

Le Répertoire Statistique des Individus et des Logements (Résil)

Un nouvel univers de référence pour les statistiques démographiques et sociales



Olivier Lefebvre*

L'Insee construit actuellement un répertoire statistique des individus et des logements (Résil), pour moderniser son dispositif de production des statistiques démographiques et sociales, notamment en tirant davantage parti des données administratives.

Ce projet, déjà bien avancé, doit aboutir fin 2025 et être utilisé dès début 2026.

Ce répertoire permettra, à l'image de son homologue pour les entreprises et leurs établissements, de construire des bases de sondage ou encore de vérifier la couverture des données administratives, mais aussi de construire plus simplement et de façon plus assurée et sécurisée des fichiers enrichis par appariement de sources diverses. Différents processus de production de données pourront ainsi répondre à des exigences croissantes de couverture, de rapidité de traitement, de réactivité.

Pour le construire, divers prérequis sont nécessaires : recourir à différentes sources pour tendre vers l'exhaustivité, pouvoir y appliquer des traitements statistiques, souvent innovants, pour en assurer la qualité, bénéficier d'un cadre juridique pour protéger les données traitées, et enfin disposer d'un « mandat social », c'est-à-dire d'une légitimité, au-delà de la capacité technique ou juridique.

 INSEE is currently setting up a Statistical Register of Individuals and Dwellings (Résil) in order to modernise its system for producing demographic and social statistics, in particular by making greater use of administrative data.

This project is already well advanced. It should be completed by the end of 2025 and be operational from the beginning of 2026.

Like its counterpart for enterprises and their local units, this register will make it possible to set up sampling bases or to check the coverage of administrative data, as well as to create files enriched by matching different sources in a simpler and more reliable way. The various data production processes will thus be able to respond to the growing demands for coverage, processing speed and responsiveness.

There are a number of prerequisites for setting up such a system: the need to use different sources in order to aim for exhaustiveness; the ability to apply statistical processing, which is often innovative, in order to ensure quality; the need for a legal framework to protect the data processed; and finally, the need for a "social mandate", i.e. a legitimacy that goes beyond technical or legal capacity.

* Maître d'ouvrage du programme Résil, Insee.
olivier.lefebvre@insee.fr

Pour établir une statistique, il est nécessaire de collecter des données (Dupont, 2023), que ce soit de manière directe (par voie d'enquête) ou indirecte (en mobilisant des données administratives, voire des données détenues par des acteurs privés). Il faut également assurer la qualité de couverture de cette collecte : couvre-t-elle bien toutes les unités statistiques de notre champ d'intérêt ? Sans doublons ni omissions ? Les informations recueillies sont-elles relatives aux « bonnes » unités statistiques ?

Pour cela, disposer d'une liste de toutes les unités statistiques du champ d'observation, sans unités présentes à tort, s'avère extrêmement utile. En effet, on peut alors tirer un échantillon d'enquête dans cette base de sondage, ou encore vérifier que les données administratives utilisées sont exhaustives sur le champ, caractériser leur représentativité et les corriger le cas échéant afin d'éviter un biais lié à un défaut de couverture. De telles listes sont appelées « univers de référence ».

► Sous certaines conditions, les répertoires peuvent constituer un univers de référence

Un répertoire, c'est une liste exhaustive d'objets, avec très peu de variables (Rivière, 2022). S'il fallait dessiner un répertoire, il serait à la fois très haut, car il contient potentiellement toutes les observations d'un champ, mais très étroit, car peu de variables y sont gérées.

Les variables présentes dans un répertoire doivent permettre d'identifier sans ambiguïté les unités qu'il contient, notamment pour éviter des doublons, faciliter les mises à jour et permettre de les relier avec d'autres éléments du système d'information. Le répertoire forme la colonne vertébrale du système d'information. Un répertoire est un objet vivant, car mis à jour le plus souvent en continu ; il est néanmoins possible d'en extraire une photo, reflétant la situation un jour donné (souvent au 1^{er} janvier) qui constituera la base de l'univers de référence.

L'Insee s'inscrit dans une longue histoire et une longue expérience en matière de répertoires. Il gère ainsi le Répertoire national d'immatriculation des personnes physiques (RNIPP) depuis 1946 (Espinasse et Roux, 2022). Plus récemment, en 2019, l'Institut a construit et pris en charge le Répertoire électoral unique (REU) pour la gestion des listes électorales (Desmotes-Mainard, 2019).

Dans le domaine des entreprises, il a été confié en 1973 à l'Insee la gestion du répertoire Sirene¹ (Système Informatique pour le Répertoire des ENtreprises et des Établissements).

Ces répertoires, notamment Sirene et le