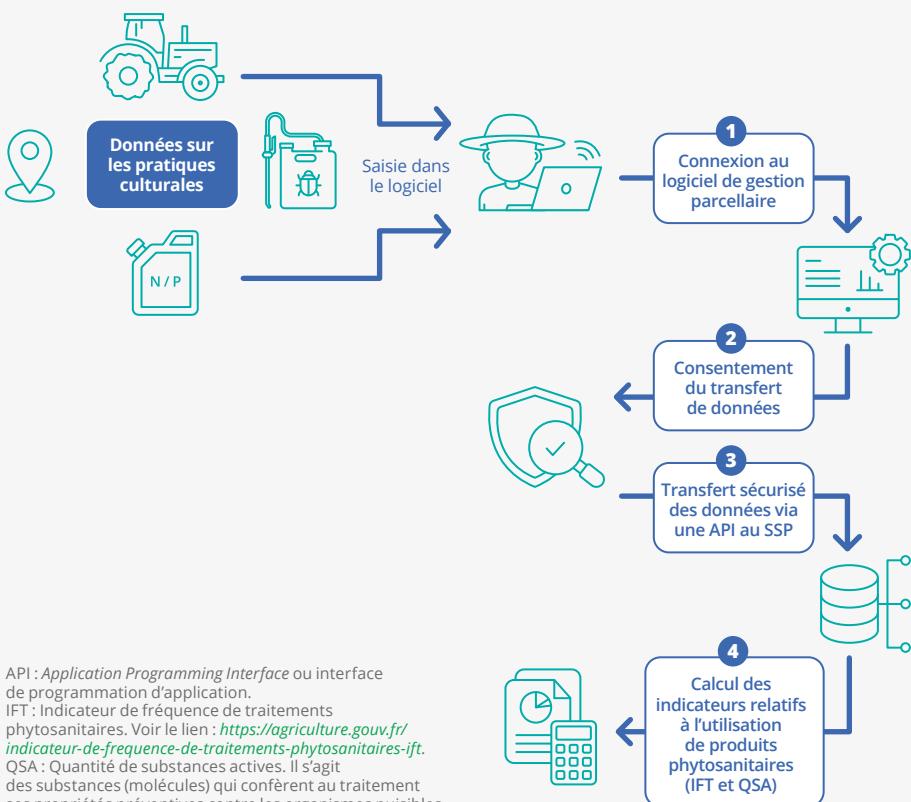
Au terme d'un assez long chemin parsemé d'étapes informatiques, administratives et organisationnelles, l'année 2024 a vu finalement la concrétisation de ce projet, avec la mise en œuvre d'une expérimentation grandeur nature (*figure 2*) rendue

possible grâce au financement du plan Écophyto. Celle-ci s'est déroulée entre juin et novembre 2024. Elle a été conduite en partenariat avec trois éditeurs majeurs de logiciels de gestion parcellaire : Isagri, Smag et Chambres d'agriculture France. Au total, 116 exploitations agricoles ont accepté de transmettre leurs données dans ce cadre. Ces dernières portent sur près de 5 000 traitements phytosanitaires réalisés au cours des campagnes agricoles de 2023 et 2024.

*Au terme
d'un assez long
chemin, l'année 2024
a vu finalement
la concrétisation
de ce projet.*

L'analyse des données recueillies a montré d'abord que les traitements phytosanitaires sont inégalement renseignés. Pour près d'un répondant sur deux, ils ne sont renseignés que pour un peu plus de la moitié des parcelles, ce qui

► **Figure 2 - Du champ à la statistique : modus operandi de l'expérimentation Prophyl**



est bien en deçà de ce que l'on connaît des pratiques culturales. Les raisons de cet état de fait n'ont pas été totalement élucidées :

- Dans certains cas, cela peut traduire le caractère différé de la saisie. En effet, certains exploitants reportent leurs données dans le logiciel avec un certain délai, parfois plusieurs mois après l'application des traitements sur les parcelles.
- Les éditeurs de logiciels indiquent également que les exploitants priorisent souvent la saisie des traitements phytosanitaires pour les parcelles gérées dans le cadre d'un contrat de production, lorsque l'itinéraire cultural fait partie intégrante de l'accord¹⁵. Les saisies peuvent même parfois être gérées dans le cadre d'une prestation de service proposée par la coopérative. Tout ceci explique pourquoi, à un instant donné, certaines parcelles d'une même exploitation peuvent avoir des traitements saisis, tandis que d'autres n'en ont pas.

Les données renseignées disponibles ont permis toutefois de reconstituer des calendriers phytosanitaires détaillés pour les campagnes agricoles de 2023 et 2024. Sur le blé d'hiver, par exemple, les doses observées s'avèrent cohérentes avec les références agronomiques. Pour cette même culture, la **quantité de substances actives** (QSA) utilisées, extrapolée au niveau national, est comparable à celle de l'enquête Pratiques

► Figure 3 - Estimation des quantités de substances actives (QSA)

		en kg	
		Expérimentation Prophyl*	
		Campagne 2023	Campagne 2024
Pour la culture de blé tendre			
Fongicides et bactéricides	3 276 599	2 370 088	2 745 847
Herbicides	8 655 327	9 211 077	8 000 658
Insecticides et acaricides	38 963	16 419	45 096
Molluscicides	40 141	67 600	27 656
Régulateur de croissance	1 957 225	1 350 662	1 508 882
Pour la culture de maïs grain			
Fongicides et bactéricides	Non renseigné	42 176	1 853
Herbicides	Non renseigné	1 062 325	1 681 917
Insecticides et acaricides	Non renseigné	22 099	34 960
Molluscicides	Non renseigné	37 629	24 804

* L'expérimentation Prophyl, réalisée de juin à novembre 2024, a permis de collecter des données sur les campagnes agricoles de 2023 et de 2024. Pour la campagne de 2023, les données n'ont pas permis d'estimer les QSA relatives à la culture de maïs grain.

Source : Service de la statistique et de la prospective (SSP), enquête Pratiques culturales en grandes cultures 2021 et expérimentation Prophyl 2024.

¹⁵ Par « parcelles gérées dans le cadre d'un contrat de production », on entend les surfaces pour lesquelles l'exploitant agricole s'engage, via un accord commercial avec un acheteur (coopérative, négociant, industriel), à suivre un itinéraire cultural défini à l'avance. Ce contrat peut préciser les pratiques à appliquer (choix des variétés, calendrier d'interventions, produits phytosanitaires autorisés, exigences de traçabilité), en échange d'un engagement d'achat ou d'un prix garanti.

culturales en grandes cultures de 2021 (*figure 3*), et ce malgré la taille très limitée de l'échantillon¹⁶.

Le calcul de l'**indicateur de fréquence de traitements phytosanitaires** (IFT)¹⁷, utilisé dans le cadre du suivi du plan Écophyto, s'est révélé plus délicat. En effet, il requiert de connaître la « cible » visée pour chaque produit appliqué, c'est-à-dire les ravageurs ou maladies des cultures contre lesquels le produit phytosanitaire est dirigé. Cette information, dont l'enregistrement n'est pas exigé par la réglementation phytosanitaire, est souvent absente des données issues des logiciels. Néanmoins, un calcul de l'IFT reste possible en l'absence du renseignement de la cible. Bien que légèrement surestimé, il aboutit à des résultats globalement comparables à ceux obtenus par voie d'enquête (*figure 4*).

► **Figure 4 - Estimation des indicateurs de fréquence de traitements phytosanitaires (IFT)**

	Expérimentation Prophyl*		Enquête Pratiques culturelles en grandes cultures 2021
	Campagne 2023	Campagne 2024	
Herbicides	4,8	4,2	4,1
Insecticides et acaricides	Non renseigné	2,2	1,6

* L'expérimentation Prophyl, réalisée de juin à novembre 2024, a permis de collecter des données sur les campagnes agricoles de 2023 et de 2024. Pour la campagne de 2023, les données n'ont pas permis d'estimer les IFT relatifs à la culture de maïs grain.

Source : Service de la statistique et de la prospective (SSP), enquête Pratiques culturelles en grandes cultures 2021 et expérimentation Prophyl 2024.

L'expérimentation a mis en évidence plusieurs défis techniques.

L'expérimentation a également mis en évidence plusieurs défis techniques : la mise en place d'une architecture informatique dédiée (*encadré 2*), l'appariement des données issues des logiciels avec les fichiers administratifs issus de la PAC, l'harmonisation des formats de fichiers hétérogènes, et la garantie de l'anonymat des exploitants. Ces défis ont été relevés grâce à un dispositif combinant :

- une clé d'appariement fondée sur le **numéro Siret**¹⁸,
- un référentiel commun automatisé (pour standardiser les formats de fichiers),
- un mécanisme de recueil du consentement explicite de la part des agriculteurs.

16 Ces résultats, ainsi que ceux de la figure 4, sont à prendre avec prudence du fait du faible nombre d'observations récupérées dans l'expérimentation Prophyl.

17 Voir le lien : <https://agriculture.gouv.fr/indicateur-de-fréquence-de-traitements-phytosanitaires-ift>.

18 Voir le lien : <https://www.insee.fr/fr/metadonnees/definition/c1841>.

► Encadré 2. Une architecture informatique au service de la fiabilité statistique

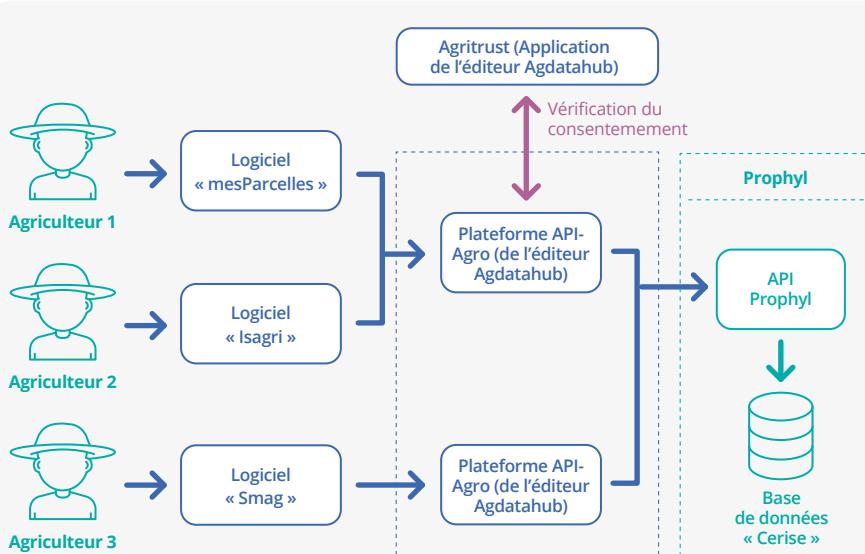
Le dispositif Prophyl, déployé lors de l'expérimentation, s'appuie sur une chaîne sécurisée en sept étapes (*figure encadré 2*) :

1. L'exploitant saisit ses traitements phytosanitaires dans son logiciel habituel.
2. L'exploitant volontaire pour prendre part à l'expérimentation Prophyl enregistre son consentement au transfert de données dans l'application Agritrust du tiers de confiance Agdatahub.
3. Agritrust vérifie l'identité numérique de l'exploitant en interrogeant le Registre national des entreprises, et enregistre le consentement le cas échéant.
4. Les données consenties sont transmises à la plateforme API-Agro du tiers de confiance Agdatahub.
5. Depuis cette plateforme, les données anonymisées sont transmises exclusivement au SSP via une interface de programmation applicative* dédiée à l'expérimentation Prophyl.
6. Le SSP reçoit uniquement des données pseudonymisées**, prêtes à être analysées.

7. Des contrôles de qualité permettent de détecter les anomalies et garantissent la robustesse des indicateurs produits.

Afin d'assurer une parfaite interopérabilité entre les différents logiciels, l'ensemble du dispositif Prophyl repose sur des standards existants : nomenclatures des substances actives, types de traitements, cultures, unités de surface, etc. Ainsi, Prophyl utilise le même modèle de données que l'*atelier de calcul de l'IFT**** du ministère de l'Agriculture, de l'Agro-alimentaire et de la Souveraineté alimentaire. Cet atelier de calcul est un outil en ligne qui permet aux agriculteurs d'évaluer leur usage des produits phytosanitaires au niveau de la parcelle ou de l'exploitation.

Globalement, les choix retenus au niveau de l'infrastructure et du processus ont visé à concilier l'efficacité technique et le respect des exigences du règlement général sur la protection des données (RGPD). L'expérimentation a cependant révélé que les étapes 2 et 3, via les outils de Agdatahub, étaient souvent rédhibitoires et conduisaient le plus souvent l'agriculteur à renoncer. À l'inverse, la solution mise en œuvre par Smag, plus simple et moins exigeante en matière de contrôle d'identité, s'est révélée beaucoup plus efficace.



API : Application Programming Interface ou interface de programmation d'application.

Lecture : Pour les utilisateurs de mesParcelles et d'Isagri, le consentement au transfert des données ainsi que leur pseudonymisation sont gérés par Agritrust, une application d'Agdatahub. Le logiciel Smag, quant à lui, gère ces étapes de manière interne. Sur la pseudonymisation, voir le lien suivant : https://www.edpb.europa.eu/sme-data-protection-guide/faq-frequently-asked-questions/answer/what-difference-between_fr.

* Ou API pour application programming interface en anglais. Il s'agit d'une interface logicielle qui permet de « connecter » un logiciel ou un service à un autre logiciel ou service, afin d'échanger des données et des fonctionnalités.

** Voir le lien suivant : https://www.edpb.europa.eu/sme-data-protection-guide/faq-frequently-asked-questions/answer/what-difference-between_fr.

*** Voir le lien suivant : <https://ecoagri.agriculture.gouv.fr/ift/>.

Pour ce dernier, la solution technique proposée par Agdatahub, utilisée par deux des trois éditeurs de logiciels, s'est révélée trop ambitieuse et exigeante, et au final inopérante. En effet, la gestion du consentement par cette société consistait non seulement à demander à l'agriculteur de donner son accord (ce qui va de soi), mais aussi à vérifier au préalable son identité numérique comme chef d'exploitation, vérification qui reposait sur une interrogation du Registre national des entreprises¹⁹. Dans les faits, cette procédure d'identification numérique était loin d'être simple : elle supposait que l'agriculteur utilise son ordinateur pour obtenir un QR code, qu'il devait ensuite scanner avec son smartphone, puis qu'il suive un parcours en ligne souvent long et fastidieux. Beaucoup d'exploitants ont ainsi rencontré des difficultés à aller au bout du processus. Au final, les données recueillies sur les 116 exploitations l'ont été essentiellement via le troisième éditeur. Sans passer par la solution Agdatahub, ce dernier avait développé dans son logiciel de gestion parcellaire un module de consentement plus léger, facilitant ainsi la participation des exploitants déjà familiers de cet outil.

► Une communication ciblée pour mobiliser et fédérer —

Les conditions de réussite de l'expérimentation ne se jouaient pas uniquement sur le terrain technique. Elles reposaient aussi sur un effort important de communication, conçu pour susciter la confiance et faciliter l'adhésion des agriculteurs. Dès la préparation du test, un plan de communication structuré a été mis en place, autour d'une signature graphique et visuelle dédiée créée pour Prophyl²⁰. Le plan visait à expliquer le dispositif,

à répondre aux interrogations et à simplifier la démarche de participation. Il reposait sur trois outils du site internet du SSP « Agreste »²¹ :

- une foire aux questions détaillée, apportant des réponses précises aux questions techniques, juridiques et méthodologiques les plus fréquentes ;
- une vidéo pédagogique, présentant en quelques minutes le fonctionnement de Prophyl, ses bénéfices pour les exploitants et les garanties offertes en matière de confidentialité ;
- un guide opérationnel (modus operandi) décrivant pas à pas la procédure de recueil du consentement et de transmission sécurisée des données depuis le logiciel de gestion parcellaire.

La mobilisation de relais de proximité, tels que les techniciens des chambres d'agriculture, avait été initialement prévue, mais n'a pas pu être déployée dans le contexte du mouvement social agricole durant l'année 2024. Parmi les partenaires, seul Smag a assuré une communication active auprès de ses utilisateurs, via notamment l'organisation d'un webinaire d'information dédié vers les distributeurs du logiciel Smag Farmer. Cette situation a permis de mesurer à quel point l'appui des

¹⁹ Ce registre a pour objectif de centraliser et de diffuser les informations économiques et juridiques des entreprises. Voir le lien : <https://www.inpi.fr/ressources/formalites-dentreprises/registre-national-entreprises>.

²⁰ Voir par exemple le lien vers la vidéo dédiée au projet : <https://www.youtube.com/watch?v=rVTypzaLdfg>.

²¹ Voir le lien : <https://agreste.agriculture.gouv.fr/agreste-web/>.

réseaux ancrés localement sera déterminant pour les prochaines étapes. L'expérience a confirmé que la communication n'est pas un simple accompagnement, mais un levier stratégique pour élargir la participation, garantir une meilleure représentativité et renforcer la confiance.

► Enquête Phyto 2026 : vers une mise à l'échelle de Prophyl —

“ Prophyl sera intégré dans le dispositif de l'enquête Phyto 2026.

Sur la base des résultats obtenus et de l'analyse des écueils rencontrés, le SSP a décidé de sauter le pas : Prophyl sera intégré dans le dispositif de l'enquête Phyto 2026 ([encadré 3](#)), en tant que mode de collecte alternatif à la saisie sur interface web. Contrairement à d'autres données privées auxquelles on accède en masse auprès d'opérateurs centralisateurs

(Lesur, 2025), il s'agit bien ici de solliciter l'accord des personnes physiques détentrices de ces données. Les exploitants de grandes cultures disposant d'un logiciel de gestion parcellaire compatible pourront ainsi répondre via Prophyl, tandis que les autres renseigneront un questionnaire en ligne.

► Encadré 3. Phyto 2026 : une enquête pour le suivi des usages des produits phytosanitaires en agriculture

L'enquête Phyto 2026 a pour objectif de décrire les pratiques d'utilisation des pesticides au cours de la campagne culturelle de 2026. Elle fournira des données détaillées sur les quantités de substances actives utilisées, ainsi que sur les superficies traitées, pour 21 cultures issues de différentes filières agricoles. Elle distinguerà par ailleurs les modes de production biologique et conventionnel. À partir de la campagne de 2028, elle sera réalisée à un rythme annuel, permettant une mise à jour plus régulière des données relatives à l'utilisation des pesticides en France.

Cette enquête répond aux exigences du règlement européen SAIO en matière de production de statistiques sur l'usage des produits phytosanitaires. Si la priorité de la Commission européenne porte sur les résultats au niveau national, des statistiques régionales sont susceptibles d'être produites pour l'agriculture conventionnelle. Ces dernières permettront de répondre aux besoins des utilisateurs, notamment pour le suivi des politiques publiques environnementales et agricoles à l'échelle territoriale (stratégie Écophyto 2030).

Le questionnaire de l'enquête se compose de trois volets principaux :

- identification de l'exploitation et du répondant, avec les informations administratives et les coordonnées ;
- caractéristiques de la parcelle enquêtée, précisant la culture, le mode de production, la localisation,

la surface, ainsi que les références cadastrales ou de la PAC ;

- pratiques phytosanitaires, visant à reconstituer l'ensemble des traitements réalisés durant la campagne de 2026, avec pour chaque intervention : la date, le nom du produit, son numéro d'autorisation, la cible, la dose appliquée, la quantité de bouillie* et la proportion de surface traitée.

La collecte des données se déroulera d'octobre 2026 à février 2027 et sera confiée à un prestataire. Un plan de sondage à deux degrés sera mis en œuvre par le SSP, comprenant un premier tirage d'exploitations, puis un second de parcelles agricoles. L'échantillon attendu comprendra environ 22 000 parcelles. Les personnes interrogées pourront répondre via un questionnaire en ligne ou via Prophyl.

À la différence de l'expérimentation Prophyl menée en 2024, qui reposait sur une participation volontaire des exploitants agricoles, l'enquête Phyto 2026 sera une enquête statistique à caractère obligatoire, permettant de répondre à une obligation réglementaire européenne. Dans ce nouveau contexte, pour les enquêtés qui choisiront de répondre via le mode de collecte Prophyl, une solution de recueil du consentement plus allégée et ergonomique leur sera proposée, dans le respect des exigences de la réglementation générale sur la protection des données (RGPD).

* La « bouillie » désigne le mélange prêt à être pulvérisé, composé du produit phytosanitaire (ou de plusieurs produits) dilué dans l'eau, éventuellement avec des adjuvants.

Ce modèle hybride d'enquête facilitera une montée en charge progressive de Prophyl. Il permettra de conserver dans un premier temps les deux modes de collecte (Prophyl et CAWI²²), et de les comparer le cas échéant pour détecter d'éventuels biais ou problèmes dans les données. Chacun de ces deux modes constitue aussi, en soi, une rupture méthodologique : jusqu'à présent, les données sur les produits phytosanitaires étaient recueillies uniquement dans des enquêtes CAPI²³ en face-à-face.

Au-delà de l'échéance de l'enquête Phyto 2026, et en fonction du bilan qui sera tiré de l'opération, cette étape pourrait ouvrir des perspectives pour une évolution plus large du dispositif :

- simplification des mécanismes de recueil du consentement ;
- potentiel enrichissement des données collectées sur les produits phytosanitaires (conditions d'application, matériel utilisé, etc.) ou sur d'autres intrants (par exemple, les engrais) ;
- élargissement de Prophyl à d'autres éditeurs, notamment dans les filières viticoles et maraîchères.

“ La réflexion est engagée également à l'échelle européenne.

La réflexion est engagée également à l'échelle européenne, dans le cadre de la stratégie 2023 pour moderniser la statistique agricole ([encadré 4](#)). Le SSP mène ce projet en partenariat avec CBS et Destatis, les instituts statistiques des Pays-Bas et de l'Allemagne. Cette coopération vise à évaluer le potentiel des logiciels de gestion parcellaire pour la production de statistiques officielles, et à identifier collectivement les verrous techniques, juridiques et éthiques à leur intégration.

La production de statistiques agricoles selon les modalités expérimentées dans le cadre de Prophyl soulève toutefois plusieurs questions fondamentales pour le statisticien.

► Des enjeux à ne pas sous-estimer

La production de statistiques agricoles selon les modalités expérimentées dans le cadre de Prophyl soulève toutefois plusieurs questions fondamentales pour le statisticien.

La première concerne la **représentativité des données**. Comme on peut l'observer dans les enquêtes Pratiques culturelles, les exploitations équipées de logiciels de gestion parcellaire ne reflètent pas nécessairement l'ensemble du tissu agricole. Souvent plus spécialisées, de taille plus importante ou davantage structurées, elles peuvent adopter des pratiques différentes de celles d'exploitations plus traditionnelles. À ce biais structurel s'ajoute celui, non négligeable, du consentement volontaire : sur un sujet aussi sensible que l'usage des pesticides, il peut renforcer les effets de sélection.

22 CAWI pour *Computer-Assisted Web Interviewing* : entretien web assisté par ordinateur. Autrement dit, l'enquêté répond directement dans une application dédiée accessible sur Internet.

23 CAPI pour *Computer-Assisted Personal Interviewing* : entretien individuel assisté par ordinateur. Autrement dit, l'enquêté répond en face-à-face à l'enquêteur, qui saisit les réponses dans une application informatique dédiée.

► Encadré 4. Vers une meilleure intégration des données des logiciels de gestion parcellaire dans les statistiques agricoles européennes

La stratégie 2023 pour moderniser la statistique agricole (Lampach et Mariñez-Solano, 2023) est une initiative européenne financée par une subvention d'Eurostat. Elle vise à moderniser la collecte et la production des statistiques agricoles, notamment par l'exploitation des données des logiciels de gestion parcellaire. On désigne aussi ces logiciels sous le nom de FMIS pour *farm management information systems* (systèmes de gestion et d'information agricoles, en français).

C'est dans ce cadre qu'un projet européen, rassemblant les deux instituts statistiques CBS (Pays-Bas) et Destatis (Allemagne) ainsi que le SSP pour la France, a été retenu par la Commission européenne. Elle porte sur l'utilisation des FMIS à des fins de production de statistiques agricoles. Dans ce projet, le SSP est chargé plus particulièrement d'explorer les leviers d'adhésion des agriculteurs, afin de favoriser un transfert

volontaire, sécurisé et contrôlé des données issues de leurs FMIS à des fins statistiques. Le projet doit livrer ses conclusions à l'automne 2026. Celles-ci alimenteront les réflexions stratégiques de l'Union européenne sur l'évolution des modes de collecte, en particulier dans la perspective de la récente entrée en application du règlement SAIO.

À terme, à l'échelle européenne, l'intégration des FMIS pourrait réduire la charge de réponse des agriculteurs, améliorer la qualité des données collectées, et ouvrir la voie à des enquêtes agricoles plus ciblées, plus fréquentes et moins intrusives. Ces perspectives restent toutefois conditionnées à la résolution de plusieurs enjeux techniques, juridiques, organisationnels et éthiques, notamment en ce qui concerne le consentement, la standardisation des formats, la sécurité des données, et l'interopérabilité entre outils et institutions.

Un deuxième enjeu tient aux **effets propres au mode de collecte**. Transmettre automatiquement des données saisies dans un logiciel, au fil de l'eau, n'est pas équivalent à répondre à un questionnaire en ligne. Certains traitements peuvent être oubliés, enregistrés tardivement ou de manière incomplète. L'expérimentation a montré que les agriculteurs, pourtant volontaires, ne renseignaient leurs traitements phytosanitaires que sur une partie de leurs parcelles (une bonne moitié). À l'inverse, le recours à un logiciel permet de structurer, fiabiliser et tracer l'information, parfois de façon plus rigoureuse que lors d'un entretien en face-à-face, visant à recenser des traitements

phytosanitaires préalablement enregistrés sur un registre. L'enquête Phyto 2026 constituera un terrain d'observation privilégié pour analyser ces effets de mode de collecte et mesurer leurs impacts.

Au-delà des enjeux techniques et méthodologiques, un défi tout aussi déterminant réside dans la communication, la construction de la confiance, et l'adhésion des acteurs au dispositif. Pour assurer l'acceptabilité et renforcer la confiance autour de

Prophyl, le SSP dispose déjà d'une rubrique dédiée sur Agreste²⁴ développée à l'occasion de l'expérimentation de 2024. Il pourra aussi s'appuyer sur son compte LinkedIn²⁵, facilitant le dialogue avec les agriculteurs et l'ensemble du monde agricole. Afin de consolider l'adhésion au dispositif, le SSP prévoit par ailleurs de diversifier les canaux tels que les webinaires et témoignages d'agriculteurs pionniers, mais aussi de renforcer l'accompagnement personnalisé via le réseau des enquêteurs des Draaf. Toute cette nécessaire pédagogie et cette communication patiente rappellent que, si les données privées semblent à portée de main, il ne suffit pas de se baisser pour les moissonner.

²⁴ Voir le lien : <https://agreste.agriculture.gouv.fr/agreste-web/statisticons/ACTIVITES-PROPHYL/listeTypeStatisticon/>.

²⁵ Voir le lien : <https://www.linkedin.com/showcase/agreste-agriculture/posts/?feedView=all>.

► Vers une statistique agricole « connectée » ?

L'expérimentation conduite en 2024, sur l'accès aux données privées des agriculteurs stockées dans leurs logiciels de gestion parcellaire, a donc confirmé la faisabilité de cette approche. Elle a aussi permis d'identifier des écueils à surmonter et des risques à maîtriser. Le succès d'une telle démarche suppose de maintenir un haut niveau d'exigence méthodologique, de garantir une gouvernance rigoureuse, et d'entretenir

un dialogue constant avec les acteurs du secteur agricole. L'enquête Phyto 2026 sera la première à la déployer à grande échelle.

Prophyl, c'est la perspective d'explorer d'autres thématiques grâce à la vaste gamme d'informations contenues dans les logiciels de gestion parcellaire.

Au-delà des traitements phytosanitaires, Prophyl offre également la perspective d'explorer d'autres thématiques, dans la mesure où les logiciels de gestion parcellaire enregistrent une vaste gamme d'informations. En effet, ces outils ne se limitent pas à la saisie des traitements phytosanitaires. Ils recueillent également des données relatives à la fertilisation, aux semis, aux travaux du sol et à

des éléments économiques (temps de travail, coût des intrants, prix de vente, etc.). Il en est de même en élevage, avec les différents outils numériques qui permettent de suivre en temps réel l'état des animaux. Ces informations issues des objets connectés, de plus en plus mobilisés par les agriculteurs, constituent un potentiel précieux pour enrichir la statistique agricole de demain. Elles permettront notamment de mieux appréhender l'utilisation des intrants, les pratiques culturales et d'élevage et, finalement, la durabilité des exploitations agricoles. Autant d'attendus des nouveaux règlements européens SAIO et RIDEA (réseau d'information sur la durabilité des exploitations agricoles)²⁶.

L'approche mise en œuvre dans le cadre du projet Prophyl invite aussi à repenser en profondeur la manière de collecter ou d'acquérir des données, et le positionnement des statisticiens dans ce processus. Plutôt que de (re-)demander à l'enquêté une information qu'il a déjà stockée quelque part sous format numérique, l'enquête se déporte pour se connecter directement au système d'information privé dans lequel se trouve la donnée recherchée, en adéquation avec les pratiques numériques. Au final, le temps économisé sur la « construction » de la donnée (questionnaire, enquête, etc.)²⁷ devra peut-être être réinvesti pour traiter une donnée qui ne se donne pas d'emblée comme telle pour la statistique.

26 Règlement d'exécution (UE) n° 2024/2746 de la Commission du 25 octobre 2024 établissant les modalités d'application du règlement (CE) n° 1217/2009 du Conseil portant création du réseau d'information sur la durabilité des exploitations agricoles et abrogeant le règlement d'exécution (UE) n° 2015/220 de la Commission (voir références juridiques en fin d'article). Le RIDEA est le dispositif qui succédera, à compter de l'exercice 2028 pour la France, à l'actuel RICA (réseau d'information comptable agricole). Il a pour objectif de compléter les données économiques collectées jusqu'à présent par des données en matière environnementale et sociale, afin de pouvoir caractériser la durabilité des exploitations.

27 Voir l'article de Lamarche et Rivière sur le temps de production des statistiques publiques dans ce même numéro.

► Fondements juridiques

- Directive 2003/4/CE du Parlement européen et du Conseil du 28 janvier 2003 concernant l'accès du public à l'information en matière d'environnement et abrogeant la directive 90/313/CEE du Conseil. In : *site de l'Union européenne*. [en ligne]. [Consulté le 4 novembre 2025]. Disponible à l'adresse : <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A32003L0004&qid=1762249461783>.
- Règlement (UE) 2022/2379 du Parlement européen et du Conseil du 23 novembre 2022 relatif aux statistiques sur les intrants et les produits agricoles. In : *site de l'Union européenne*. [en ligne]. [Consulté le 4 novembre 2025]. Disponible à l'adresse : <https://eur-lex.europa.eu/eli/reg/2022/2379/oj?locale=fr>.
- Règlement d'exécution (UE) 2023/1537 de la Commission du 25 juillet 2023 portant modalités d'application du règlement (UE) 2022/2379 du Parlement européen et du Conseil en ce qui concerne les statistiques sur l'utilisation des produits phytopharmaceutiques qui doivent être transmises pour l'année de référence 2026 pendant le régime transitoire 2025-2027 et les statistiques sur les produits phytopharmaceutiques mis sur le marché. In : *site de l'Union européenne*. [en ligne]. [Consulté le 4 novembre 2025]. Disponible à l'adresse : <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A32023R1537&qid=1762249618728>.
- Règlement d'exécution (UE) 2024/2746 de la Commission du 25 octobre 2024 établissant les modalités d'application du règlement (CE) 1217/2009 du Conseil portant création du réseau d'information sur la durabilité des exploitations agricoles et abrogeant le règlement d'exécution (UE) 2015/220 de la Commission. In : *site de l'Union européenne*. [en ligne]. [Consulté le 5 décembre 2025]. Disponible à l'adresse : <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:32024R2746>.
- Code de l'environnement du 21 septembre 2000 – Articles L124-1 à L124-4. In : *site de Légifrance*. [en ligne]. [Consulté le 4 novembre 2025]. Disponible à l'adresse : https://www.legifrance.gouv.fr/codes/section_lc/LEGITEXT000006074220/LEGISCTA000006159211/#LEGISCTA000006159211/#LEGISCTA0.

► Bibliographie

- ANDRIST RANGEL, Ylva, EISERMAN, Daniel, OTTERSKOG, Lena, et REDNER, Anna, 2019. New tools for data collection in Swedish surveys on use of fertilisers and animal manure and cultivation measures in agriculture. In : *62nd ISI Statistics World Congress*. [en ligne]. 18-23 août 2019. Special Topic Session, Volume 1, pp. 162-172. [Consulté le 4 novembre 2025]. Disponible à l'adresse : <https://isi-web.org/proceedings-abstracts>.
- ANXIONNAZ, Isabelle et MAUREL, Françoise, 2021. Le Conseil national de l'information statistique – La qualité des statistiques passe aussi par la concertation. In : *Courrier des statistiques*. [en ligne]. 8 juillet 2021. Insee. N° N6, pp. 123-142. [Consulté le 13 novembre 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/information/5398693?sommaire=5398695>.
- CRAVATTE, Céline, 2025. L'utilisation des produits phytopharmaceutiques dans l'agriculture : cultiver les échanges pour simplifier et enrichir les collectes. In : *Chroniques du Cnis*. [en ligne]. À paraître en décembre 2025. N° 36. Disponible à l'adresse : <https://www.cnis.fr/publications/liste-des-chroniques-du-cnis/>.
- ESPINOSA, Marc, 2023. Moins de substances actives utilisées en arboriculture entre 2012 et 2018 – Enquête Pratiques culturelles en arboriculture. In : *Primeur*. [en ligne]. 28 juin 2023. Agreste. N° 2023-11. [Consulté le 13 novembre 2025]. Disponible à l'adresse : <https://www.agreste.agriculture.gouv.fr/agreste-web/disaron/Pri2311/detail/>.
- FIRQUET, Sandrine et MONTCOUDIOL, Noémie, 2024. Enquête Pratiques culturelles en grandes cultures 2021 – IFT et nombre de traitements. In : *Chiffres et données*. [en ligne]. 24 avril 2024. Agreste. N° 2024-7. [Consulté le 13 novembre 2025]. Disponible à l'adresse : <https://agreste.agriculture.gouv.fr/agreste-web/disaron/Chd2407/detail/>.
- LAMPACH, Nicolas et MARÍNEZ-SOLANO, Domingo, 2023. Strategy to modernise agricultural statistics: new pathways for the future – 2023 edition. In : *Documents de travail statistiques*. [en ligne]. 9 novembre 2023. Eurostat. [Consulté le 21 novembre 2025]. Disponible à l'adresse : <https://ec.europa.eu/eurostat/fr/web/products-statistical-working-papers/w/ks-tc-23-003>.
- LESUR, Romain, 2025. Sources de données privées : panorama et perspectives. In : *Courrier des statistiques*. [en ligne]. 23 juin 2025. Insee. N° N13, pp. 73-94. [Consulté le 4 novembre 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/information/8546943?sommaire=8546949>.
- SINÉ, Mehdi, SABOT, Philippe-Michel et PARISSE, Sandrine, 2025. L'usage des produits phytopharmaceutiques dans l'agriculture. In : *Rapport du groupe de travail du Cnis*. [en ligne]. Cnis. Octobre 2025. N° 171. [Consulté le 4 novembre 2025]. Disponible à l'adresse : <https://www.cnis.fr/instances/groupe-de-travail-lusage-des-produits-phytopharmaceutiques-dans-lagriculture/>.

- SNIJKERS, Ger, PÉREZ José Gómez et DE JONG, Tim, 2022. System-to-system data collection in business surveys applied to an agricultural survey: a proof of concept. In : *UNECE Expert meeting on Statistical Data Collection – Towards a new normal?*. [en ligne]. 3 octobre 2025. [Consulté le 4 novembre 2025]. Disponible à l'adresse : <https://unece.org/statistics/documents/2022/10/presentations/dc2022s2netherlandssnijkers-et-al>.
- SNIJKERS, Ger, SARAIVA, Paulo, VAN WIJK, Derk, MARTINEZ-SOLANO Jose Domingo, et RODRIGUES, Sofia, 2025. Automated business data collection: Experiences in Portugal and The Netherlands. In : *Statistical Journal of the IAOS*. [en ligne]. 25 mai 2025. Volume 41, n° 2. [Consulté le 4 novembre 2025]. Disponible à l'adresse : <https://journals.sagepub.com/doi/abs/10.1177/18747655251341465>.

À propos du temps de production des statistiques publiques



Pierre Lamarche* et Pascal Rivière**

Si la production d'informations conjoncturelles s'opère dans des délais très brefs, nombre de statistiques publiques requièrent un temps significatif. Qu'elles proviennent d'enquêtes auprès de ménages ou d'entreprises, ou de sources administratives ou privées, trois grandes phases se dégagent. D'abord celle de la préparation : un temps de travail en amont, par exemple de concertation, qui n'entre pas dans le « délai de diffusion » des chiffres. Puis celle de la construction des données : via un processus de collecte pour les enquêtes, ou de transformation dans le cas de sources externes. Enfin, une dernière phase de traitement statistique et de diffusion, qui inclut en particulier des activités de validation.

Les phases de préparation et de traitement statistique sont d'autant plus chronophages que le nombre de variables est élevé. Le temps de la préparation dépend aussi fortement de la complexité et de la sensibilité du dispositif à établir. Celui de la construction croît généralement avec le nombre d'unités. Enfin, le temps passé sur la dernière phase dépend en particulier du niveau de détail auquel les statistiques sont diffusées.

De manière générale, la statistique publique produit certes des données, mais surtout du sens. Ainsi, même si une meilleure automatisation est envisageable, une large part du temps de travail réside dans l'analyse, l'interprétation, les vérifications, et toutes les boucles de rétroaction que cela induit.

 Even though short-term business statistics can be produced very quickly, many official statistics require a significative amount of time. Be they originating in surveys towards households or businesses, or in administrative or private data, three main stages appear. First, preparation: a work period beforehand, dedicated to consultations for instance, that is not accounted for in the timeliness of statistics. Then, data construction: through a collecting process for surveys or a transformation one for external sources. Finally, data processing and distribution, which includes validation works in particular.

Preparation and data processing stages are more time-consuming the higher the amount of variables. Preparation time also depends greatly on the complexity and sensitivity of the process to devise. The construction time usually increases with the amount of units. Finally, time spent on the last stage depends specifically on the statistical granularity.

Generally speaking, statistical institutes produce insights as well as numbers, they also construct meaning and interpretation. Thus, even if a better automation may be worth considering, a large amount of the time spent on production resides in analysing, interpreting, verifying, and all the feedback loops that ensue.

* Chef de l'unité Innovation et stratégie du système d'information, DSI, Insee.
pierre.lamarche@insee.fr

** Chef de l'Inspection générale, Insee.
pascal.riviere@insee.fr

Dans un monde où les données sont abondantes et accessibles, dans un contexte d'essor du big data et de l'open data, mais aussi de possibilités d'appariement accrues (Benichou et al., 2023), un nouveau champ des possibles s'ouvre pour la statistique publique (Elbaum, 2018). Des plateformes innovantes de traitement des données rendent tout le processus de production statistique plus adaptable à de gros volumes et à des configurations hétérogènes (Comte et al., 2022). Elles le rendent aussi plus réplicable, avec la logique des *pipelines*¹ (Cotton et Haag, 2023). On observe également une prolifération d'outils facilitant le développement, entre autres via une utilisation croissante de l'intelligence artificielle.

On pourrait en déduire hâtivement que, quelle que soit l'origine des données, les statistiques publiques deviennent de plus en plus « faciles à produire » et que, profitant des nouvelles sources et de la *data science*, on devrait pouvoir désormais les réaliser beaucoup plus rapidement. Cette volonté de réduire les délais de production est d'ailleurs régulièrement affichée comme un objectif fort au niveau d'Eurostat².

Nouvelles données ou pas, nouvelles techniques ou pas, l'élaboration de statistiques ayant une qualité suffisante pour servir de référence dans le débat public requiert toute une gamme d'opérations.

Dans de nombreux cas, la rapidité de réalisation des statistiques est inhérente au processus, lorsque ce dernier a justement été conçu avec cette exigence principale (indice de production industrielle ou enquêtes de conjoncture, par exemple). Dans d'autres cas, des gains de temps ont effectivement été observés ces dernières années (statistiques sur l'emploi ou comptabilité nationale, par exemple). Pourtant, dans beaucoup de situations, il y a loin de la coupe aux lèvres. Nouvelles données ou pas, nouvelles techniques ou pas, l'élaboration de statistiques ayant une qualité suffisante

pour servir de référence dans le débat public requiert toute une gamme d'opérations. Or, une grande partie d'entre elles ne se prête pas à une quelconque automatisation et leur accélération paraît de fait moins évidente.

L'objet du présent papier est de mieux comprendre tous les éléments qui expliquent le temps de production des statistiques publiques, et de déterminer ainsi les marges de manœuvre pour améliorer cet aspect sans dégrader les autres critères.

- Les marges de manœuvre peuvent différer selon les dispositifs mis en place. On distinguera ainsi **trois types de processus**³ : les enquêtes, piliers historiques des statistiques officielles ; les processus fondés sur l'usage de données « administratives » (appartenant à des administrations, par exemple des organismes de protection sociale) ; enfin, ceux qui s'appuient sur des données « privées » (appartenant⁴ à des entreprises privées, par exemple à des opérateurs de téléphonie mobile⁵). Il arrive, comme dans

¹ Traitements de données standardisés.

² Eurostat est l'institut statistique communautaire, direction générale de la Commission européenne.

³ Il existe aussi des situations hybrides, par exemple le cas des données relatives aux droits de mutation à titre gratuit (où il y a historiquement un échantillonnage, mais qu'on ne maîtrise pas).

⁴ Avec « appartenant », on veut simplement dire que ces données figurent dans des bases de données qui sont sous la responsabilité de ces entreprises (ou, respectivement, de ces administrations).

⁵ Lesur (2025) décrit les différentes sources privées qui ont été mobilisées par l'Insee dans le cadre de partenariats.

le cas des statistiques structurelles d'entreprises, que la production statistique mixe plusieurs types de processus.

- Le processus de production lui-même, de façon générale, peut être subdivisé en **trois grandes phases** : préparation, construction des données, traitement statistique et diffusion.

Pour chacune de ces phases, on analysera les déterminants du temps de réalisation des statistiques selon le type de dispositif. On verra également que la situation n'est pas la même selon que l'on diffuse des statistiques pérennes, régulières, ou au contraire ponctuelles, *one-shot*. On en tirera quelques leçons sur les principaux facteurs explicatifs du temps de réalisation, et sur les possibilités de gain de temps.

Mais avant cela, il nous faudra répondre à une autre question, a priori plus basique : au fait, de quel « temps » parle-t-on exactement ?

► **Le temps nécessaire : plusieurs définitions, plusieurs moments-clés**

Dans le processus complexe qui mène à la publication de statistiques sur un sujet, on peut distinguer quatre dates essentielles, qu'on notera de t1 à t4 (**figure 1**) :

- t1 : le démarrage de la réflexion – qui va se matérialiser, par exemple, par l'explicitation du besoin et le recensement des sources d'information disponibles ;
- t2 : la date de référence des statistiques – par exemple, on récupère des données fiscales portant sur le mois de décembre 2024, ou bien, on réalise une enquête auprès d'entreprises relativement à l'année 2024⁶ ;
- t3 : la date de fin de récupération des données – c'est-à-dire la fin de collecte, pour une enquête, ou la date d'obtention de la dernière version du fichier administratif ;
- t4 : la date de publication des statistiques.

Le critère privilégié pour apprécier l'actualité (ou *timeliness*) des statistiques produites, en référence au treizième principe du **Code de bonnes pratiques de la statistique européenne**⁷, est le **délai entre la fin de la période de référence d'un phénomène statistique et la date de publication des résultats**. C'est donc t4 – t2.

Cette mesure, dont on comprend bien l'enjeu quant à la « fraîcheur » des données, ne rend pas compte de l'entièreté du temps de production statistique. En effet, elle fait l'impasse sur tout le travail effectué avant la période de référence, donc sur l'écart t2 – t1. Ainsi, pour l'indice des prix à la consommation, il a fallu une dizaine d'années pour intégrer des **données de caisse**⁸ dans la production, entre la mise en place des conventions avec les enseignes de la grande distribution alimentaire et l'investissement

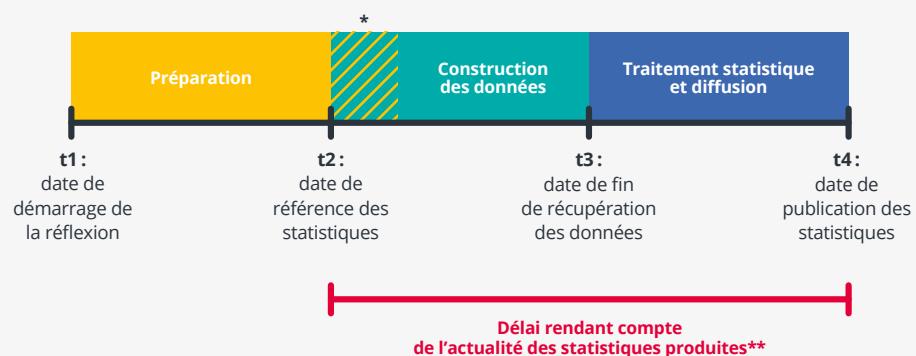
⁶ Lorsque la référence temporelle est en réalité une période, la date t2 que l'on considère est la fin de la période, donc le 31 décembre dans le cas d'une année, le dernier jour du mois dans le cas d'un mois. Pour la définition de la période de référence, voir le lien suivant :

https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Reference_period.

⁷ Voir le lien : <https://www.insee.fr/fr/information/4140105>.

⁸ Voir le lien : <https://www.insee.fr/fr/metadonnees/definition/c2159>.

► Figure 1 - Les trois phases du processus de production des statistiques publiques



* Pour certains processus, la fin de la préparation / le début de la collecte peut être postérieur à la date de référence des statistiques. C'est le cas par exemple pour les données fiscales, les déclarations de revenus fiscaux étant postérieures à la période de référence.

** L'actualité (*timeliness* en anglais) est un critère de qualité des statistiques publiques mis en avant dans le Code de bonnes pratiques de la statistique européenne.

informatique requis (Leclair, 2019). Bien sûr, ce temps d'investissement n'est plus nécessaire pour les périodes suivantes (ici, les mois suivants), mais on ne peut passer sous silence cette longue préparation.

Par ailleurs, lorsqu'on utilise des données administratives, leur date d'obtention est toujours décalée par rapport à la date de référence. Par exemple, les données fiscales définitives relatives à l'année 2023 arrivent à l'Insee début 2025. En d'autres termes, t₃ – t₂ n'est pas nul. Dans les faits, les statisticiens n'ont donc de prise que sur le résidu consacré au traitement « statistique » des données, c'est-à-dire t₄ – t₃.

Commençons alors par nous intéresser au temps de préparation t₂ – t₁⁹, qui a son importance en particulier dans le contexte d'intégration de données nouvelles dans le système d'information statistique.

► La phase de préparation

Pour les enquêtes, ménages ou entreprises, concerter pour convaincre

La concertation est la clé de la crédibilité de la statistique publique : en réunissant autour d'une même table producteurs et utilisateurs des statistiques publiques, elle assure qu'un sujet a été correctement évalué dans l'ensemble des dimensions pertinentes. Grâce à elle, le système d'information bâti pour mesurer et analyser un phénomène

⁹ À noter que t₂ – t₁ est une approximation du temps de préparation, car t₂ est la date de référence des statistiques et non la fin de préparation. C'est donc un minorant, mais c'est un bon ordre de grandeur.

social ou économique résulte bien d'un consensus, dans la limite des moyens alloués. L'objectif est de vérifier l'adhésion aux concepts, ainsi qu'au dispositif qui doit permettre de les mettre en œuvre.

La concertation est la clé de la crédibilité de la statistique publique : en réunissant autour d'une même table producteurs et utilisateurs des statistiques publiques, elle assure qu'un sujet a été correctement évalué dans l'ensemble des dimensions pertinentes.

On reviendra plus bas sur les instances mises en place dans cet objectif. Cette étape-clé peut s'avérer longue, car elle suppose de la réflexion et une discussion entre personnes venues d'horizons professionnels très différents. C'est souvent un processus itératif, qui demande un certain nombre de va-et-vient entre les statisticiens qui construisent l'outil de mesure et les personnes consultées (*figure 2*).

Ainsi, dans le cadre des enquêtes auprès des ménages, l'élaboration du questionnaire résulte d'échanges très poussés entre statisticiens, chercheurs, experts des domaines concernés et partenaires sociaux¹⁰. Ces considérations

s'appliquent naturellement pour des dispositifs d'enquêtes en construction ou en refonte. Mais même pour les enquêtes régulières bien installées, il convient de s'assurer que le questionnaire soumis aux ménages est toujours d'actualité et n'omet pas d'éventuelles nouveautés. Il faut dans bien des cas satisfaire des objectifs parfois contradictoires : construire un questionnaire le plus complet possible, permettant d'analyser l'ensemble des dimensions d'un phénomène donné, et qui soit en même temps intelligible et d'une complexité et d'une longueur maîtrisées. Par exemple, si l'on s'intéresse aux mécanismes de transmission intergénérationnelle du patrimoine, il est important de demander aux personnes répondant à l'*enquête Histoire de vie et Patrimoine*¹¹ de décrire leur fratrie et plus particulièrement leur rang de naissance ; ou encore de décrire leurs conditions de vie durant leur enfance et les revenus de leurs parents. Il s'agit donc d'un exercice d'équilibrisme, entre richesse de questionnement et attention portée à la facilité de réponse.

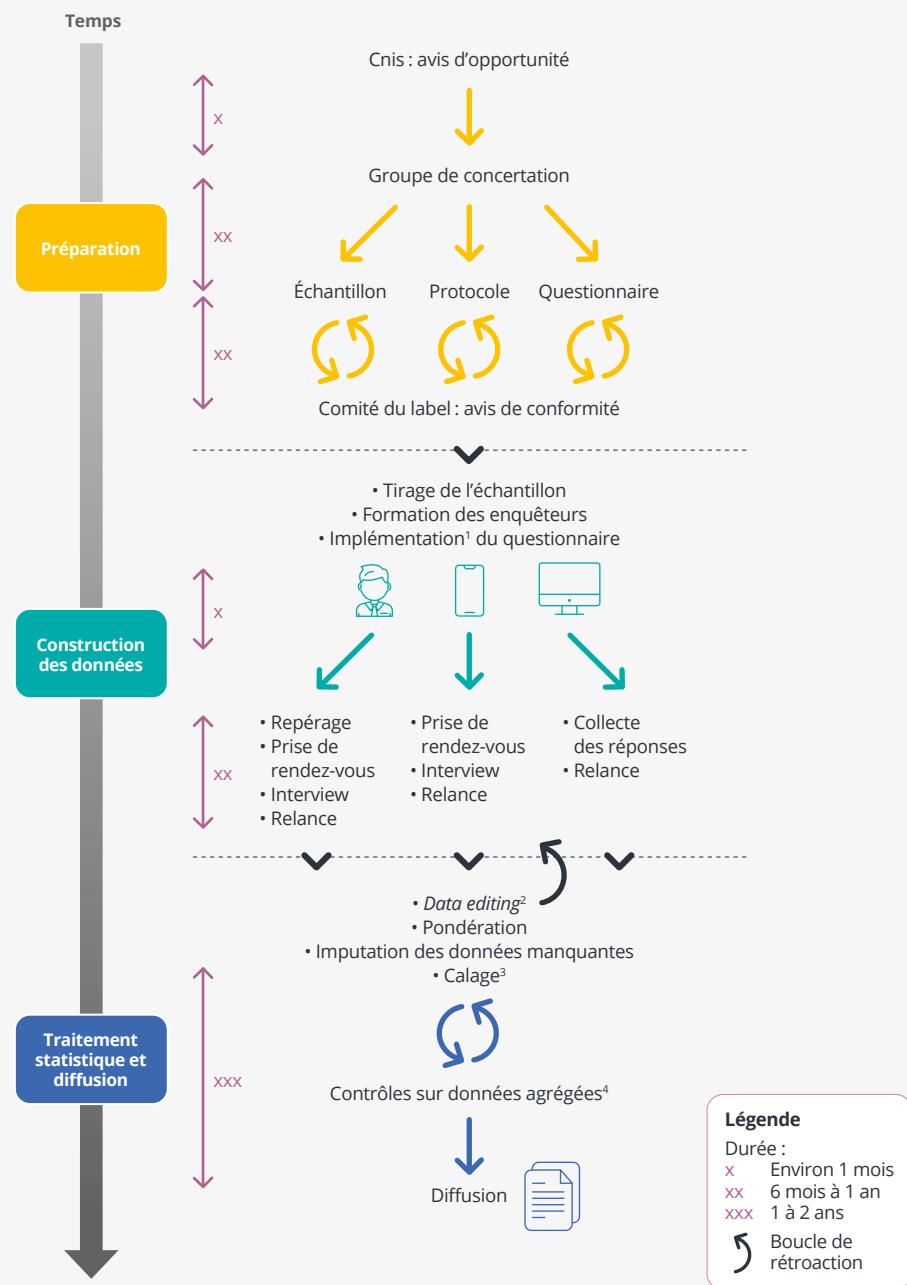
Rechercher la pertinence des concepts...

Les concepts sur lesquels reposent les indicateurs statistiques construits à partir de ces données d'enquêtes ne vont pas nécessairement de soi. L'élaboration de ces concepts, dont la nature est largement séculaire, nécessite un effort de réflexion qui peut s'inscrire dans la durée et qui peut naturellement prêter à débat. La pauvreté en est un exemple emblématique : cette notion dépend largement du contexte social et elle varie beaucoup dans l'espace et dans le temps. Il faut alors s'assurer de la pertinence des indicateurs la caractérisant. Les réflexions sur la mesure de la pauvreté débutent dans les années 1970, avec les premiers travaux de chercheurs comme ceux d'Anthony B. Atkinson (Atkinson, 1975). Les recherches se poursuivent durant les années 1980 et 1990. Face au caractère

¹⁰ Sans compter les enquêtes encadrées par un règlement de l'Union européenne, qui font également l'objet d'une concertation à ce niveau.

¹¹ Voir le lien : <https://www.insee.fr/fr/metadonnees/source/serie/s1005>.

► Figure 2 - La production de statistiques publiques fondées sur une enquête



Cnis : Conseil national de l'information statistique.

1. Programmation (quand il ne s'agit pas d'un questionnaire à remplir sur papier).

2. Processus de vérification des données. Cette étape peut conduire à retourner à la construction des données (rappeler des entreprises...).

3. Processus visant à modifier légèrement les pondérations de manière à retrouver des grandeurs connues (comme le nombre d'individus).

4. Contrôles de cohérence avec les résultats d'autres sources officielles.

multidimensionnel de la pauvreté, plusieurs approches sont développées : pauvreté absolue ou relative, monétaire, en conditions de vie, subjective, administrative... (Insee, 1998). Dans les années 2000, des travaux conséquents ont lieu au niveau européen avec la [stratégie de Lisbonne](#)¹². Ils aboutissent à la mise en place du dispositif d'enquête européen EU-SILC¹³. Ce dernier permet des analyses comparatives selon différentes approches, dont en premier lieu la pauvreté monétaire et la privation matérielle et sociale.

... ainsi que la crédibilité du dispositif

Une fois les concepts bien établis, il faut également concevoir le dispositif de collecte garantissant l'information la plus fiable à moindre coût. Ainsi, une fois que l'on a défini un indicateur de pauvreté basé sur les revenus, il faut trouver la meilleure façon de mesurer ces revenus ; ce que l'on a d'abord fait par le biais d'un questionnaire, dont il a fallu s'assurer qu'il couvrait correctement les différentes composantes du revenu et qu'il était compréhensible du grand public. Ceci passe par le test du questionnaire auprès d'un petit échantillon d'individus ou d'entreprises, qui permet également d'évaluer la bonne compréhension des concepts et la disponibilité de l'information auprès des unités enquêtées.

Pour incorporer les réflexions conceptuelles dans un questionnaire d'enquête, entre les différents tests et les passages devant les instances de concertation, il peut s'écouler jusqu'à un ou deux ans pour un dispositif complètement nouveau.

Combien de temps alors pour construire un questionnaire permettant d'incorporer les réflexions conceptuelles ? Entre les différents tests et les passages devant les instances de concertation que sont le Conseil national de l'information statistique (Cnis) (Anxionnaz et Maurel, 2021) et le Comité du label (Christine et Roth, 2020), les réflexions peuvent ainsi s'étendre jusqu'à un ou deux ans pour un dispositif complètement nouveau. Le passage devant le Cnis doit permettre d'abord d'examiner l'opportunité du dispositif de collecte dans le contexte social et statistique.

Moment clé dans la vie d'une enquête, il permet bien souvent d'initier des premiers contacts avec le monde de la recherche et la société civile, pour constituer ensuite le groupe de concertation qui sera sollicité pour l'élaboration du questionnaire. S'il n'est pas évident de caractériser un questionnaire achevé¹⁴, on peut estimer que la concertation est terminée lorsque l'ensemble des remarques formulées par les parties prenantes, ainsi que les résultats des tests sur le terrain, ont pu être instruits et ont fait l'objet d'arbitrages sur le questionnaire. Décrit ainsi, l'exercice peut sembler simple : il n'en demeure pas moins chronophage et laborieux, l'élaboration d'un questionnaire résultant d'une tension constante entre parcimonie et exhaustivité.

¹² Voir le lien : <https://www.europarl.europa.eu/highlights/fr/1001.html>.

¹³ European Union statistics on income and living conditions (statistiques sur les revenus et les conditions de vie dans l'Union européenne). Voir le lien : <https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions>.

¹⁴ Dans la mesure où les problématiques qui traversent la société et l'économie sont en évolution constante et que celles-ci peuvent mériter des éclairages qui évoluent au cours du temps.

Données administratives : comprendre l'univers et élaborer un cadre

Lorsque les données sont fournies à la statistique publique, en provenance d'entités administratives, le travail de préparation est radicalement différent : il ne s'agit plus de se mettre d'accord sur des variables à collecter, encore moins sur un questionnaire, puisque le contenu des données est imposé aux statisticiens publics (Hand, 2018). Il faut donc d'abord comprendre un environnement externe, un contenu métier étranger, ce qui requiert un long travail de « découverte » de l'univers en question (*figure 3*). Cette familiarisation avec un autre langage, d'autres manières de travailler, d'autres codes, se révèle nécessaire si l'on veut utiliser ces données sans commettre d'erreur d'interprétation grossière. Elle est à ce point importante que l'*Office for National Statistics (ONS)*¹⁵ est allé jusqu'à mettre en place des détachements de statisticiens, pendant plusieurs mois, au sein d'administrations détentrices des données (Fermor-Dunman et Parsons, 2022).

Cette phase de découverte est l'occasion d'acquérir une bonne connaissance de la source administrative, de son statut, du contexte dans lequel elle s'inscrit (juridique, notamment) et des liens existants avec d'autres entités. De ce point de vue, les discussions peuvent prendre beaucoup de temps. En effet, le cadre juridique dans lequel s'exerce l'activité de la statistique publique n'est pas toujours bien connu des autres administrations.

Aussi, la sensibilité des données susceptibles de faire l'objet d'un échange peut conduire leur détenteur à une certaine prudence. Puis, peu à peu, des modalités de coopération s'établissent entre l'administration, en tant que futur « fournisseur » des données, et le service statistique, en tant que futur « récepteur » de ces données.

Il ne s'agit pas seulement d'accéder à une documentation « technique », mais également de comprendre ce qui la sous-tend.

les métadonnées avant tout travail statistique. Il ne s'agit pas seulement d'accéder à une documentation « technique », mais également de comprendre ce qui la sous-tend. Ainsi, utiliser des données fiscales requiert un minimum de maîtrise de la législation fiscale. Il s'agit aussi de décider des variables nécessaires. Par exemple, une réflexion fine a été menée pour décider des variables issues de fichiers administratifs à intégrer dans le *répertoire statistique des individus et des logements (Résil)*¹⁶ tout en limitant les risques d'identification des personnes (Lefebvre, 2024).

Le dialogue avec l'administration vise enfin à mettre en place un processus opérationnel de fourniture des données, ce qui n'est pas une mince affaire : jeux de données complexes, volumineux, souvent transmis en plusieurs fois, avec possibilités de retour en arrière si les données sont incomplètes ou erronées. Il en résulte une convention entre administration et service statistique, dont l'élaboration, issue de multiples allers-retours,

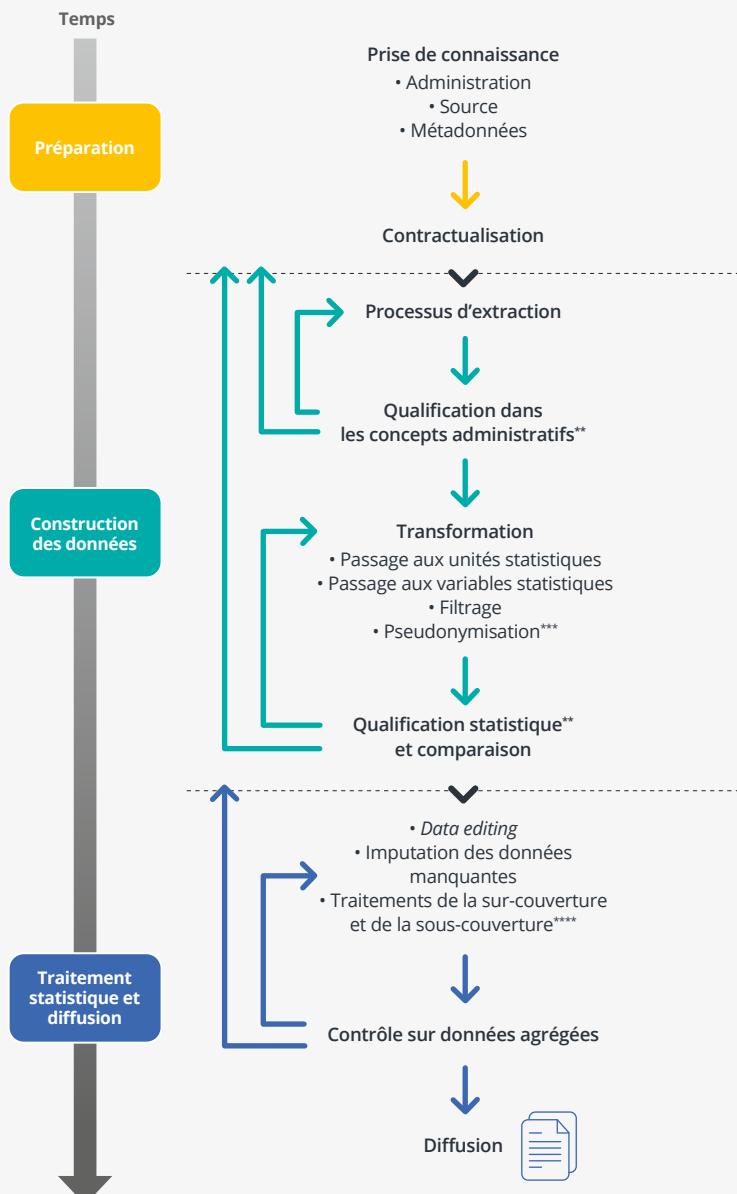
¹⁵ Institut national de statistique du Royaume-Uni.

¹⁶ Une métadonnée est une donnée qui définit et décrit une autre donnée. Voir à ce sujet Bonnans (2019).

¹⁷ Par exemple, la norme d'échanges NEODES pour la déclaration sociale nominative (Humbert-Bottin, 2018).

¹⁸ Voir le lien : <https://www.insee.fr/fr/information/7748881>.

► **Figure 3 - La production de statistiques publiques fondées sur des données administratives – schéma très simplifié***



- * Ce schéma est une synthèse simplifiée de trois schémas successifs de Koumarianos et Rivière (2025).
- ** La première qualification des données, essentielle, veille à s'assurer du bon état des données a priori, à aboutir à une « source », tout en restant strictement dans le cadre des concepts administratifs. La qualification statistique ayant lieu après l'étape de transformation veille quant à elle à s'assurer que les variables et unités statistiques sont conformes et cohérentes avec l'objectif recherché.
- *** La pseudonymisation est un traitement de données à caractère personnel de manière qu'on ne puisse pas attribuer les données à une personne physique sans avoir recours à des informations supplémentaires (Wikipedia).
- **** Sous-couverture : le champ couvert représente moins que la population cible, « il en manque ». Sur-couverture : le champ couvert va au-delà de la population cible, « il y en a trop ».

Lecture : Les phases de qualification et de contrôle des données peuvent entraîner des boucles de rétroaction.

peut s'avérer chronophage. Mais on acceptera d'y passer un temps très significatif si l'échange est régulier, pérenne, comme c'est le cas pour la déclaration sociale nominative (DSN) (Renne, 2018).

Données privées : de nouvelles exigences

L'utilisation de données privées peut sembler très similaire à celle des données administratives : dans les deux cas, il s'agit de récupérer une masse d'informations provenant d'un autre environnement, et donc de comprendre un univers métier, d'accéder aux métadonnées, et d'aboutir à la rédaction d'une convention. Tout le temps de découverte, de compréhension de la source, d'obtention des métadonnées, etc., demeure ici valable.

Mais ces sources soulèvent des défis supplémentaires. La fourniture de données entre administrations est d'autant plus naturelle que celles-ci sont de plus en plus soumises à une exigence d'ouverture. Ce n'est pas le cas des acteurs privés, qui sont dans les faits beaucoup moins tenus de transmettre des informations, et qui n'en voient pas nécessairement l'utilité. La donnée est par ailleurs bien identifiée par ces acteurs privés comme un actif de plus en plus stratégique. Il n'est donc pas simple de les convaincre de partager cet actif avec une administration, sans intérêt évident.

Ici, le dialogue relève donc aussi, en quelque sorte, de la négociation, dans laquelle il faut penser aux contreparties.

Ici, le dialogue relève donc aussi, en quelque sorte, de la négociation, dans laquelle il faut penser aux contreparties. En particulier, accepte-t-on ou non de payer les données, comme le fait par exemple l'*Instituto Nacional de Estadística* (INE)¹⁹ auprès d'opérateurs espagnols de téléphonie mobile ? Si oui, quels sont les critères pour déterminer dans quels cas on paye ? Si non, il faut que l'entreprise y trouve un autre intérêt, et d'autres contreparties sont alors possibles (Joubert, 2025). Au niveau européen,

un effort a également été entrepris pour accéder à certaines données existantes au niveau supranational, comme en témoigne le projet sur les plateformes d'économie collaborative d'Eurostat (Eurostat, 2025).

Quel que soit l'accord mis en place, comme pour les sources administratives, le statisticien doit comprendre le contexte de génération de la donnée, ce qui nécessite de nombreux échanges avec le « propriétaire ». Ceux-ci se feront de manière plus naturelle dans un cadre de coopération, à travers un partenariat entre statistique publique et entreprise privée détentrice des données. L'accord doit par ailleurs intégrer des exigences de confidentialité particulière : à celle des données relatives aux personnes, s'ajoute ici le secret des affaires, qui limite, ou plutôt encadre, l'acceptation du propriétaire à transférer des données (Bonnet et Loisel, 2024).

La convention, cadre contractuel contraignant, est donc tout à fait essentielle. Mais elle sera beaucoup plus longue à réaliser et à mettre en œuvre, comme ce fut le cas pour les données de caisse (Leclair, 2019). Souvent, on passe par des phases d'expérimentation,

¹⁹ Institut national de statistique de l'Espagne.

nécessitant de faire intervenir d'autres partenaires, tels que des instituts de recherche (Boittelle et al., 2025) : la multiplicité des acteurs allonge d'autant plus le processus. Par ailleurs, rien ne garantit que la convention soit pérenne : par exemple, la fourniture gratuite de données de téléphonie mobile par Orange a été possible en 2020, au plus fort de la crise sanitaire (Tavernier, 2020), mais n'a pas été poursuivie.

Ainsi, quel que soit le dispositif, il existe toujours un temps significatif de préparation, de cadrage, de négociation : un investissement indispensable avant tout travail effectif sur les données. Pour des opérations pérennes, régulières, celui-ci bénéficie à toutes les opérations ultérieures, et les travaux de préparation sont appelés à se répéter dans le temps, avec une ampleur généralement plus faible qu'au départ.

► La phase de construction des données

On entre ici dans une phase plus opérationnelle, le temps t3 – t2, où l'on va s'attacher à construire les données individuelles : soit à travers un processus de collecte (enquêtes auprès des ménages ou des entreprises), soit via un travail d'analyse, de qualification et de transformation de données obtenues par ailleurs (sources administratives ou privées).

Enquêtes auprès des ménages : un temps de collecte incontournable

Pour ce qui est des enquêtes, la statistique publique a une maîtrise assez complète de ce processus, qu'elle passe par un prestataire pour réaliser le terrain ou qu'elle se charge de celui-ci par ses moyens propres. Néanmoins, ce temps n'est pas le même pour toutes les opérations. En effet, mesurer un phénomène saisonnier comme la consommation nécessite une collecte sur une année complète. Par ailleurs, mesurer un phénomène structurel – objet d'enquêtes souvent très espacées dans le temps, voire ponctuelles – suppose de collecter plus d'information. En effet, le processus de concertation porte en général sur un ensemble de variables plus large. Pour ces enquêtes, il faut également davantage de préparation du terrain dans la mesure où le réseau d'enquêteurs n'est pas toujours familier du questionnaire.

De manière générale, la collecte peut se découper en différentes phases. Dans un premier temps, il y a la **phase d'organisation**, qui peut être plus ou moins courte selon que l'enquête est un dispositif récurrent ou non. Elle inclut l'élaboration du protocole de collecte et la rédaction d'une documentation le décrivant. Cette dernière englobe notamment les consignes aux enquêteurs. Il faut également former l'ensemble du réseau des enquêteurs, ce qui peut prendre du temps. Ainsi, lorsque le réseau de l'Insee est sollicité, il faut d'abord former les gestionnaires au niveau national, lesquels forment à leur tour les enquêteurs dans chaque direction régionale. Le processus de formation peut donc prendre un peu de temps, un mois, voire plus dans le cas d'enquêtes comme **Sans Domicile**²⁰ : à la fois parce qu'il se fait de manière décentralisée et parce qu'il faut accommoder les agendas très chargés du réseau d'enquêteurs.

20 Voir le lien : <https://www.insee.fr/fr/metadonnees/source/serie/s1002>.

L'organisation inclut également des étapes essentielles comme la sélection de l'échantillon de personnes à enquêter, l'impression des fiches-adresses (coordonnées des personnes échantillonnées) et leur répartition entre les enquêteurs. L'envoi des lettres-avis annonçant l'enquête auprès des ménages est aussi un jalon essentiel, qui doit s'anticiper pour tenir compte des délais d'acheminement du courrier. Si ces éléments préparatoires peuvent, pour certains, se dérouler en parallèle, la programmation infra-annuelle des enquêtes impose un calendrier permettant d'identifier bien en amont du terrain les zones d'enquête concernées et la volumétrie associée.

Dans une collecte en face-à-face ou par téléphone, les enquêteurs doivent entrer en contact avec les ménages sélectionnés, les convaincre de répondre à l'enquête et souvent prendre rendez-vous avec eux.

Puis vient la **phase du terrain** proprement dite, qui dépend fortement du mode de collecte. Dans une collecte en face-à-face ou par téléphone, les enquêteurs doivent entrer en contact avec les ménages sélectionnés, les convaincre de répondre à l'enquête et souvent prendre rendez-vous avec eux. Lorsque la taille du réseau d'enquêteurs est fixe, le temps associé à cette phase dépend directement de la volumétrie totale de l'échantillon. Par ailleurs, le protocole peut être plus ou moins chronophage selon le mode de collecte. Ainsi,

une collecte en face-à-face nécessite une étape préliminaire, dite de repérage, pendant laquelle l'enquêteur s'assure sur le terrain de l'existence du logement et s'attache à entrer en contact avec ses occupants. Aujourd'hui, les données de contact sont généralement disponibles dans les bases de sondage utilisées pour les enquêtes « ménages », ce qui rend cette phase moins cruciale. Elle reste cependant importante, car elle demeure un moyen essentiel pour atteindre les ménages concernés par l'enquête.

De manière générale, la phase de prise de contact peut durer, car les ménages ne sont pas toujours faciles à joindre. Notamment, certains profils peuvent être rarement présents à leur domicile durant les temps de collecte. Le caractère représentatif de l'enquête, et donc sa qualité, sont donc déterminés par le temps et l'effort passés à contacter l'ensemble des ménages de l'échantillon, qu'ils soient faciles ou non à contacter. Dans le cadre d'une collecte par web, ces questions se posent moins. Mais la collecte uniquement selon ce mode n'est pas un protocole envisageable pour des raisons de représentativité. En effet, les taux de réponse pour la collecte web sont très largement inférieurs à ceux d'une collecte intermédiaire par un enquêteur, le rôle de ce dernier étant déterminant dans le processus d'adhésion à l'enquête. Le web est plutôt mobilisé dans le cadre de protocoles multimodes (Beck et al., 2022). Dans le cas où ces derniers sont mobilisés de manière séquentielle, par exemple web, sinon téléphone, il est nécessaire de finir la collecte dans un mode avant de passer à un autre, ce qui aboutit à rallonger significativement le temps de collecte.

Enfin, de façon très variable selon les enquêtes, il existe un **travail d'apurement réalisé par les gestionnaires** : vérification de la bonne compréhension des consignes par les enquêteurs, ainsi que de la cohérence de quelques données collectées. Cette phase, si elle se déroule pour une bonne partie durant la collecte proprement dite, peut nécessiter des allers-retours entre les gestionnaires et les enquêteurs, et dans certains cas avec les enquêtés (Forteza et García-Uribe, 2025).

Enquêtes auprès des entreprises : la place centrale du *data editing*

En matière de collecte, la réalisation d'enquêtes auprès d'entreprises présente des différences importantes par rapport à celles effectuées auprès de ménages. Tout d'abord, pour l'essentiel, ce sont des questionnaires auto-administrés²¹, historiquement transmis par courrier, et désormais quasi systématiquement sur le web. Le temps de l'organisation du réseau d'enquêteurs et des interactions entre enquêteurs et enquêtés ne joue donc plus. Il y a bien un suivi d'avancement de la collecte, mais d'une autre nature : on peut connaître à tout moment la liste des non-répondants, ce qui

permet d'organiser auprès d'eux un rappel, éventuellement en plusieurs étapes. Cependant, si les relances améliorent le taux de réponse et donc la précision, elles détériorent le délai : c'est donc un compromis à trouver.

Ce qui caractérise les enquêtes auprès d'entreprises, c'est l'importance considérable du *data editing*, c'est-à-dire du processus de vérification des données, automatique ou manuel, et des modifications que cela induit.

Ce qui caractérise également les enquêtes auprès d'entreprises, c'est l'importance considérable du *data editing*, c'est-à-dire du processus de vérification des données, automatique ou manuel, et des modifications que cela induit. Sur ce sujet, toute une littérature s'est développée depuis longtemps dans la communauté statistique internationale (Granquist, 1997 ; De Waal et al., 2011 ; Pannekoek et al., 2013)²². Dans le cas des

enquêtes « entreprises », il s'agit de vérifier non seulement les questionnaires jugés douteux, mais aussi les données à fort impact sur les agrégats ; car l'univers observé est très hétérogène, et une seule erreur peut tout dégrader (Lawrence et McKenzie, 2000).

Contrairement aux enquêtes « ménages » où l'on peut souvent imaginer d'automatiser le contrôle et l'imputation²³, opérations qui relèvent alors plutôt des traitements statistiques, le contrôle manuel se révèle ici indispensable et prend inévitablement du temps. À nouveau, c'est un compromis à déterminer entre précision statistique et délai de production (Granquist et Kovar, 1997). Ce temps de vérification porte sur toutes les « cases » des tableaux finaux. En effet, la demande des utilisateurs, par exemple celle des fédérations professionnelles, peut concerner un secteur d'activité très précis (au niveau le plus détaillé de la *nomenclature d'activités française (NAF)*²⁴, par exemple). Dès lors, le fait que les statistiques produites portent sur des granularités²⁵ très fines conduit mécaniquement à un temps de travail significatif pour les gestionnaires d'enquête, car proportionnel au nombre de cases de diffusion. Il est cependant possible de prioriser le travail pour corriger avant tout les erreurs les plus influentes sur les résultats, grâce à

²¹ Remplis directement par les enquêtés, sans l'intermédiation d'un enquêteur.

²² Caron (2025) donne une première approche des principes généraux de mise en œuvre du *data editing* qui existent dans la littérature.

²³ Voir l'article fondateur de Fellegi et Holt (1976). L'imputation est une technique statistique utilisée pour remplacer les données manquantes par des valeurs de substitution, permettant ainsi d'obtenir un ensemble de données plus complet pour l'analyse. L'exemple de Forteza et García-Uribe (2025) cité plus haut montre cependant l'intérêt d'un traitement manuel dans les enquêtes « ménages » pour des variables dont la distribution est très asymétrique.

²⁴ Voir le lien : <https://www.insee.fr/fr/information/2406147>.

²⁵ Pour les statistiques, la granularité correspond à la finesse décrite par les « cases » des tableaux : par région, département, commune ? Par tranche d'âge de 5 ans, 10 ans ? Par secteur d'activité en 21 sections, 88 divisions, ou 615 classes ?

des procédures de vérification sélective des données (*selective editing*). Ainsi, les équipes en charge du dispositif Ésane²⁶ ont mis en œuvre une telle méthode, à l'appui du calcul de scores permettant de caractériser la criticité des variables à contrôler.

En pratique, pour les variables à contrôler in fine, le travail des gestionnaires d'enquête va souvent consister à rappeler les unités enquêtées, afin de compléter ou corriger la réponse initiale, ou de mieux comprendre les données transmises et s'assurer qu'il n'y a pas d'erreur. Ce retour à l'enquêté n'existe pas pour les ménages. D'abord parce que l'interaction avec l'enquêté s'effectue à la source, lors du questionnement. Mais aussi parce que l'univers des ménages est beaucoup plus homogène, de sorte qu'une erreur est moins préjudiciable pour la qualité des résultats de l'enquête.

Données administratives : itérations avec l'administration et transformations

Il peut paraître étrange de parler de temps passé à la « collecte » dans le cas des données administratives. En effet, on pourrait considérer que, la convention ayant été écrite, il suffit de l'appliquer pour récupérer les données, avant de passer à la tabulation

proprement dite. Le temps nécessaire serait donc, sinon nul, du moins négligeable. Ce serait méconnaître la réalité !

Les données administratives ont une finalité bien précise. Elles servent avant tout à faire fonctionner des processus opérationnels, à piloter, ou à prendre des décisions. Elles possèdent donc une structure liée à ces objectifs.

Les données administratives ont une finalité bien précise. Elles servent avant tout à faire fonctionner des processus opérationnels, à piloter, ou à prendre des décisions. Elles possèdent donc une structure liée à ces objectifs. Le traitement qu'il faut leur faire subir pour les intégrer dans un processus statistique est très généralement conséquent, et la conception de celui-ci prend du temps. Il faut bien souvent un certain nombre d'allers-retours entre les administrations détentrices de ces données et la statistique publique avant d'arriver à identifier

les données pertinentes, puis les transformations qu'il faut leur imposer... afin de convertir une donnée à visée opérationnelle en une donnée statistique. Cela passe aussi par une phase de consolidation de la donnée visant à éliminer les doublons, lesquels ne manquent pas d'apparaître dans des données à visée opérationnelle.

Par ailleurs, du côté de l'administration, l'**élaboration de la donnée administrative** peut se révéler assez longue, et c'est là un délai de disponibilité sur lequel la statistique publique n'a aucune prise. Ainsi, les données fiscales sur les revenus des ménages ne sont disponibles qu'un an après la fin de la période de référence. En effet, pour mesurer le revenu des ménages dans l'année N, il faut d'abord que ces derniers réalisent leur déclaration de revenus au printemps N+1. Vient ensuite le traitement de l'information par l'administration fiscale, qui inclut les correctifs liés aux déclarations tardives et aux contrôles. Celui-ci dure jusqu'à la fin de l'année N+1.

²⁶ Élaboration des statistiques annuelles d'entreprises. Voir le lien : <https://www.insee.fr/fr/metadonnees/source/serie/s1188>.

De ce fait, la statistique publique dépend du processus de recueil de la donnée administrative. Elle peut parfois décider de recourir à des données plus provisoires, au prix d'une qualité moindre de l'information. C'est par exemple ce qui a été fait pour les données sur les revenus intégrées dans le dispositif *Statistiques sur les revenus et les coûts (SRVC)*²⁷, partie française de EU-SILC. En effet, afin d'atteindre les nouvelles exigences en matière de fraîcheur demandées par Eurostat, il a été décidé de recourir à une information fiscale moins consolidée, de manière à gagner quelques mois dans le processus d'élaboration de la donnée statistique.

Néanmoins, l'essentiel de l'effort va porter sur les travaux méthodologiques nécessaires à la construction du processus de **transformation de la donnée opérationnelle en donnée statistique** (Koumarianos et Rivière, 2025). L'objectif est d'adapter l'information collectée par l'administration aux concepts et au cadre de référence de la statistique publique. Ainsi, dans le *Fichier démographique sur les logements et les individus (Fidéli)*²⁸ (Lamarche et Lollivier, 2021), l'identification des résidences principales fait appel à des concepts issus du recensement de population, définis par l'Insee (et plus ou moins facilement mis en œuvre dans les processus de collecte dont il a la maîtrise). Mais la source fiscale, elle, s'appuie sur une notion de résidence principale très adhérente à la problématique du recouvrement de l'impôt. Toute la difficulté réside donc dans la capacité à adapter l'information sur la localisation des individus dans la source fiscale pour se rapprocher le plus possible du concept statistique.

Données privées : un cadre plus contraignant pour la transformation des données

Pour les données privées, la situation est très similaire à celle des données administratives. Ce qui change, c'est l'exigence accrue de confidentialité des données. Le statisticien public est souvent conduit à travailler directement sur un poste de travail de l'entreprise concernée, car celle-ci n'envoie pas de fichier de données individuelles, même anonymisées. C'est le cas par exemple pour les données de comptes bancaires (Bonnet et Loisel, 2024). Les entreprises peuvent aussi souhaiter que les statisticiens n'accèdent pas directement aux données des personnes, mais uniquement à des données transformées, comme pour les données de téléphonie mobile (Joubert, 2025). Le développement d'un *pipeline* conçu en collaboration entre l'entreprise et les statisticiens publics – pour s'entendre sur la méthode de transformation – est alors particulièrement crucial (Lesur, 2025).

Nouvel usage implique nouvelle convention.

Le respect rigoureux des conventions constitue également une contrainte forte : on ne peut utiliser les microdonnées²⁹ hors du cadre invoqué dans la convention simplement parce qu'on considère que c'est intéressant. Nouvel usage implique nouvelle convention. Ce fut le cas, par exemple, pour les données de caisse : si ces dernières contribuent de manière réglementaire au calcul de l'indice des prix à

²⁷ Voir le lien : <https://www.insee.fr/fr/metadonnees/source/serie/s1220>.

²⁸ Voir le lien : <https://www.insee.fr/fr/metadonnees/source/serie/s1019>.

²⁹ Données individuelles, non agrégées : par exemple, des données par ménage, et non par catégorie de ménage, ou des données par entreprise, et non par secteur d'activité.

la consommation (Leclair 2019), elles peuvent être utiles également pour estimer la consommation des ménages, ou encore le chiffre d'affaires des grandes enseignes de distribution.

► La phase finale de traitement statistique

Corriger les processus de sélection, en particulier dans les enquêtes

Pour les enquêtes, à l'issue de l'étape de collecte, on dispose de microdonnées « originelles », qu'on qualifie souvent de « brutes »³⁰, et qui ne sont pas les microdonnées finales. En particulier, les données présentent de manière très générale un certain nombre de manques, qu'il va falloir redresser de manière statistique.

**Pour les enquêtes,
à l'issue de l'étape
de collecte, les données
présentent de manière
très générale un certain
nombre de manques,
qu'il va falloir redresser
de manière statistique.**

Le premier d'entre eux vient du fait que les enquêtes portent le plus souvent sur des **échantillons**, qui doivent donc être pondérés pour être « représentatifs » de la population dans laquelle ils sont sélectionnés (au sens de refléter au final sa structure et ses effectifs).

Cependant, à la sélectivité maîtrisée par le statisticien qu'est l'échantillonnage, s'ajoutent d'autres phénomènes de sélection, tels que la **non-réponse totale**. Le fait qu'un ménage ou une entreprise refuse de répondre, en dépit de l'obligation légale, est une

réalité avec laquelle la profession statistique doit composer. Il faut alors comprendre le mécanisme de non-réponse – les déterminants observables de l'absence de réponse – et incorporer ces déterminants dans le calcul des pondérations. Cette étape est essentielle pour préserver la « représentativité » de l'échantillon de répondants (au sens de la capacité à reconstruire une image fidèle de la population). La compréhension du mécanisme de « non-réponse » ne peut pas se faire a priori : dans bien des situations, ce phénomène reste difficile à anticiper et nécessite une expertise statistique.

Remplir les « trous », une phase potentiellement fastidieuse

Il faut ensuite faire de même pour ce qu'on appelle la **non-réponse partielle**. Un ménage ou une entreprise échantillonnée – même en ayant accepté de répondre – peut pour une raison ou une autre ne pas fournir de réponse à une question particulière : par choix, parce que la personne interrogée ne sait pas, ou parce que la question ne la concerne pas. Ce phénomène est généralement traité par imputation. On va chercher à comprendre les déterminants de la non-réponse, en étudiant la corrélation entre la variable non collectée et les autres. Il s'agit ensuite d'attribuer la valeur la plus probable au regard des informations connues pour l'unité enquêtée. Le travail peut être très important, non

³⁰ Même si le terme est inadéquat, car les données sont toujours une construction, et ne sont en réalité jamais « brutes » (Gitelman, 2013).

seulement en matière de modélisation des mécanismes de non-réponse, mais aussi selon l'exigence d'exhaustivité de la correction. En effet, l'ensemble des variables collectées peut être affecté par la non-réponse partielle. Il existe des méthodes bayésiennes³¹ pour redresser de manière systématique tout un corpus de variables (Rubin, 1987), mais celles-ci n'exonèrent pas d'une analyse souvent fastidieuse des mécanismes de non-réponse affectant chacune des variables.

La situation est très similaire pour les enquêtes et pour les données administratives ou privées. Simplement, les « données de base » ne sont pas issues d'une collecte, mais, comme on l'a vu, d'un processus de récupération de données du système d'information, et de transformations. Néanmoins, les « traitements statistiques » sont les mêmes.

Le *fine tuning*³² des pondérations des données

Reste enfin, pour les enquêtes, le « calage sur marges », qui est un processus permettant de modifier légèrement les pondérations de manière à retrouver dans l'échantillon de répondants des grandeurs connues (comme le nombre d'individus). Si la loi des grands nombres et la théorie du sondage assurent que ces grandeurs seront approchées de manière satisfaisante, il peut être efficace de caler l'échantillon sur telle ou telle grandeur, de manière à améliorer la précision de l'échantillon sur les variables d'intérêt de l'enquête. Ce processus de calage peut lui aussi être fastidieux et impliquer un nombre élevé de variables. Pour les statistiques d'entreprises, se posent également des problèmes spécifiques liés aux unités à fort impact sur les agrégats publiés, nécessitant un travail approfondi sur les pondérations.

À première vue, l'ensemble de ces traitements (contrôle, imputation, repondération) se ramène finalement à une transformation. Dans l'idéal, pour gagner du temps, on pourrait imaginer de l'automatiser, c'est-à-dire de développer un *pipeline*. Une fois conçu et mis au point, celui-ci n'aurait plus qu'à être appliqué de manière systématique, comme tout traitement de données mis au point par un *data scientist* (Comte et al., 2022). Mais de tels travaux exigent une analyse qui peut s'avérer chronophage, selon le niveau d'exigence en matière de complétude et de cohérence. Ils peuvent ainsi requérir un temps significatif lorsqu'ils sont mis en œuvre. Il n'en demeure pas moins qu'un effort en matière de standardisation des outils et des pratiques est déjà un gage d'optimisation.

Exigences de cohérence et limites de l'automatisation

Attardons-nous sur les exigences de cohérence qui pèsent sur les données produites par la statistique publique, afin de mieux comprendre les obstacles à une automatisation complète.

On pense d'abord aux « **microcontrôles** », dans lesquels on s'assure que les données relatives à une même unité statistique ne présentent pas d'incohérences manifestes entre elles. On a déjà évoqué, dans le cas des enquêtes auprès des entreprises, l'enjeu

31 Méthodes s'appuyant sur le théorème de Bayes, qui permettent d'estimer la probabilité de phénomènes à partir de la connaissance de certaines informations.

32 Réglage fin.

“ Les exigences de cohérence qui pèsent sur les données produites par la statistique publique rendent difficiles une automatisation complète des traitements.

pour juger de la qualité des statistiques et pour les qualifier également d'officielles. C'est d'ailleurs le quatorzième principe du Code de bonnes pratiques de la statistique européenne. Toute source d'incohérence peut être une raison légitime de discrédit des chiffres produits.

Or, les exigences de cohérence peuvent porter sur des domaines de diffusion³³ qui peuvent être très fins. En effet, non seulement la statistique publique publie ses propres analyses, mais elle met également les microdonnées à disposition du monde de la recherche. Ainsi, les statisticiens doivent s'assurer en amont, autant que possible, que les données produites peuvent être utilisées sans contredire de manière rédhibitoire d'autres statistiques publiées – ou à tout le moins comprendre et fournir des éléments d'explication aux contradictions apparentes.

Par exemple, les sources statistiques contenant de l'information sur les revenus sont multiples. En effet, pour de nombreuses enquêtes « ménages », il s'agit d'une dimension d'analyse très importante, soit parce qu'elle est au cœur des thématiques de l'enquête, soit parce que c'est une source d'explication déterminante pour les phénomènes auxquels elle s'intéresse. Il faut alors s'assurer que les chiffrages implicitement fournis

par chacune de ces enquêtes sont cohérents avec le dispositif de référence sur le revenu³⁴.

“ Même avec un pipeline automatisant les contrôles de cohérence, la majorité du temps passé consistera à analyser les résultats, à comprendre d'où viennent les éventuelles incohérences et à déterminer sur quoi intervenir.

De la même manière, il faut s'assurer d'une certaine cohérence entre information macroéconomique (issue des comptes nationaux) et information microéconomique (issue des enquêtes). Cette exigence s'est renforcée ces quinze dernières années, à la suite du rapport de la commission Stiglitz-Sen-Fitoussi sur la mesure de la performance économique et du progrès social (Stiglitz et al., 2009). Ce rapport insiste en effet sur la nécessité, pour la statistique publique, de mesurer à la fois les agrégats économiques et leur distribution au sein de la population.

³³ Regroupements au niveau desquels les données sont agrégées pour être diffusées, i.e. les « cases » de tableaux évoquées plus haut : par exemple, regroupement selon le croisement département / tranche d'âge.

³⁴ Qui dans ce cas précis est l'enquête Revenus Fiscaux et Sociaux (ERFS).

Ainsi, même si on a réalisé un *pipeline* automatisant les contrôles de cohérence, la majorité du temps passé consistera à en analyser les résultats, à essayer de comprendre d'où proviennent les éventuelles incohérences et à déterminer sur quoi intervenir : enlever une unité douteuse ou hors champ, compléter le champ, revoir des modalités de codage, etc. C'est là un travail de fourmi, et chercher à trop réduire cette durée ferait courir des risques concernant la qualité.

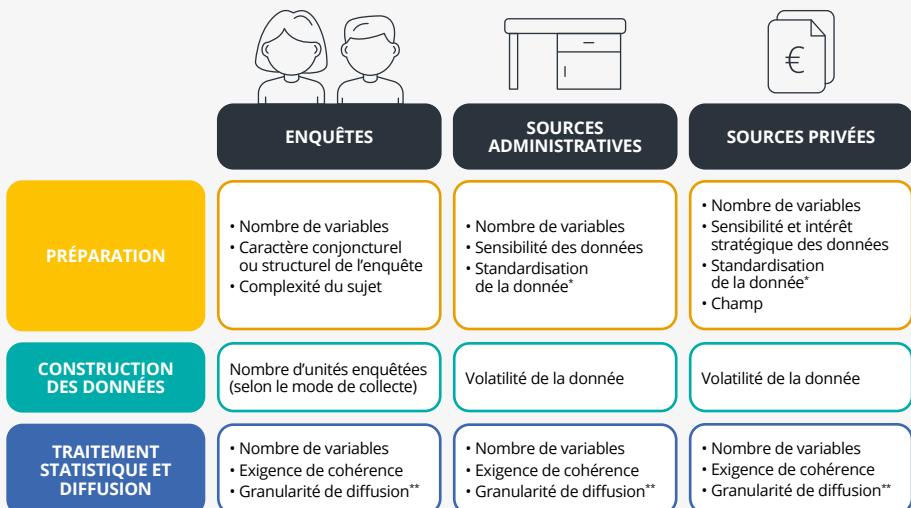
► Quels enseignements tirer ?

L'explication du temps nécessaire à l'élaboration de statistiques de qualité, vue comme une fonction de plusieurs variables explicatives, est paradoxalement difficile à fonder sur des données (i.e. difficile à modéliser). Comme le souligne Parkinson (1955), « le travail s'étale de façon à occuper le temps disponible pour son achèvement »³⁵. En d'autres termes, si on nous octroie plus de temps pour une tâche, on peut avoir tendance à remplir le temps en question. Une mesure fondée sur des temps observés est donc sujette à caution. C'est la raison pour laquelle le présent article tente plutôt de raisonner de façon qualitative ou par ordre de grandeur.

Dès lors, quels sont les facteurs explicatifs de ce temps (*figure 4*) ?

- Lors de la **préparation**, ce sont la complexité du sujet, son caractère polysémique, mais aussi le manque de standardisation des données, la sensibilité des données et les délais de livraison propres au fournisseur, en cas de sources externes. Pour une

► **Figure 4 - Facteurs influant sur le temps de production des statistiques publiques**



* Transformation pour adapter les données aux concepts et au cadre de référence de la statistique publique.

** Niveau de détail le plus fin des données diffusées.

³⁵ L'analyse de l'auteur est connue sous le nom de « loi de Parkinson ».

opération régulière, dès la seconde opération, on va clairement gagner du temps de compréhension des données, sauf si la réglementation évolue significativement. Les données administratives ou privées peuvent s'avérer très chronophages car il s'agit de s'assurer, avant transformation, que l'on comprend bien le sens et les limites des données.

- **La construction des données**, lorsqu'il s'agit d'enquêtes, est sans doute la phase la plus maîtrisable dans un institut statistique, où les procédures de collecte sont très organisées. Le nombre d'unités statistiques est alors décisif, ainsi que la complexité du questionnement. Cela dépend du protocole : pour les enquêtes par Internet, donc sans l'intermédiation d'un enquêteur, ce temps de collecte sera court. Cependant, on « paie » ce gain de temps ultérieurement, par une étape méthodologique supplémentaire : prise en compte de la baisse du taux de réponse et du biais de sélection, harmonisation avec les résultats issus d'autres modes de collecte. Car pour produire des statistiques publiques, il est aujourd'hui difficile d'imaginer collecter de l'information exclusivement par Internet. Pour la construction de données issues de sources externes, un *pipeline* de traitement automatique, répllicable, et un environnement facilitant sa mise au point améliorent les choses. Mais dans tous les cas, enquêtes ou sources externes, il faut déjà consacrer des ressources aux vérifications : par les enquêteurs, par les gestionnaires et par les statisticiens en charge du dispositif, par exemple, s'il faut rappeler des entreprises parce que leurs réponses paraissent incohérentes.

- **Lors du traitement statistique**, on pourra certes réaliser un autre *pipeline*, menant cette fois des microdonnées aux statistiques, qui peut permettre de gagner un peu de temps par la systématisation des opérations et par la mutualisation plus ou moins forte des outils. Mais il sera inévitablement interrompu par des interventions manuelles, car on devra en permanence s'assurer de la cohérence des résultats. Pour chaque variable, pour chaque domaine de diffusion, on devra regarder s'il y a une rupture de série ou s'il existe des écarts significatifs avec d'autres statistiques et, si oui, chercher à comprendre pour quelle raison. La statistique n'est pas produite seule, elle s'inscrit dans un environnement, c'est une pièce dans un puzzle. Il faut pouvoir interpréter, donner un sens, et plus on a de variables et de strates de diffusion, plus longues et nombreuses sont les vérifications.

Avec la prégnance de l'activité de vérification, et le temps passé dans les boucles de rétroaction (*feedback loops*) qui lui sont associées, il paraît illusoire de gagner beaucoup

de temps « à exigences constantes ». En jouant sur les exigences, en revanche, « si l'enjeu du temps est particulièrement important », on peut trouver un compromis : en réduisant le nombre de variables, de questions, ou bien en produisant rapidement des statistiques sur un petit nombre d'entre elles. On peut aussi imposer des limites sur les usages à granularité de diffusion fine. D'une façon ou d'une autre, gagner du temps, c'est accepter de « perdre » sur l'une ou l'autre des dimensions...

“
Avec la prégnance de l'activité de vérification, et le temps associé, il paraît illusoire de gagner beaucoup de temps « à exigences constantes ».
”

► Bibliographie

- ANXIONNAZ, Isabelle et MAUREL, Françoise, 2021. Le Conseil national de l'information statistique – La qualité des statistiques passe aussi par la concertation. In : *Courrier des statistiques*. [en ligne]. 8 juillet 2021. Insee. N° N6, pp. 123-142. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/5398693?sommaire=5398695>.
- ATKINSON, Anthony B., 1975. *The Economics of Inequality*. Mai 1975. Clarendon Press, Oxford University Press. ISBN 0-19-877076-6.
- BECK, François, CASTELL, Laura, LEGLEYE, Stéphane et SCHREIBER, Amandine, 2022. Le multimode dans les enquêtes auprès des ménages : une collecte modernisée, un processus complexifié. In : *Courrier des statistiques*. [en ligne]. 20 janvier 2022. Insee. N° N7, pp. 7-28. [Consulté le 6 octobre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/6035934?sommaire=6035950>.
- BÉNICHOU, Yves-Laurent, ESPINASSE, Lionel et GILLES, Séverine, 2023. Le code statistique non signifiant (CSNS) : un service pour faciliter les appariements de fichiers. In : *Courrier des statistiques*. [en ligne]. 30 juin 2023. Insee. N° N9, pp. 64-85. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/7635825?sommaire=7635842>.
- BOITTELLE, Mathieu, CUPILLARD, Émilie, JACQUOT, Alain, JOUBERT, Marie-Pierre et LE GOFF Florian, 2025. Les données de transaction par carte bancaire CB – Quels apports possibles aux analyses conjoncturelles et territoriales ? In : *Courrier des statistiques*. [en ligne]. 23 juin 2025. Insee. N° N13, pp. 119-142. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/8546947?sommaire=8546949>.
- BONNANS, Dominique, 2019. RMéS, le référentiel de métadonnées statistiques de l'Insee. In : *Courrier des statistiques*. [en ligne]. 27 juin 2019. Insee. N° N2, pp. 46-57. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/4168396?sommaire=4168411>.
- BONNET, Odran et LOISEL, Tristan, 2024. L'économie racontée par les données bancaires – Ce que nos relevés de comptes disent de nous. In : *Courrier des statistiques*. [en ligne]. 16 décembre 2024. Insee. N° N12, pp. 115-136. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/8264558?sommaire=8264562>.
- CARON, Nathalie, 2025. Le *data editing* : définition et principes généraux. In : *Documents de travail*. [en ligne]. 20 octobre 2025. Insee. N° M2025/06. [Consulté le 3 novembre 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/statistiques/8658331>.
- CERRONI, Fulvia, DI BELLA, Grazia et GALIÈ, Lorena, 2014. Evaluating administrative data quality as input of the statistical production process. In : *Rivista di statistica ufficiale*. [en ligne]. Octobre 2014. Istat. Volume 16, N° 1-2, pp. 117-146. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://www.istat.it/produzione-editoriale/rivista-di-statistica-ufficiale-1-22014/>.

- CHRISTINE, Marc et ROTH, Nicole, 2020. Le Comité du label – Un acteur de la gouvernance au service de la qualité des statistiques publiques. In : *Courrier des statistiques*. [en ligne]. 31 décembre 2020. Insee. N° N5, pp. 39-52. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/5008698?sommaire=5008710>.
- COMTE, Frédéric, DEGORRE, Arnaud et LESUR, Romain, 2022. Le SSPCloud : une fabrique créative pour accompagner les expérimentations des statisticiens publics. In : *Courrier des statistiques*. [en ligne]. 20 janvier 2022. Insee. N° N7, pp. 68-85. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/6035940?sommaire=6035950>.
- COTTON, Franck et HAAG, Olivier, 2023. L'intégration des données administratives dans un processus statistique – Industrialiser une phase essentielle. In : *Courrier des statistiques*. [en ligne]. 30 juin 2023. Insee. N° N9, pp. 104-125. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/7635829?sommaire=7635842>.
- De WAAL, Ton, PANNEKOEK, Jeroen et SCHOLTUS, Sander, 2011. *Handbook of Statistical Data Editing and Imputation*. Janvier 2011. John Wiley & Sons, Inc.. ISBN 9780470542804.
- ELBAUM, Mireille, 2018. Les enjeux des nouvelles sources de données. In : *Chroniques du Cnis*. [en ligne]. Septembre 2018. N° 16. [Consulté le 15 septembre 2025]. Disponible à l'adresse : <https://www.cnis.fr/publications/liste-des-chroniques-du-cnis/>.
- EUROSTAT, 2025. Short-stay accommodation offered via online collaborative economy platforms. In : *site d'Eurostat*. [en ligne]. Juin 2025. Statistics Explained. [Consulté le 24 octobre 2025]. Disponible à l'adresse :
<https://ec.europa.eu/eurostat/statistics-explained/index.php?oldid=545642>.
- FELLEGI, Ivan P. and HOLT, Daniel T., 1976. A Systematic Approach to Automatic Edit and Imputation. In : *Journal of the American Statistical Association*. Mars 1976. Taylor & Francis Ltd.. Volume 71, n° 353, pp. 17-35.
- FERMOR-DUNMAN, Verena et PARSONS, Laura, 2022. Data Acquisition Processes Improving Quality of Microdata at the Office for National Statistics. In : *European Conference on Quality in Official Statistics (Q2022)*. [en ligne]. 8 juin – 10 juin 2022. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://q2022.stat.gov.lt/scientific-information/papers-presentations/session-25>.
- FORTEZA, Nicolás et GARCÍA-URIBE, Sandra, 2025. A Score Function to Prioritize Editing in Household Survey Data: A Machine Learning Approach. In : *Journal of Official Statistics*. [en ligne]. Mars 2025. Volume 41, n° 1, pp. 144-171. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://journals.sagepub.com/doi/10.1177/0282423X241309971>.
- GITELMAN, Lisa, 2013. *"Raw Data" Is an Oxymoron*. The MIT Press. ISBN 978-0-262-51828-4.

- GRANQUIST, Leopold, 1997. The New View on Editing. In : *International Statistical Review*. [en ligne]. Décembre 1997. International Statistical Institute. Volume 65, n° 3, pp. 381-387. [Consulté le 15 septembre 2025]. Disponible à l'adresse : <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1751-5823.1997.tb00315.x>.
- GRANQUIST, Leopold et KOVAR, John G., 1997. Editing of Survey Data: How Much Is Enough?. In : *Survey Measurement and Process Quality*. Février 1997. John Wiley & Sons, Inc.. Pp. 415-435. ISBN 978-0-471-16559-0.
- HAND, David J., 2018. Statistical Challenges of Administrative and Transaction Data. In : *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. [en ligne]. Juin 2018. Volume 181, n° 3, pp. 555-605. [Consulté le 15 septembre 2025]. Disponible à l'adresse : <https://academic.oup.com/jrssa/article/181/3/555/7072022?login=false>.
- HUMBERT-BOTTIN, Élisabeth, 2018. La déclaration sociale nominative – Nouvelle référence pour les échanges de données sociales des entreprises vers les administrations. In : *Courrier des statistiques*. [en ligne]. 6 décembre 2018. Insee. N° N1, pp. 25-34. [Consulté le 15 septembre 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/information/3647025?sommaire=3647035>.
- INSEE, 1998. Mesurer la pauvreté aujourd'hui. In : *Économie et statistique*. [en ligne]. Octobre 1998. N° 308-310. [Consulté le 15 septembre 2025]. Disponible à l'adresse : https://www.persee.fr/issue/estat_0336-1454_1998_num_308_1?sectionId=estat_0336-1454_1998_num_308_1_2588.
- JOUBERT, Marie-Pierre, 2025. Les données de téléphonie mobile – Une source de connaissance sur la population et ses déplacements. In : *Courrier des statistiques*. [en ligne]. 23 juin 2025. Insee. N° N13, pp. 95-118. [Consulté le 15 septembre 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/information/8546945?sommaire=8546949>.
- KOUAMARIANOS, Heidi, RIVIÈRE, Pascal, 2025. Statistiques fondées sur des données administratives : esquisse d'un cadre général. In : *Documents de travail*. [en ligne]. 23 juin 2025. Insee. N° M2025/03. [Consulté le 15 septembre 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/statistiques/8596376>.
- LAMARCHE, Pierre et LOLLIVIER, Stéfan, 2021. Fidéli, l'intégration des sources fiscales dans les données sociales. In : *Courrier des statistiques*. [en ligne]. 8 juillet 2021. Insee. N° N6, pp. 28-46. [Consulté le 15 septembre 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/information/5398683?sommaire=5398695>.
- LAWRENCE, David, et McKENZIE, Richard, 2000. The General Application of Significance Editing. In : *Journal of Official Statistics*. [en ligne]. Janvier 2000. Volume 16, n° 3, pp. 243-253. [Consulté le 15 septembre 2025]. Disponible à l'adresse : https://www.researchgate.net/publication/256199660_The_General_Application_of_Significance_Editing.
- LECLAIR, Marie, 2019. Utiliser les données de caisses pour le calcul de l'indice des prix à la consommation. In : *Courrier des statistiques*. [en ligne]. 19 décembre 2019. Insee. N° N3, pp. 61-75. [Consulté le 15 septembre 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/information/4254225?sommaire=4254170>.

- LEFEBVRE, Olivier, 2024. Le Répertoire Statistique des Individus et des Logements (Résil) – Un nouvel univers de référence pour les statistiques démographiques et sociales. In : *Courrier des statistiques*. [en ligne]. 8 juillet 2024. Insee. N° N11, pp. 73-94. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/8203040?sommaire=8203072>.
- LESUR, Romain, 2025. Sources de données privées : panorama et perspectives. In : *Courrier des statistiques*. [en ligne]. 23 juin 2025. Insee. N° N13, pp. 73-94. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/8546943?sommaire=8546949>.
- PANNEKOEK, Jeroen, SCHOLTUS, Sander, et VAN DER LOO, Mark, 2013. Automated and Manual Data Editing: a View on Process Design and Methodology. In : *Journal of Official Statistics*. [en ligne]. Novembre 2013. Volume 29, n° 4, pp. 511-537. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://journals.sagepub.com/doi/abs/10.2478/jos-2013-0038>.
- PARKINSON, Cyril Northcote, 1955. Parkinson's Law. In : *The Economist*. 19 novembre 1955. The Riverside Press. Volume 19.
- RENNE, Catherine, 2018. Bien comprendre la déclaration sociale nominative pour mieux mesurer. In : *Courrier des statistiques*. [en ligne]. 6 décembre 2018. Insee. N° N1, pp. 35-44. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/3647029?sommaire=3647035>.
- RUBIN, Donald B., 1987. *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons Inc.. ISBN 0-471-65574-0.
- STIGLITZ, Joseph E., SEN, Amartya et FITOUSSI, Jean-Paul, 2009. *Vers de nouveaux systèmes de mesure – Performances économiques et progrès social*. Odile Jacob. ISBN 978-2-7381-2463-0.
- TAVERNIER, Jean-Luc, 2020. Fonctionnement de l'Insee dans la période de confinement. In : *Courrier des statistiques*. [en ligne]. 31 décembre 2020. Insee. N° N5, pp. 6-20. [Consulté le 15 septembre 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/5008681?sommaire=5008710>.

Des démarches qualité pour sécuriser et optimiser la production de statistiques publiques



Pierrette Briant* et Élodie Kranklader**

La qualité des statistiques publiques est une notion multidimensionnelle : elle prend en compte non seulement les produits statistiques (données, cartes, publications, etc.), mais aussi leur processus de production et l'environnement dans lequel ils sont élaborés. Le code de bonnes pratiques de la statistique européenne fournit un guide pour réguler la qualité des statistiques publiques en Europe ; l'Insee l'applique à l'ensemble de ses statistiques.

Pour évaluer la qualité, plusieurs dispositifs existent, institutionnels, mais aussi opérationnels avec notamment les « démarches qualité ». Fondées sur le principe d'amélioration continue, celles-ci permettent aux équipes de définir et mettre en œuvre des actions pour sécuriser et améliorer régulièrement leurs processus. Elles conduisent à analyser le contexte de production, révélant souvent la nécessité de mieux connaître les besoins des utilisateurs et de sécuriser la réception de données externes, de plus en plus prégnantes dans la production statistique. Elles fournissent une vue d'ensemble structurée des processus, accessible à tous les agents du service statistique public, et mettent en lumière les points forts et axes d'amélioration. Elles recensent aussi les risques auxquels les processus sont exposés (sur la qualité des produits, le respect des échéances, etc.). L'approche collective adoptée permet d'échanger sur les difficultés de chacun et d'établir un plan d'actions consensuel.

Des travaux équivalents sont menés dans d'autres pays, faisant ressortir ces mêmes bénéfices, à condition que les plans soient suivis et actualisés.

 The quality of official statistics is a multidimensional concept: it takes into account not only the statistical outputs themselves (data, maps, publications, etc.) but also their production processes and the environment in which they are produced. The European Statistics Code of Practice provides a framework for regulating the quality of official statistics in Europe; Insee applies it to all of its statistical operations.

Several mechanisms are in place for quality assessment, both at the institutional level and the operational level, notably through quality reviews. Based on the principle of continuous improvement, these approaches allow teams to define and implement actions to regularly secure and enhance their processes. They involve analysing the production context, often revealing the need to better understand user needs and secure the reception of external data, which are increasingly critical in statistical production. These approaches provide a structured overview of processes, accessible to all official statistical service staff, and highlight strengths as well as areas for improvement. They also identify the risks to which the processes are exposed (regarding product quality, meeting deadlines, etc.). The collective approach adopted fosters discussion about each actor's challenges and helps establish a consensual action plan.

Similar initiatives are carried out in other countries, highlighting the same benefits, provided that they are subsequently monitored and updated.

* Cheffe de la division Expertise qualité, DMCSI, Insee.
pierrette.briant@insee.fr

** À la date de rédaction de l'article, experte qualité, DMCSI, Insee.
elodie.kranklader@insee.fr

Nombre d'habitants en France, taux d'inflation, prévisions de croissance, taux de chômage et d'emploi, taux d'innovation des entreprises... Comment être assuré de la qualité de ces statistiques publiques et des nombreuses autres produites par l'Insee ?

Au-delà de la simple précision des indicateurs, il faut d'abord s'entendre sur ce que la qualité recouvre en matière de statistiques publiques. À travers les thématiques qu'elles abordent, les enquêtes de satisfaction menées par l'Insee auprès de ses utilisateurs donnent une première idée des dimensions qui entrent en jeu. Très majoritairement, l'institut inspire confiance, et les personnes interrogées considèrent qu'il est utile, qu'il produit et diffuse des informations aisément compréhensibles, fiables, faciles à trouver, indépendantes du pouvoir politique (Hillaireau et Zana Rouquette, 2024). Comme l'illustre la variété des questions posées, la qualité revêt de nombreuses facettes et est liée à des attentes, des usages, qu'il est donc important de connaître. Ces enquêtes ne mesurent cependant que la perception globale qu'ont les utilisateurs des productions de l'institut. Il faut aussi s'assurer, dans les faits, que ces productions se réalisent dans les meilleures conditions et aboutissent bien, in fine, à des statistiques de bonne qualité. Dans cet objectif, un cadre institutionnel spécifique régule le fonctionnement de l'Insee, et plus largement celui du service statistique public (SSP)¹. L'institut s'organise également en interne et mène une politique proactive en faveur de la qualité.

► Qu'est-ce que la qualité ? —

La qualité, une notion multidimensionnelle orientée vers les utilisateurs

Les critères énoncés dans les enquêtes de satisfaction renvoient à la **qualité observable des statistiques produites**. Ils reprennent largement une définition arrêtée en 1995 par le [système statistique européen²](https://ec.europa.eu/eurostat/fr/web/european-statistical-system/overview). Cette dernière, intégrée dans l'article 12 du règlement

n° 223/2009 relatif aux statistiques européennes³, s'articule autour de sept critères perceptibles par les utilisateurs ou interprétables grâce à la documentation méthodologique : pertinence, exactitude, actualité, ponctualité, accessibilité et clarté, comparabilité, cohérence.

Or, depuis les travaux d'Edwards Deming⁴, il est établi que la qualité des produits dépend aussi de la manière dont ils sont élaborés, autrement dit de la **qualité de leur processus de production**.

Cette dernière est difficilement observable par les

utilisateurs. Par exemple, dans le cas des statistiques, elle passe par des garanties sur la fourniture des données sources, associées à des moyens humains et informatiques adéquats. Ce sont en effet des prérequis pour maîtriser la ponctualité, mettre en œuvre

¹ Le service statistique public (SSP) français est composé de l'Insee et de 16 services statistiques ministériels (SSM) qui réalisent les opérations statistiques dans leur domaine de compétence.

² <https://ec.europa.eu/eurostat/fr/web/european-statistical-system/overview>.

³ Voir les références juridiques en fin d'article.

⁴ Statisticien américain qui se pencha, dans les années 1930, sur les contrôles de qualité dans l'industrie (Desrosières, 2003).

une méthodologie qui respecte les règles de l'art, etc. Dans le domaine des technologies, du management et de la production de biens et services, l'**Organisation internationale de normalisation**⁵ (ISO dans son acronyme anglais, pour *International Organization for Standardization*) a formalisé une telle approche de la qualité en construisant des normes⁶, qui constituent des garanties pour les usagers.

Quant à la question de l'indépendance professionnelle, abordée dans les enquêtes de satisfaction, elle renvoie à la **qualité du contexte** dans lequel les statistiques sont produites. Celle-ci dépend essentiellement de l'environnement institutionnel.

La qualité globale d'un produit est donc une notion relative à des usages et multidimensionnelle. Elle est précisée par des critères qui sont rassemblés dans un référentiel. Pour les statistiques publiques européennes, ce référentiel est le **code de bonnes pratiques de la statistique européenne**⁷ (CoP dans son acronyme anglais, pour *Code of Practice*). Il allie les trois dimensions évoquées précédemment : produits statistiques, processus statistiques et cadre institutionnel. Il s'applique aux autorités statistiques du système statistique européen, qui comprennent à la fois l'autorité statistique de l'Union européenne (Eurostat)⁸, les instituts nationaux de statistique (INS) et d'autres autorités nationales⁹.

Un référentiel et des dispositifs pour évaluer la qualité des statistiques publiques

En France, le SSP, dont l'Insee, a choisi d'étendre l'application du CoP à l'ensemble des statistiques qu'il produit et diffuse, y compris celles dont l'élaboration ne relève pas des obligations européennes. Le **cadre d'assurance qualité du système statistique européen**¹⁰ fournit en outre une approche plus opérationnelle pour appliquer le CoP, à travers des recommandations de méthodes, d'outils et de bonnes pratiques.

Le CoP est un référentiel pour de nombreuses instances. Au niveau européen, bien sûr, il l'est pour les audits réalisés tous les 7 à 8 ans dans le cadre des revues par les pairs (Chaleix et Durr, 2024). En France, c'est aussi le référentiel de l'**Autorité de la statistique publique**¹¹. Dans ses avis et son rapport annuel sur les faits marquants du SSP, celle-ci interroge la qualité globale de notre organisation à l'aune de ces principes.

⁵ <https://www.iso.org/fr/home.html>.

⁶ Une norme définit des exigences, des spécifications, des lignes directrices ou des caractéristiques visant à s'assurer que les produits et services sont sûrs, fiables et de qualité. En France, l'Association française de normalisation (AFNOR) rédige des normes françaises (logo NF) et représente la France à l'ISO pour créer et réviser des normes internationales.

⁷ <https://ec.europa.eu/eurostat/web/quality/european-quality-standards/european-statistics-code-of-practice>.

⁸ Eurostat, l'Office statistique de l'Union européenne, est chargé de publier des statistiques et des indicateurs européens, permettant d'effectuer des comparaisons entre les pays et les régions.

⁹ En France, il y a 12 autres autorités nationales (en anglais, *other national authority*, ONA) : 10 des 16 SSM ainsi que le CépiDc, Centre d'épidémiologie sur les causes médicales de décès, et le CASD, Centre d'accès sécurisé aux données.

¹⁰ <https://ec.europa.eu/eurostat/documents/64157/4392716/ESS-QAF-V2.0-final.pdf>.

¹¹ <https://www.autorite-statistique-publique.fr>.

Tous examinent la qualité du fonctionnement de l'organisation statistique, mais pas ou peu celle des processus ou celle des produits. Or ce sont ces dernières dont les statisticiens sont responsables au quotidien. L'Insee a en effet adopté une organisation du travail qui responsabilise pleinement le concepteur d'une opération statistique, et ce jusqu'à la diffusion de ses résultats (Chaleix et Lefebvre, 2024).

Certains dispositifs spécialisés, tels que les audits métier d'Eurostat, sont plus en prise avec la réalité des processus. Ainsi, la France, comme les autres INS, élabore à l'attention d'Eurostat une documentation détaillée de la manière dont est calculé le revenu national brut¹², sachant que le calcul des contributions des États membres au budget de l'UE s'appuie principalement sur cet indicateur. Au niveau national, le **Comité du label de la statistique publique**¹³ vérifie à la fois la méthodologie, le protocole de collecte, le traitement post-collecte et la diffusion des projets d'enquêtes du SSP (Christine et Roth, 2020). À l'Insee, le Comité des investissements réalise également une analyse fine des processus pour des projets de création ou de modification d'opérations statistiques requérant des moyens supplémentaires. Quant aux homologations obligatoires de sécurité informatique sur les applications, par définition, elles observent la production statistique à travers son système d'information.

Ainsi, il existe plusieurs dispositifs d'évaluation de la qualité, portant sur l'organisation du SSP ou ses processus, qui orientent le travail de production statistique. Ils présentent l'avantage de fournir des éléments de preuve au public sur la qualité et les efforts fournis par le SSP pour l'améliorer encore.

Cependant, **aucun de ces dispositifs ne se penche en profondeur sur la mise en œuvre par les producteurs**. Par exemple :

- Quelle est la répartition des tâches au sein des équipes ? Comment y circule l'information ?

À l'Insee, pour évaluer la qualité des processus de production, des initiatives informelles ont progressivement laissé place à une organisation interne dédiée, au plus près des producteurs.

- Les besoins des utilisateurs, y compris internes, sont-ils bien recensés ?
- À quels aléas les processus sont-ils possiblement exposés et comment y faire face ou les anticiper ?
- Comment optimiser les processus dans le contexte de moyens contraints¹⁴ ?

À l'Insee, pour répondre à ce besoin, des initiatives informelles ont progressivement laissé place à une organisation interne dédiée, au plus près des producteurs de l'ensemble du

SSP. Dans la vision d'Eurostat, l'approche opérationnelle des démarches qualité portant sur les processus de production statistique complète l'approche institutionnelle plus globale que constituent les revues par les pairs (Bechev, 2024).

¹² https://www.insee.fr/fr/metadonnees/source/fichier/France_GNI_inventory.pdf.

¹³ <https://www.comite-du-label.fr/>.

¹⁴ Par exemple, à périmètre constant, les effectifs de l'Insee ont diminué d'un quart depuis 2006 (Madelin et Guéné, 2024).

► Évaluer la qualité des processus

L'introduction des démarches qualité à l'Insee, encouragée par l'Europe

La présence de protocoles de qualité a toujours été très forte à l'Insee. Lors de la première revue par les pairs européens organisée en France, en 2007, les auditeurs l'avaient reconnu. Ils avaient toutefois noté qu'il s'agissait d'initiatives décentralisées, qui n'étaient ni coordonnées, ni mutualisées. Par exemple, dans les années 1990, les équipes nationales et régionales en charge du répertoire d'identification des entreprises *Sirene*¹⁵ s'étaient mobilisées ensemble, en totale autonomie, pour définir comment améliorer sa qualité et le service rendu à ses utilisateurs.

Une première expérimentation de coordination a vu le jour au début des années 2000. Les producteurs des statistiques départementales d'emploi salarié avaient décidé d'adapter les contrôles réalisés sur les données. L'objectif était de raccourcir les délais de publication pour répondre à la demande des utilisateurs. Pour cette action, menée là encore aux niveaux national et régional, ils avaient bénéficié de l'appui de services transversaux, incluant notamment des méthodologues.

Les réorganisations de certains dispositifs ont constitué, également, des occasions pour chercher à améliorer la qualité. Ce fut le cas en 2009, lorsque l'Insee a remplacé les enquêtes annuelles d'entreprises par l'exploitation des liasses fiscales de la Direction générale des Finances publiques (DGFiP), combinée à une unique enquête structurelle annuelle allégée : l'*enquête sectorielle annuelle*¹⁶. En effet, de telles réorganisations offrent aux équipes concernées l'occasion de repenser et d'optimiser leurs dispositifs.

La professionnalisation des démarches qualité a été notamment impulsée par la deuxième revue par les pairs européens menée en France, en 2014.

En 2012, l'Insee s'est doté d'une unité transversale dédiée à la qualité : l'unité Qualité. C'est cette dernière qui pilote aujourd'hui les démarches qualité de processus telles qu'elles sont désormais pratiquées : fondées sur des méthodes harmonisées et un accompagnement spécifique des équipes. Cette professionnalisation des démarches qualité a été notamment impulsée par la deuxième revue par les pairs européens menée

en France, en 2014. Les auditeurs avaient alors demandé au SSP français de s'engager à formaliser davantage sa gestion de la qualité, relevant par exemple la nécessité d'homogénéiser la description des processus statistiques en suivant des normes internationales (voir infra). Pour l'Insee, c'était aussi une opportunité pour optimiser les processus dans une approche centrée sur les utilisateurs. Le pilotage par une unité transversale permet par ailleurs de mutualiser les meilleures pratiques.

¹⁵ <https://www.insee.fr/fr/metadonnees/source/serie/s1020>.

¹⁶ <https://www.insee.fr/fr/metadonnees/source/serie/s1269>.

Au même moment, l'institut a créé un département dédié aux méthodologies de traitement des données (tirage d'échantillon, correction de la non-réponse, économétrie, etc.). Ce dernier propose aux services producteurs une offre d'appui axée sur la qualité statistique des méthodes et des résultats, que l'unité Qualité complète par une offre axée sur la qualité du fonctionnement des processus.

Cette distinction entre, d'une part, la qualité de la méthodologie et, d'autre part, celle des processus n'est pas spécifique à la France. Par exemple, l'INS britannique conduit des démarches ciblées pour identifier les points de vulnérabilité des processus liés à leur organisation : systèmes difficiles à maintenir, boîtes noires non maîtrisées, rotations de personnel, etc. (Lound et Mais, 2022). D'autres INS allient qualité des méthodes et des processus dans leurs revues des productions statistiques. La Suède identifie par exemple les sources d'erreurs possibles et en évalue l'impact potentiel sur la qualité des données produites (Biemer et al., 2014). L'Irlande et la Norvège associent également des méthodologues à leurs examens des processus, ainsi que des spécialistes du code informatique pour les Irlandais (Forde et Moore, 2022) et des utilisateurs pour les Norvégiens (Berglund et Haugen, 2024).

Recenser les processus, objets d'étude des démarches qualité

Afin de mettre en œuvre des démarches qualité à l'Insee, l'unité Qualité s'est d'abord attachée à recenser les processus. Selon la norme internationale ISO 9000:2015¹⁷, « **un processus est un ensemble d'activités corrélées ou en interaction qui utilise des éléments d'entrée pour produire un résultat escompté** ». Cette définition laisse une grande latitude dans la finesse du découpage associé. L'Insee a ainsi considéré comme étant des processus des ensembles d'activités aussi larges que ceux liés à la production et à la diffusion du recensement de la population (RP) ou de l'indice des prix à la consommation (IPC). Il en est de même des ensembles d'activités nécessaires à la réalisation d'enquêtes auprès d'entreprises, différencier par leur thématique (innovation, technologies de l'information et de la communication, etc.). Les processus de l'Insee relèvent évidemment de la production statistique, mais certains concernent également la gestion de répertoires, la réalisation d'études ou des activités support. Au total, une centaine de processus ont été recensés.

Par ailleurs, l'institut, dans sa politique de gestion des risques, a considéré que les efforts de sécurisation devaient porter d'abord sur certains « processus essentiels ». Cela concerne une vingtaine de processus répondant, notamment, aux enjeux suivants :

- assurer les missions régaliennes de l'Insee que sont, par exemple, la gestion des répertoires d'identification des entreprises (Sirene) et des personnes physiques (RNIPP¹⁸) ;
- mettre à disposition des communes leurs populations de référence (estimées à partir du RP) ;
- produire les grands agrégats économiques, dont le produit intérieur brut (PIB) ;
- alimenter le diagnostic conjoncturel (indices mensuels, dont l'IPC, etc.).

¹⁷ <https://www.iso.org/obp/ui/#iso:std:iso:9000:ed-4:v2:fr>.

¹⁸ <https://www.insee.fr/fr/information/5019311>.

Pour définir ensuite les actions nécessaires pour sécuriser les processus, il a été jugé que le cadre des démarches qualité était à la fois le plus adapté et le plus fédérant. La troisième revue par les pairs européens organisée en France, en 2021, a conforté ce choix. Les auditeurs ont enjoint à l'Insee de déployer des démarches sur tous les processus essentiels d'ici 2027. Des recommandations du même type ont été prodiguées à d'autres INS, par exemple à l'Irlande (Forde et Moore, 2022).

Au cours des cinq dernières années, l'Insee a mené des démarches qualité sur la moitié de ses processus essentiels, soit une dizaine d'entre eux. Ce taux de couverture est à mettre en regard du large spectre d'activités qu'englobe chacun de ces processus, mais aussi des contraintes calendaires des équipes, lesquelles doivent être prises en compte pour la mise en œuvre des démarches. Ce taux ne reflète pas, par ailleurs, le fait que certains processus ont bénéficié au cours de ces années d'une actualisation d'une démarche initiale.

Plus globalement, sur la centaine de processus recensés, une vingtaine a fait l'objet d'une démarche qualité, qui a parfois été actualisée. En effet, en sus des processus essentiels, d'autres processus ont été examinés en raison d'enjeux particuliers (statistiques soumises à un règlement européen, par exemple) et/ou de besoins exprimés par les producteurs allant au-delà de la sécurisation. En particulier, ceux-ci recherchent également, dans la mise en œuvre de ces démarches, l'opportunité de mutualiser les savoirs de chacun dans le cadre collectif des équipes. L'enjeu est alors de construire une compréhension globale et partagée des processus.

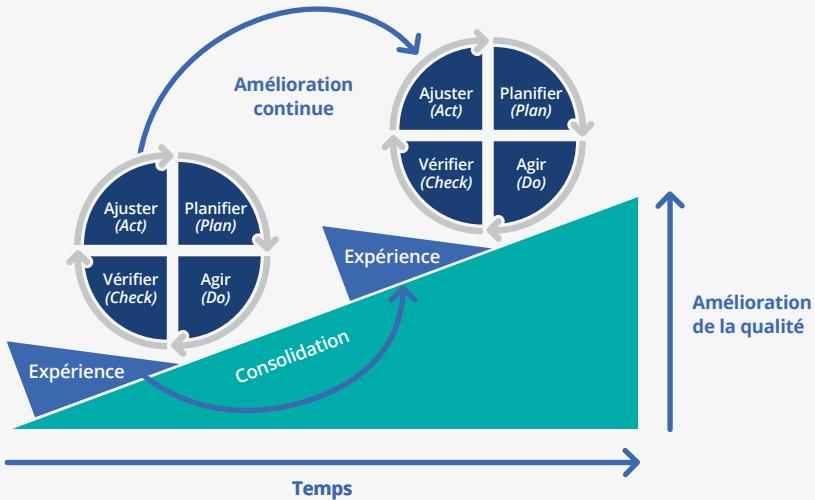
Enchaîner démarches et revues, pour une amélioration continue des processus

L'idée est d'améliorer progressivement le processus par des actions ciblées, priorisées, mises en œuvre et régulièrement adaptées par les équipes. C'est ce que l'on appelle l'amélioration continue.

La stratégie des démarches qualité est celle des petits pas. L'idée est d'améliorer progressivement le processus par des actions ciblées, priorisées, mises en œuvre et régulièrement adaptées par les équipes. C'est ce que l'on appelle l'amélioration continue, promue dans le CoP et formalisée par la **roue de Deming** (*figure 1*). La roue symbolique qu'a imaginée le statisticien américain est divisée en quatre étapes :

- 1) *Plan* (planifier) : planifier les actions d'amélioration, définies lors de la démarche qualité ;
- 2) *Do* (agir) : mettre en œuvre les actions décidées ;
- 3) *Check* (vérifier) : vérifier que les objectifs visés sont atteints ou comprendre les écarts, grâce à une actualisation de la démarche appelée « revue » ;
- 4) *Act* (ajuster) : adapter le plan d'actions initial, au regard des écarts constatés.

► **Figure 1 - La roue de Deming**



Lecture : La roue de Deming symbolise une méthode de travail pour améliorer en continu la qualité d'un processus. Elle consiste à dérouler successivement quatre étapes : « planifier » les actions d'amélioration, « agir » en les mettant en œuvre, « vérifier » leur efficacité et « ajuster » si besoin. Ce cycle doit ensuite être renouvelé régulièrement.

La révision régulière du plan d'actions au fil des cycles « *Plan Do Check Act* » permet progressivement de réduire les risques encourus par le processus. À titre d'exemple, le processus Réfigéo, qui centralise et actualise les nomenclatures de données géographiques, a déjà fait l'objet de trois démarches qualité : une démarche initiale et deux revues. Lors de la démarche initiale, une dizaine de risques avaient été identifiés, en particulier sur le respect des délais de diffusion (liés aux fournisseurs, à l'application informatique, à un circuit de traitement des données insuffisamment organisé, etc.). Lors de la deuxième revue réalisée, cinq ans après, les risques évoqués sur les délais de diffusion avaient disparu, témoignant du travail de sécurisation accompli. Il ne demeure aujourd'hui que trois risques à maîtriser davantage : le suivi des textes législatifs relatifs aux communes, l'identification des utilisateurs, mais aussi une nouvelle difficulté. En effet, en cinq ans, les canaux de diffusion se sont diversifiés et il est possible que leurs contenus ne soient pas strictement identiques.

Mais, concrètement, comment se déroule une démarche qualité ?

Une démarche qualité se déroule en trois temps :

- 1) **La description globale du processus** : celle-ci permet de mettre en avant les forces et faiblesses du processus et, partant, les axes d'amélioration.
- 2) **L'analyse des risques à surveiller** : elle vise à identifier les risques potentiels ou avérés susceptibles d'affecter le fonctionnement du processus et les résultats ; ils sont classés selon leur criticité¹⁹ afin de définir des actions prioritaires de sécurisation.
- 3) **Le plan d'actions à mettre en œuvre** : il comprend une liste d'actions visant à optimiser ou sécuriser le processus.

¹⁹ Indicateur prenant en compte à la fois la fréquence du risque et la gravité des conséquences de sa réalisation.

Une approche collective, associant experts qualité et producteurs

Comme les producteurs travaillent dans des délais tendus et n'ont qu'un temps limité à investir dans la démarche, les experts qualité s'adaptent à leurs contraintes et les accompagnent au maximum.

Les démarches qualité sont pilotées par des agents de l'unité Qualité, appelés « experts qualité ». Ils travaillent en interaction avec les producteurs, impliquant tout ou partie d'une équipe en charge d'un processus (*figure 2*). Il s'agit en effet d'une **démarche collective**, à la différence des démarches de type audit. Cependant, comme les producteurs travaillent dans des délais tendus et n'ont qu'un temps limité à investir dans la démarche, les experts qualité s'adaptent à leurs contraintes et les accompagnent au maximum. Ils mènent une

première analyse sur documents, préparent les premières versions des rapports, organisent et animent les réunions pour en discuter, améliorent les documents à partir de ces échanges et restituent les résultats de la démarche. Les réunions ont lieu avec les producteurs et éventuellement d'autres partenaires, notamment des informaticiens.

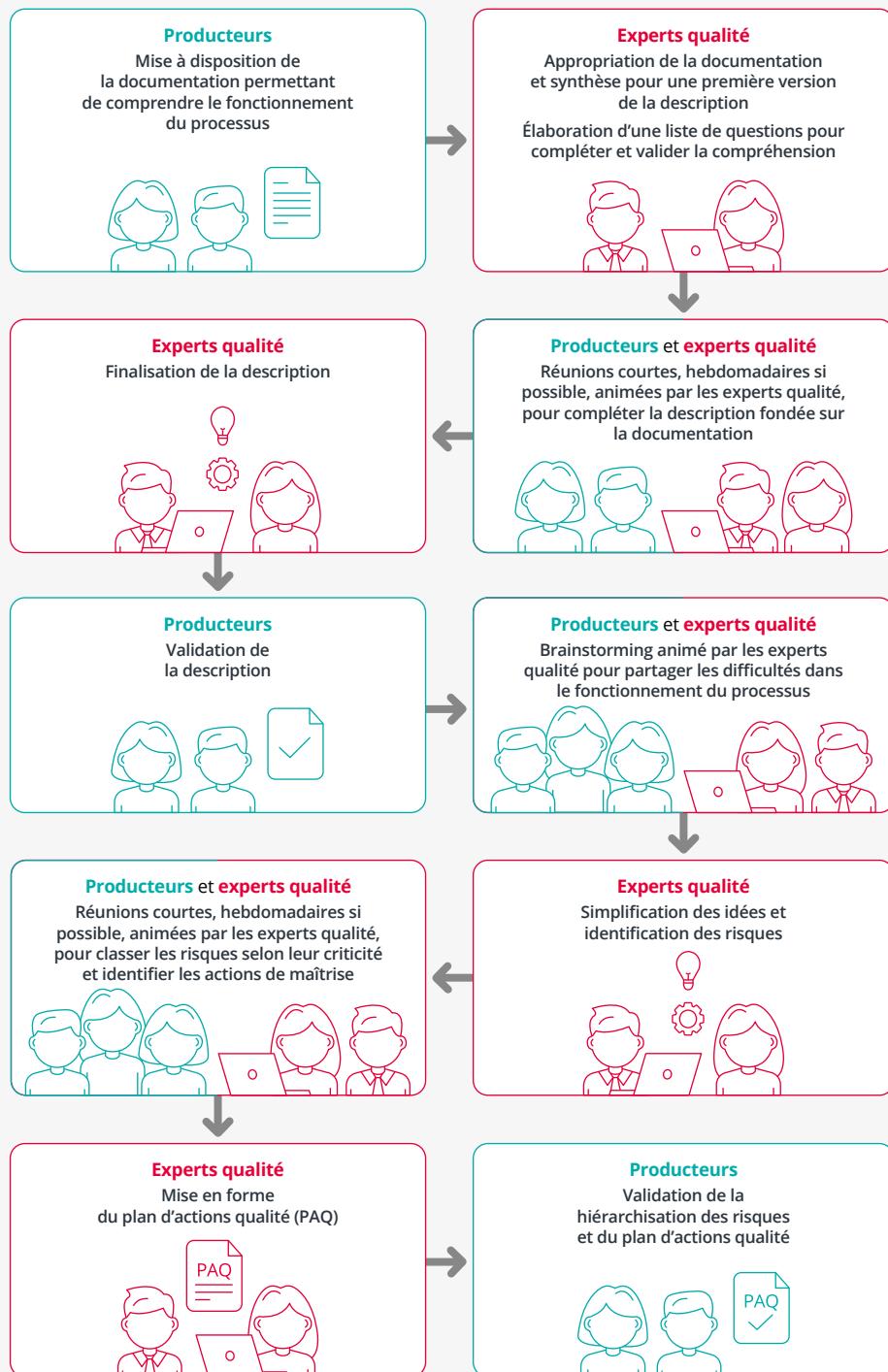
Au début de la démarche, les producteurs mettent à disposition des experts qualité toute la documentation existante, pour que les experts puissent s'acculturer au sujet et commencer à décrire le processus. Les producteurs apportent ensuite essentiellement leur connaissance et leur expérience du processus lors des réunions, sans qu'aucune préparation ne leur soit demandée. Ils doivent également valider les différents documents élaborés par les experts. Le plan d'actions sera ensuite mis en œuvre par les producteurs, avant de refaire ultérieurement le point à l'occasion d'une revue de la démarche.

En général, une démarche initiale s'étale sur un peu plus d'un an (*encadré 1*), un temps proche de celui des analyses réalisées par Eurostat sur ses propres processus statistiques (Bechey, 2024). Les équipes de production consacrent quant à elles environ une vingtaine d'heures à la démarche. Ce peut être un peu plus long quand le processus est complexe ou un peu moins quand il est déjà bien documenté.

Certains instituts, y compris Eurostat, demandent au préalable aux producteurs de réaliser eux-mêmes une première évaluation de leur processus à partir de questionnaires inspirés du CoP²⁰. L'intérêt est de disposer d'un bilan a minima, mais pour tous les principaux processus. Ceci permet d'établir des premiers plans d'actions. L'idée est ensuite de mener des démarches plus approfondies, en commençant par les processus les plus stratégiques. Les avantages d'un tel fonctionnement sont à mettre en balance avec la charge que représente la double évaluation pour les producteurs (Lound et Mais, 2022), ainsi que la difficulté pour ces derniers de mener seuls une analyse de leur processus, même avec l'aide de tels outils d'autoévaluation (Bechey, 2024).

20 Par exemple, le questionnaire DESAP est une liste générique de contrôles pour évaluer la qualité des enquêtes du système statistique européen.
https://ec.europa.eu/eurostat/documents/64157/4373903/07-Checklist-for-Survey-Managers_DESAP-EN.pdf/ec76e3a3-46b5-409e-a7c3-52305d05bd42.

► Figure 2 - Déroulement d'une démarche qualité



► Encadré 1. Une démarche s'étale sur environ un an, mais ne mobilise les producteurs qu'une vingtaine d'heures

Quel que soit le processus étudié, il s'écoule généralement un peu plus d'un an entre le lancement d'une démarche qualité et la fin de cette dernière (**figure encadré**). Cette durée peut occasionnellement être allongée quand les conditions au sein des équipes de production sont particulières. Ce fut le cas pour les enquêtes thématiques auprès des entreprises, du fait du fort renouvellement de l'équipe et de périodes de vacance sur certains postes.

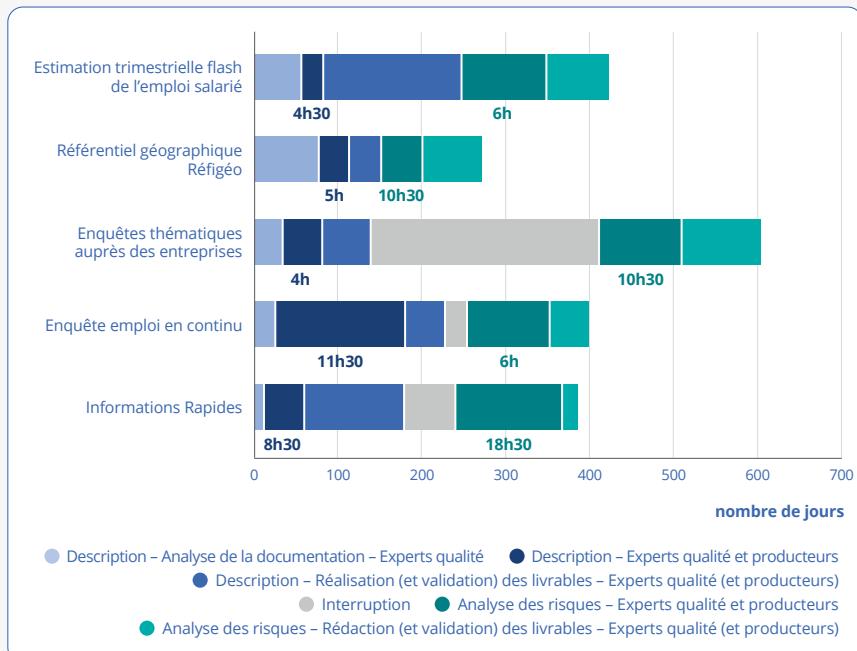
Une démarche qualité mobilise les équipes de production pendant une vingtaine d'heures en moyenne. Il s'agit de la durée totale des réunions avec les experts qualité, mais tous les acteurs ne sont pas toujours sollicités. La configuration dépend des étapes de la démarche : de quelques acteurs principaux pour la description à un cercle plus large pour l'analyse des risques, incluant représentants informatiques et autres partenaires internes. L'investissement des acteurs dépend de la complexité du processus (plusieurs sources, plusieurs modes de collecte, richesse des produits de diffusion). Un nombre élevé d'acteurs génère aussi plus d'échanges et le consensus peut être un

peu plus long à trouver. À l'inverse, l'investissement des équipes peut être moindre si la documentation fournie en amont est riche. Par exemple, le processus des enquêtes thématiques auprès des entreprises est complexe (une dizaine d'enquêtes aux contraintes différentes), mais bien documenté. Pour le décrire, les experts qualité ont passé un peu plus d'un mois à analyser la documentation, mais seulement quatre heures à échanger avec le responsable du processus. L'analyse des risques a en revanche mobilisé une dizaine d'agents de l'équipe pendant un peu plus de dix heures.

Après l'analyse des risques, il peut parfois s'écouler du temps avant la fin de la démarche. Il n'est en effet pas possible de planifier le comité de pilotage de clôture tant que la démarche n'est pas assez avancée et il est souvent compliqué de trouver rapidement une date qui convienne à l'ensemble des acteurs.

Lorsqu'il s'agit d'actualiser une démarche qualité initiale, la durée des travaux se raccourcit significativement. Ainsi, la deuxième revue du processus Réfigéo, intervenue trois ans après la précédente et cinq ans après la démarche initiale, a été réalisée en neuf mois.

Chronologie de cinq démarches qualité de processus



Lecture : La phase de description du processus est en bleu et celle de l'analyse des risques, y compris la définition du plan d'actions d'amélioration, est en vert. Pour chacune de ces phases, les zones plus foncées correspondent aux périodes où les producteurs sont plus particulièrement sollicités, les nombres d'heures de réunion étant précisés au-dessous. Entre les deux phases de la démarche qualité, il existe parfois des interruptions plus ou moins longues pour des raisons d'organisation : elles sont représentées en gris.

► Les apports des démarches qualité

Questionner les besoins des utilisateurs et les choix de diffusion

Les démarches qualité débutent par l'examen du contexte dans lequel les statistiques publiques sont élaborées. Elles identifient d'abord les réglementations auxquelles répondent les statistiques, ainsi que les besoins des utilisateurs. Il s'agit à cette étape de faire la part entre les contraintes que les démarches peuvent questionner et celles qui sont intangibles.

Deux tiers des statistiques de l'Insee sont encadrées par des règlements européens. Ces derniers imposent des contraintes fortes sur les concepts, méthodes et calendriers de diffusion sur lesquelles les producteurs ont peu de marge. L'Insee participe aux échanges avec Eurostat, mais doit appliquer, comme les autres INS, les décisions qui en résultent.

Les choix de diffusion visant à répondre aux besoins des utilisateurs nationaux peuvent davantage être questionnés. Pour certains processus, les besoins nationaux sont bien identifiés, grâce aux commissions du [Conseil national de l'information statistique²¹](#) (Cnis) et à des *focus groups* avec des utilisateurs. C'est le cas par exemple pour le recensement de la population. Pour d'autres processus, les utilisateurs autres qu'institutionnels sont difficiles à connaître et leurs besoins encore plus. Aussi, un grand nombre de produits de diffusion sont mis à la disposition du grand public sur insee.fr. Les utilisateurs du site peuvent en effet avoir des profils très variés et leurs besoins différer fortement, sans pour autant qu'ils soient connus.

Les experts qualité sont parfois amenés à analyser la « popularité » des produits pour savoir s'ils répondent effectivement à un besoin. De fait, certains produits sont très peu consultés. Par exemple, le processus d'élaboration de l'indice des prix à la consommation générale, en plus des indices mesurant l'inflation, les prix moyens d'une soixantaine de produits à un niveau très fin. En 2023, les prix moyens mensuels de la baguette ont été consultés 7 000 fois alors que ceux du citron l'ont été seulement 60 fois.

“ Les démarches qualité incitent alors les producteurs à réajuster leur offre de diffusion et à dialoguer davantage avec les utilisateurs. ”

Les démarches qualité incitent alors les producteurs à réajuster leur offre de diffusion et à dialoguer davantage avec les utilisateurs : pour identifier les besoins devenus obsolètes et, à l'inverse, les besoins émergents. En Grande-Bretagne, une démarche qualité sur les indices des prix à la production avait conduit de même à la recommandation de rationaliser l'offre de diffusion (Dunn, 2024).

²¹ <https://www.cnis.fr/>.

Être attentif au cadre de réception des données externes

L'analyse du contexte se poursuit par le recensement des données externes nécessaires au processus. Il s'agit le plus souvent de données administratives, fiscales ou sociales. Pour sécuriser leur transmission, il convient de vérifier l'existence et le contenu des conventions qui cadrent les échanges avec les fournisseurs. Les démarches qualité incitent alors à mettre à jour certaines conventions qui peuvent être incomplètes, voire caduques.

Les conventions ne sont toutefois pas suffisantes. La non-réception de données externes ou un contenu non conforme à l'attendu (à la suite, par exemple, d'un changement de concept) restent des sources d'inquiétude majeures pour les producteurs. La première pourrait compromettre la diffusion des résultats et la deuxième dégraderait la qualité des statistiques diffusées. Les démarches qualité donnent l'opportunité aux producteurs de réfléchir à des actions pour limiter la survenue de ces difficultés, ou à défaut pour en limiter l'impact.

Par exemple, l'élaboration des estimations trimestrielles flash²² de l'emploi salarié privé s'appuie principalement sur les déclarations sociales nominatives (DSN)²³. Pour sécuriser le processus, il est prévu de sensibiliser systématiquement les fournisseurs externes aux délais de diffusion auxquels l'Insee est contraint par Eurostat. D'autres actions visent à davantage anticiper et maîtriser des changements de contexte qui pourraient avoir un impact sur l'élaboration des estimations, à travers notamment des veilles sur les évolutions de l'emploi (réglementation, emplois aidés, etc.) et leur bonne prise en compte dans la DSN.

Les démarches qualité ont ainsi pour ambition d'améliorer et de sécuriser les processus au regard des contraintes externes, mais également de maîtriser les éventuels risques internes. Pour cela, il est indispensable de connaître le fonctionnement global des processus et d'en faire une description accessible à l'ensemble des acteurs impliqués.

Offrir une description structurée pour une meilleure vision globale

Les descriptions réalisées dans le cadre des démarches qualité répondent aux besoins exprimés par tous les producteurs : documenter le fonctionnement des processus pour faciliter la transmission des connaissances au sein des équipes. Les démarches qualité proposent de cartographier les processus selon une grille d'analyse spécifique : le **modèle générique du processus de production statistique**, ou GSBPM dans son acronyme anglais, pour *Generic Statistical Business Process Model* (Unece, 2019) (**encadré 2**). Cette grille d'analyse permet de construire une vue d'ensemble structurée et accessible à tous, comme l'illustre l'exemple relatif au processus d'élaboration de l'indice des prix à la consommation (IPC) (**figure 3**). Elle offre une vision simplifiée, mais complète, de processus parfois complexes, du fait par exemple de la diversité des sources d'informations et des modes de collecte.

²² L'Insee réalise deux estimations trimestrielles d'emploi salarié : une estimation « flash », publiée 40 jours environ après la fin du trimestre considéré, qui couvre uniquement l'emploi salarié privé, puis une estimation « détaillée », publiée 70 jours environ après la fin du trimestre et qui porte sur l'ensemble de l'emploi salarié.

²³ La déclaration sociale nominative (DSN) permet à l'employeur de déclarer et payer ses cotisations sociales via les données de l'entreprise sur ses salariés. Voir à ce sujet Humbert-Bottin (2018).

Le GSBPM est un outil indispensable à la description d'un processus statistique, mais il présente certaines limites. En particulier, la chronologie des activités ne peut pas être restituée dans le détail de leurs itérations. Comme c'est un point important pour les services producteurs, les démarches offrent la possibilité de compléter cette cartographie par des **diagrammes d'activités** : ces derniers précisent comment les activités s'enchaînent, ou au contraire sont menées en parallèle sur des parties de processus (*figure 4*). Cette représentation a ainsi été utilisée pour compléter certaines descriptions : par exemple pour les activités liées au changement d'année de l'IPC ou

► Encadré 2. Le GSBPM : un standard utilisé pour la description des processus

En matière d'évaluation de la qualité, certains outils internationaux ont fait la preuve de leur puissance d'analyse. L'exemple le plus représentatif d'entre eux est le modèle générique du processus de production statistique (GSBPM dans son acronyme anglais, pour *Generic Statistical Business Process Model*). Développé en 2008 par l'Unece*, Eurostat et l'OCDE**, et inscrit parmi la liste des standards européens depuis 2017, il permet aux organismes statistiques d'adopter une terminologie commune et de s'entendre sur la description détaillée du cycle de vie d'une opération statistique.

Le GSBPM décompose toute opération de production statistique en huit phases : définition des besoins, conception, élaboration, collecte, traitement, analyse, diffusion, évaluation (*figure encadré*). Enfin, trois processus transverses

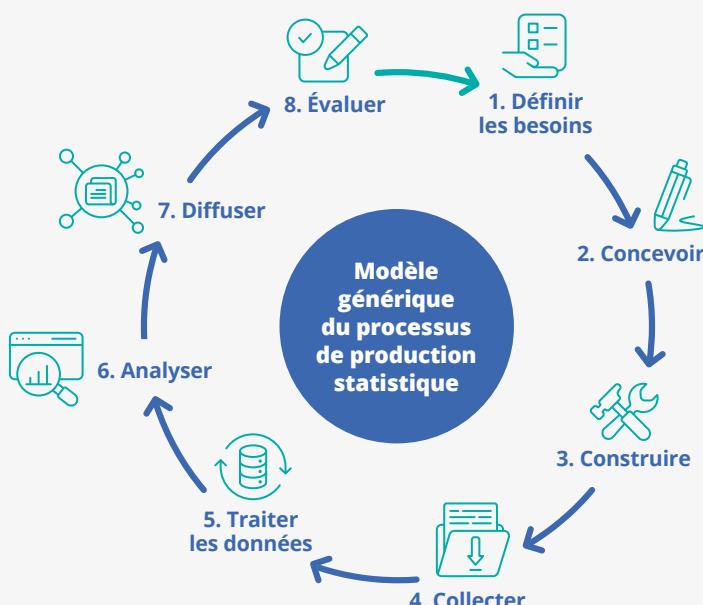
aux huit phases complètent ce modèle : gestion de la qualité, gestion des métadonnées et gestion des données.

Le GSBPM est une grille souple qui permet un usage au plus près des besoins. Par exemple, il permet de décrire une enquête du début à la fin en déroulant les huit phases ou seulement la partie « production » d'un indicateur mensuel (phases allant de la collecte à la diffusion).

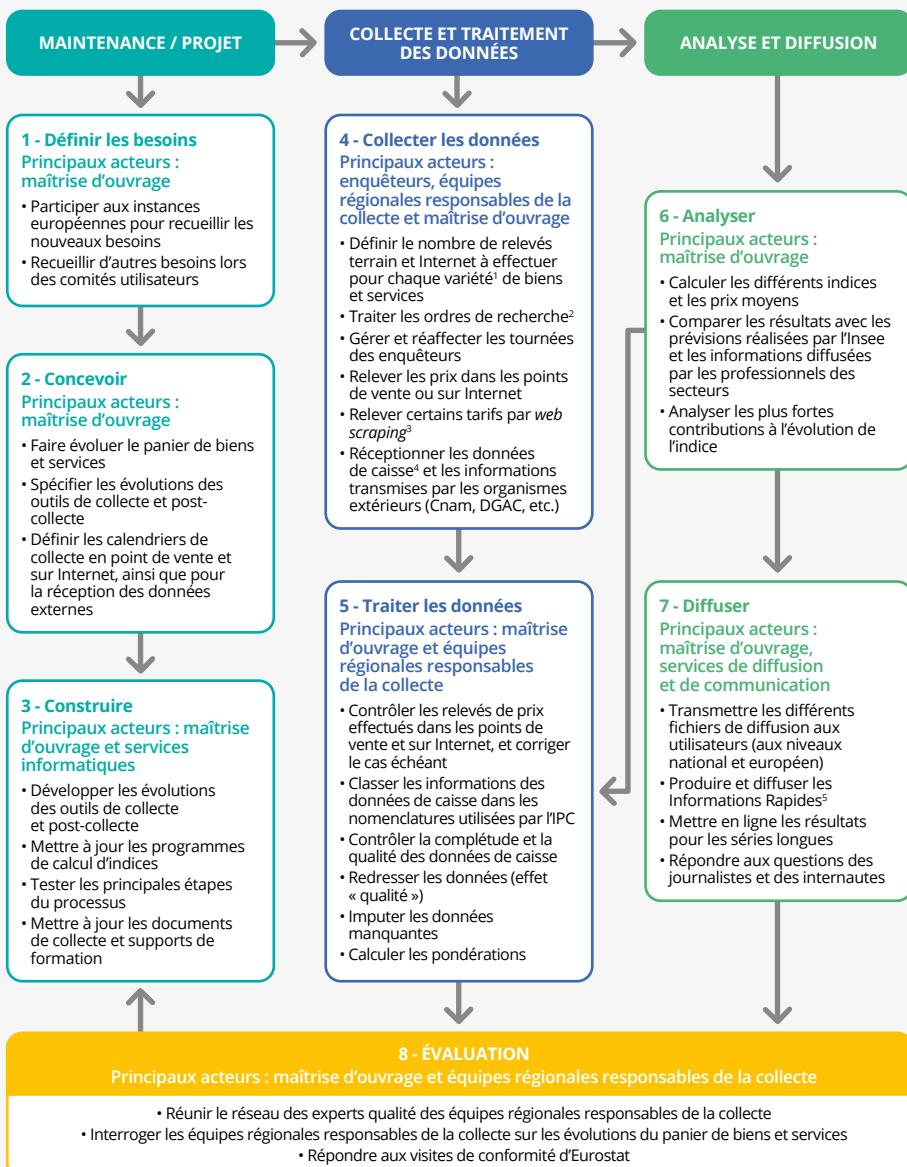
Utilisé par la plupart des pays pour décrire et analyser leurs processus statistiques, il fait l'objet de partages d'idées et d'expériences. Les échanges permettent de constater que certains pays vont plus loin dans son utilisation, en le rattachant à leur système de documentation ou de pilotage (Erikson, 2020).

* Unece : Commission économique des Nations unies pour l'Europe (et l'Amérique du nord).

** OCDE : Organisation de coopération et de développement économique.



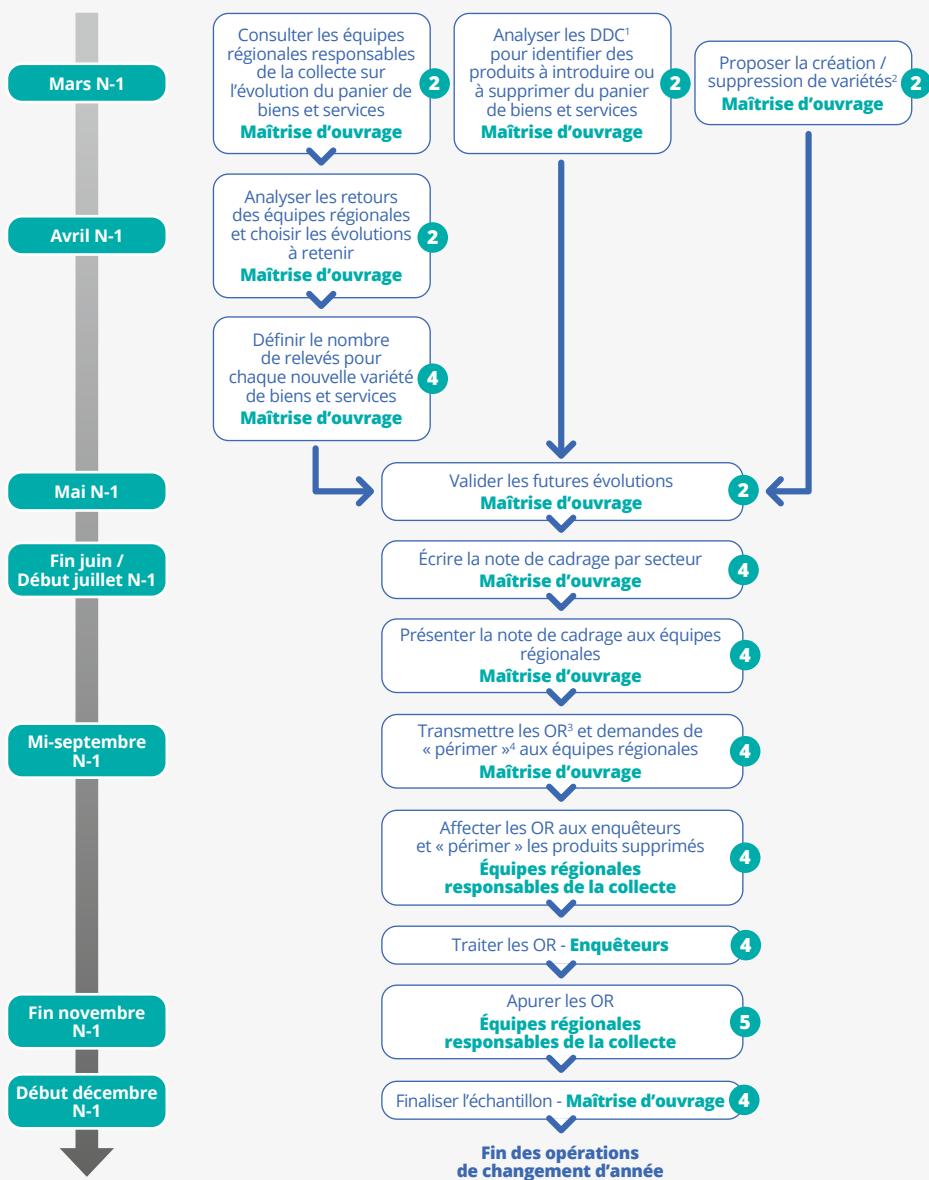
► **Figure 3 - Cartographie du processus d'élaboration de l'indice des prix à la consommation**



- Les biens et services suivis par l'IPC découlent d'une nomenclature européenne dont le niveau le plus fin est appelé sous-classe. Chaque sous-classe est partitionnée en postes pour l'IPC et les variétés sont des biens et services représentant ces postes. Par exemple, la variété « croissant au beurre » est un représentant du poste « viennoiserie ».
- Lorsque de nouveaux biens ou services sont introduits dans le panier suivi par l'IPC, il convient de repérer les articles associés à suivre dans les points de vente ou sur les sites Internet de vente. La demande de recherche de ces articles est appelée ordre de recherche.
- Technique d'extraction automatisée de données de sites web.
- Données recueillies par les enseignes de la grande distribution au moment du passage des clients à la caisse.
- Les Informations Rapides sont des publications qui présentent les derniers indices et les résultats les plus récents des enquêtes de conjoncture de l'Insee.

Sigles : Cnam : Caisse nationale de l'assurance maladie, DGAC : Direction générale de l'aviation civile, IPC : indice des prix à la consommation.

► **Figure 4 - Indice des prix à la consommation : principales opérations de changement d'année pour la collecte dans les points de vente**



1. Données de caisse : données recueillies par les enseignes de la grande distribution au moment du passage des clients à la caisse.

2. Les biens et services suivis par l'indice des prix à la consommation (IPC) découlent d'une nomenclature européenne dont le niveau le plus fin est appelé sous-classe. Chaque sous-classe est partitionnée en postes pour l'IPC et les variétés sont des biens et services représentant ces postes. Par exemple, la variété « croissant au beurre » est un représentant du poste « viennoiserie ».

3. Pour chaque nouveau bien à collecter dans les points de vente, un ordre de recherche (OR) est généré. Il s'agit alors de repérer dans les points de vente les articles dont les prix devront être collectés.

4. Pour chaque suppression de bien collecté auparavant dans des points de vente, une demande de « périmer » est générée. Il s'agit alors de supprimer les articles correspondants dans l'application de collecte.

Lecture : Pour chaque activité décrite, les acteurs qui en sont responsables sont mentionnés en vert, de même que le numéro de la phase du GSBPM concernée (voir encadré 2). Le diagramme permet ainsi à la fois de situer dans le temps les différentes activités et de visualiser les allers-retours entre les phases 2, 4 et 5 du GSBPM.

encore pour celles liées à l'élaboration des publications d'indicateurs conjoncturels d'activités économiques (collection *Informations Rapides* de l'Insee²⁴). Dans ce dernier cas, le diagramme a mis en évidence des points de tensions calendaires : pour certains indicateurs, les publications peuvent nécessiter jusqu'à trois niveaux de relecture en l'espace de trois heures, avec l'intégration d'éventuelles modifications demandées à chaque niveau de relecture.

Ces descriptions structurées sont importantes dans un contexte où les mobilités dans les services producteurs sont fréquentes.

continuité de service en servant de documentation de base, autour de laquelle peut s'articuler une documentation plus détaillée.

Ces descriptions offrent non seulement une vue globale à partager entre tous les acteurs impliqués, mais elles permettent aussi aux experts qualité d'acquérir une connaissance suffisante pour analyser objectivement les processus. Ils sont alors en mesure de mettre en avant leurs forces et faiblesses.

Mettre en lumière les forces et faiblesses du processus

L'enjeu sur les points forts peut paraître moins stratégique. Pourtant, il est important de consigner ces derniers pour qu'ils perdurent dans le temps.

Du fait de leur positionnement extérieur aux équipes de production, les experts qualité ont un regard neutre sur le fonctionnement du processus. Ils peuvent ainsi aider les équipes à en identifier les points forts et ceux à améliorer.

En première approche, l'enjeu sur les points forts peut paraître moins stratégique.

Pourtant, il est important de consigner ces

derniers pour qu'ils perdurent dans le temps, notamment au-delà des renouvellements d'équipe. Les démarches qualité en sont l'occasion. Pour certains processus, par exemple, certaines évolutions ont permis de réduire la charge des enquêtés et/ou d'enrichir la caractérisation du phénomène étudié. C'est le cas, pour plusieurs processus, du recours accru aux données administratives, mais aussi de la mise en place d'organisations permettant que des processus s'alimentent entre eux. Ainsi, les questionnaires des enquêtes thématiques auprès des entreprises sont préremplis avec les chiffres d'affaires issus d'un autre processus : celui de l'élaboration des *statistiques structurelles d'entreprises*²⁵.

²⁴ <https://www.insee.fr/fr/information/3701213>.

²⁵ <https://www.insee.fr/fr/information/3291202>.

Les démarches qualité permettent également d'identifier les **points à améliorer**. Elles incitent par exemple à automatiser certaines tâches manuelles ou à repenser une partie de l'organisation. Pour reprendre le cas des enquêtes thématiques auprès des entreprises, la démarche qualité a mis en évidence le nombre conséquent d'acteurs impliqués dans les travaux d'élaboration de la base de sondage et d'échantillonnage, et la multiplicité des allers-retours entre eux. Or, le risque de dérapage dans les délais croît avec le nombre d'acteurs. Même s'il n'y a pas eu jusque-là d'incident, la démarche qualité a mis en évidence l'importance de répartir les rôles de façon plus optimale, en recentrant davantage les travaux des équipes en charge de la collecte et de l'échantillonnage sur leurs cœurs de métier respectifs.

Les experts qualité s'appuient sur ces points forts à entretenir et ces points à améliorer pour proposer des actions qui pourraient être mises en œuvre. Ils ne sont cependant que force de proposition : **l'élaboration du plan d'actions revient aux équipes en charge du processus**, dans un souci d'adhésion. Ces dernières sont alors invitées à échanger ensemble pour déterminer collégialement les risques qu'il convient en priorité de maîtriser davantage.

Croiser les regards pour aboutir à un plan d'actions consensuel

Croiser les regards des différents acteurs permet à ces derniers de partager leurs difficultés.

L'aspect collectif sur lequel insistent les experts qualité lors des démarches est primordial. Croiser les regards des différents acteurs permet à ces derniers de partager leurs difficultés et subséquemment d'identifier les risques vraiment problématiques sur lesquels agir.

Les analyses de risques menées dans le cadre des démarches qualité sont basées sur une méthodologie de type « **analyse des modes de défaillance, de leurs effets et de leur criticité** » (**encadré 3**). Elles visent à élaborer collectivement un plan d'actions pour sécuriser davantage le processus. Le rôle différent de chaque acteur dans la mise en œuvre de ce dernier conduit inévitablement à une perception très personnelle de la nature des risques et de leur gravité. Par exemple, les acteurs en charge de la collecte d'une enquête seront plus sensibles aux risques pouvant affecter le taux de collecte que ne le seront les responsables des traitements en aval. Ces derniers peuvent en effet y pallier, dans une certaine mesure, par des méthodes d'imputation ou de repondération. Les échanges entre les différents acteurs permettent ainsi de relativiser certains problèmes. L'enjeu principal de cette phase de la démarche qualité est donc d'aboutir à **un consensus sur les effets des risques auxquels sont exposés les processus et sur la criticité de ces risques**. Ce consensus est indispensable pour que les acteurs adhèrent au plan d'actions et le mettent en œuvre in fine.

Les analyses de risques incitent donc à prioriser des actions répondant aux risques les plus critiques. D'autres actions peuvent être intégrées pour résoudre certains points qui ne relèvent pas de la sécurité du processus, mais sont mal vécus au quotidien pour certains acteurs. Il peut s'agir par exemple de mieux formaliser un circuit de relecture interne avant publication pour faciliter la gestion des remarques issues des diverses lectures (Briant et al., 2024).

► Encadré 3. Les analyses de risques de processus à l'Insee

Deuxième étape des démarches qualité, l'analyse des risques d'un processus s'appuie sur une méthodologie inspirée de l'analyse des modes de défaillance, de leurs effets et de leur criticité (AMDEC). La méthode AMDEC a été élaborée dans les années 1940 par l'armée américaine et a été reprise ensuite par l'industrie aéronautique, puis par l'industrie automobile. La méthode mise en œuvre par l'Insee en reprend les grands principes et s'inscrit en cohérence avec les lignes directrices de l'Unece* pour la gestion des risques. Le déroulé est le suivant (*figure encadré*) :

1) Lors d'un brainstorming, l'ensemble des acteurs du processus recensent les événements perturbateurs qui ont empêché ou pourraient empêcher son bon déroulement.

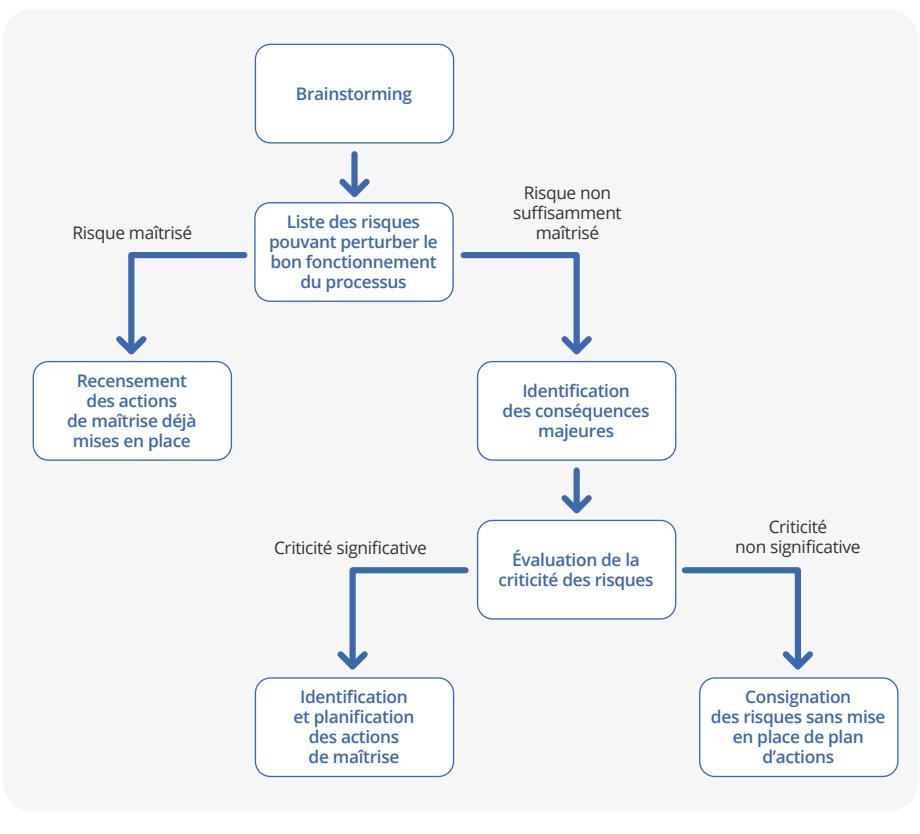
2) À partir du matériau ainsi collecté, les experts qualité établissent une liste des risques, puis distinguent avec les acteurs ceux qui sont insuffisamment maîtrisés de ceux qui le sont suffisamment.

3) Pour les risques jugés suffisamment maîtrisés, les actions actuelles de maîtrise sont consignées pour ne pas les perdre.

4) Pour les risques jugés insuffisamment maîtrisés, les acteurs identifient leur conséquence majeure et notent collectivement leur criticité : ils positionnent sur une échelle allant de 1 à 3, d'une part, la probabilité de leur réalisation, d'autre part, la gravité de la conséquence principale associée, puis font le produit de ces deux notes.

5) Enfin, les acteurs établissent en commun un plan d'actions pour maîtriser davantage les risques jugés les plus inacceptables.

* <https://unece.org/statistics/documents/2023/01/presentations/2c-guidelines-risk-management-practices-statistical>.



► Au-delà des apports « directs » des démarches qualité —

Les démarches qualité constituent ainsi un levier pour améliorer en continu la qualité des statistiques publiques via celle de leur processus d'élaboration. Leur mise en œuvre repose sur un corpus de normes et de méthodes éprouvées, au premier rang desquelles le code de bonnes pratiques de la statistique européenne et le GSBPM. Ce dernier fournit un langage et un cadre réutilisables d'un processus à l'autre, favorisant leur rapprochement et la transmission des connaissances.

“ *Les démarches offrent un temps de réflexion collectif aux équipes, hors contingences de production.* ”

Les démarches sont aussi importantes en elles-mêmes que les livrables auxquels elles permettent d'aboutir. Elles offrent un temps de réflexion collectif aux équipes, hors contingences de production, sur le fonctionnement de leur processus. Elles mettent en regard les perceptions des différents acteurs et les équipes disposent à la fin de feuilles de route pragmatiques à mettre en œuvre, dont le contenu et le calendrier ont été déterminés de manière collégiale. Il est préconisé ensuite de revoir régulièrement les

processus pour les ajuster, selon le cycle « *Plan Do Check Act* » de Deming. L'exemple des démarches sur le processus Réfigéo montre que les difficultés s'amoindrissent grâce à l'application de ce cycle. Encore faut-il que l'effort pour mettre en œuvre les actions nécessaires se maintienne dans le temps, ce qui est considéré comme le défi le plus délicat des démarches qualité (Forde et Moore, 2022).

L'accompagnement par un collègue expert en qualité facilite le déroulé des démarches et garantit autant que possible son objectivité et son efficacité. Celui-ci impulse, organise et anime ces temps d'échange, selon un protocole et des méthodes souvent peu familiers aux équipes. Ces dernières expriment en général leur satisfaction d'avoir été accompagnées dans ce travail de fond et d'avoir été déchargées de l'organisation des échanges et de la rédaction des conclusions.

Les démarches qualité contribuent à optimiser et sécuriser les processus, mais elles ne peuvent pas tout :

- Les équipes ont certes la main sur des actions visant à fluidifier leur communication interne, à compléter la documentation des procédures, à approfondir la connaissance de leurs utilisateurs et de leurs besoins. Mais certaines actions identifiées relèvent de décisions sur lesquelles elles n'ont pas prise. C'est le cas de la disponibilité des moyens (ressources humaines et informatiques), qu'elles évoquent régulièrement. Les démarches peuvent néanmoins en être l'écho, parmi d'autres canaux de remontée.
- Par ailleurs, les démarches ambitionnent de définir des actions graduelles d'amélioration. Cependant, des évolutions de plus grande ampleur sont parfois nécessaires, comme l'illustrent les projets de refonte de processus régulièrement lancés. D'autres modes de travail sont alors mobilisés, de type gestion de projet. Quand une démarche qualité a été précédemment menée, elle peut alors fournir un précieux état des lieux pour asseoir les réflexions.

Les problématiques mises en évidence par les démarches qualité mettent en avant des préoccupations communes à l'ensemble des processus. Il convient donc de capitaliser sur ces enseignements pour en déduire des principes, lignes directrices, qui facilitent le travail des équipes. C'est ce que s'efforce de faire le SSP. Ce dernier a ainsi élaboré des **règles communes pour la diffusion des statistiques²⁶**, incluant notamment un protocole standard pour gérer les erreurs dans les publications. De même, l'INS du Danemark a mis au point un standard pour ce que les statisticiens appellent le *data editing*, c'est-à-dire les contrôles de données individuelles et agrégées aux différentes phases d'un processus (Andersen et Blix, 2022).

D'autres pays européens réalisent des démarches qualité, parfois appelées autrement (audits, examens, etc.). Les pratiques sont similaires à celles de l'unité Qualité, avec le recours aux mêmes standards (CoP et GSBPM, en particulier) et des équipes dédiées à ces travaux. Ces dernières sont cependant parfois pluridisciplinaires (incluant des méthodologues). Par ailleurs, certaines démarches sont précédées d'autoévaluations des processus par les équipes. Autant de variantes qu'il conviendrait d'instruire, inscrivant aussi les démarches qualité françaises dans l'amélioration continue. Notamment, l'autoévaluation, certes moins approfondie que les démarches qualité décrites ici, pourrait permettre d'évaluer a minima la qualité d'un spectre plus large de processus à l'Insee, mais aussi dans les services statistiques ministériels. En effet, comme évoqué précédemment, les démarches représentent un investissement pour les équipes par ailleurs fortement occupées, malgré les bénéfices reconnus a posteriori.

²⁶ <https://www.insee.fr/fr/information/1300614>.

► Fondements juridiques

- Règlement (CE) n° 223/2009 révisé du Parlement européen et du Conseil du 11 mars 2009 relatif aux statistiques européennes. In : *site de l'Union européenne*. [en ligne]. [Consulté le 4 juin 2025]. Disponible à l'adresse :
<https://eur-lex.europa.eu/legal-content/FR/ALL/?uri=CELEX:32009R0223>.

► Bibliographie

- ANDERSEN, Anne Kaag et BLIX, Karin, 2022. Managing quality assurance. In : *European Conference on Quality in Official Statistics (Q2022)*. 8 juin – 10 juin 2022. [en ligne]. [Consulté le 4 juin 2025]. Disponible à l'adresse suivante :
<https://q2022.stat.gov.lt/scientific-information/papers-presentations/session-11>.
- BECHEV, Ilcho, 2024. Quality Reviews in Eurostat, 2024. In : *European Conference on Quality in Official Statistics (Q2024)*. 4 juin – 7 juin 2024. [en ligne]. [Consulté le 4 juin 2025]. Disponible à l'adresse suivante :
<https://airdrive.eventsair.com/eventsairwesteuprod/production-leading-public/4f460e098c064057b6eb41a99b8e4fd6>.
- BERGLUND, Frode et HAUGEN, Per Ola, 2024. Put the users in focus! How focus-groups with users in quality reviews contributes to improved statistics. In : *European Conference on Quality in Official Statistics (Q2024)*. 4 juin – 7 juin 2024. [en ligne]. [Consulté le 4 juin 2025]. Disponible à l'adresse suivante :
<https://airdrive.eventsair.com/eventsairwesteuprod/production-leading-public/37392a4cd2664426a716b2b05aa33907>.
- BIEMER, Paul, TREWIN, Dennis, BERGDAHL, Heather et JAPEC, Lilli, 2014. A System for Managing the Quality of Official Statistics. In : *Journal of Official Statistics*. [en ligne]. Septembre 2014. Vol. 30, N° 3, pp. 381-415. [Consulté le 4 juin 2025]. Disponible à l'adresse : <https://doi.org/10.2478/jos-2014-0022>.
- BRIANT, Pierrette, JLASSI, Mahmoud et KRANKLADER, Élodie, 2024. The French approach to quality through processes: A key role in risk management at INSEE. In : *European Conference on Quality in Official Statistics (Q2024)*. 4 juin – 7 juin 2024. [en ligne]. [Consulté le 4 juin 2025]. Disponible à l'adresse suivante :
<https://airdrive.eventsair.com/eventsairwesteuprod/production-leading-public/6b4bfc688b3f4929becccd0307bdb7d04>.
- CHALEIX, Mylène et DURR, Jean-Michel, 2024. Les revues par les pairs : un levier au service de la statistique publique. In : *Courrier des statistiques*. [en ligne]. 16 décembre 2024. Insee. N° N12, pp. 137-157. [Consulté le 4 juin 2025]. Disponible à l'adresse :
<https://www.insee.fr/fr/information/8264560?sommaire=8264562>.
- CHALEIX, Mylène, LEFEBVRE, Olivier, 2024. Éthique et pratique, pratique de l'éthique, l'exemple de l'Insee. In : *site de l'Unece*. [en ligne]. Mars 2024. Unece. [Consulté le 4 juin 2025]. Disponible à l'adresse :
https://unece.org/sites/default/files/2024-03/Ethics2024_France_Chaleix_Lefebvre_D-FR.pdf.
- CHRISTINE, Marc et ROTH, Nicole, 2020. Le Comité du label – Un acteur de la gouvernance au service de la qualité des statistiques publiques. In : *Courrier des statistiques*. [en ligne]. 31 décembre 2020. Insee. N° N5, pp. 39-52. [Consulté le 4 juin 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/information/5008698?sommaire=5008710>.

- DESROSIÈRES, Alain, 2003. Les qualités des quantités. In : *Courrier des statistiques*. [en ligne]. Juin 2003. Insee. N° 105-106, pp. 51-63. [Consulté le 4 juin 2025]. Disponible à l'adresse : <https://www.bnsp.insee.fr/ark:/12148/bc6p06xt32g.r=%28prOx%3A%20%22desrosi%C3%A8re%22%201%20%22alain%22%29?rk=1287560;0>.
- DUNN, Marianthi, 2024. GDP revisions, unemployment and factory gate prices: Regulating the quality of UK economic statistics. In : *European Conference on Quality in Official Statistics (Q2024)*. 4 juin – 7 juin 2024. [en ligne]. [Consulté le 4 juin 2025]. Disponible à l'adresse suivante : <https://airdrive.eventsair.com/eventsairwesteuprod/production-leading-public/a39b951a53d8464f837197218f423c61>.
- ERIKSON, Johan, 2020. Le modèle de processus statistique en Suède – Mise en œuvre, expériences et enseignements. In : *Courrier des statistiques*. [en ligne]. 29 juin 2020. Insee. N° N4, pp. 122-141. [Consulté le 4 juin 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/information/4497085?sommaire=449709>.
- FORDE, Don et MOORE, Ken, 2022. From audit to assurance – An approach to the delivery of statistical quality assurance and improvement in the CSO. In : *European Conference on Quality in Official Statistics (Q2022)*. 8 juin – 10 juin 2022. [en ligne]. [Consulté le 4 juin 2025]. Disponible à l'adresse : <https://q2022.stat.gov.lt/scientific-information/papers-presentations/session-11>.
- HILLAIREAU, Fabrice et ZANA ROUQUETTE, Céline, 2024. Qui a confiance dans les chiffres de l'Insee ? In : *Le blog de l'Insee*. [en ligne]. 22 octobre 2024. [Consulté le 4 juin 2025]. Disponible à l'adresse : <https://blog.insee.fr/qui-a-confiance-dans-les-chiffres-de-l-insee/>.
- HUMBERT-BOTTIN, Élisabeth, 2018. La déclaration sociale nominative – Nouvelle référence pour les échanges de données sociales des entreprises vers les administrations. In : *Courrier des statistiques*. [en ligne]. 6 décembre 2018. Insee. N° N1, pp. 25-34. [Consulté le 4 juin 2025]. Disponible à l'adresse : <https://www.insee.fr/fr/information/3647025?sommaire=3647035>.
- LOUND, Charles et MAIS, David, 2022, Managing risks to statistical quality in the UK Office for National Statistics. In : *European Conference on Quality in Official Statistics (Q2022)*. 8 juin – 10 juin 2022. [en ligne]. [Consulté le 4 juin 2025]. Disponible à l'adresse : <https://q2022.stat.gov.lt/scientific-information/papers-presentations/session-11>.
- MADELIN, Virginie et GUÉNÉ, Stéphane, 2024. Comment mesurer l'« inestimable » production de l'Insee ? In : *le blog de l'Insee*. [en ligne]. 22 octobre 2024. [Consulté le 4 juin 2025]. Disponible à l'adresse : <https://blog.insee.fr/comment-mesurer-l-iestimable-production-de-l-insee>.
- UNECE, 2019. Generic Statistical Business Process Model. In : *site statswiki de l'Unece*. [en ligne]. [Consulté le 4 juin 2025]. Disponible à l'adresse suivante : <https://statswiki.unece.org/display/GSBPM/GSBPM+v5.1>.



PRÉSENTATION DU NUMÉRO N14

Ce numéro N14 du Courrier des Statistiques nous fait voyager dans le temps. L'aventure commence au milieu du XX^e siècle, avec le récit de la construction du système statistique européen, au gré des avancées et crises de l'Europe, avant le tournant de 2004 lié à la falsification des comptes nationaux grecs. Nous remontons ensuite plus loin dans le passé, à une époque où les ordinateurs n'existaient pas. Comment les statisticiens communiquaient-ils avec les autres pays sur leurs travaux et résultats ? Quel rôle jouaient les bibliothèques statistiques dans ces échanges ? Enfin, le troisième papier raconte l'histoire du service statistique ministériel de la fonction publique à partir des années 1970, à travers notamment l'évolution de sa publication phare : le rapport annuel sur l'état de la fonction publique.

Nous revenons au présent avec le récit de l'expérimentation « Prophyl », menée en 2024 par le service de la statistique et de la prospective (SSP) du ministère de l'Agriculture. Ce dispositif original a pour ambition de simplifier le recueil d'informations sur l'usage de produits phytosanitaires (pesticides). Puis, un article à caractère général propose une grille d'analyse pour comprendre le temps de production des statistiques publiques, identifier les facteurs influents et mettre en avant les compromis à faire s'il fallait réduire les délais. Enfin, le dernier papier explique les principes et enjeux des démarches qualité menées à l'Insee pour assurer la sécurité et l'optimisation de ses processus de production. Ces démarches s'inscrivent dans les principes du code de bonnes pratiques de la statistique européenne.



ISSN 2107-0903

ISBN 978-2-11-162504-4



Insee
Mesurer pour comprendre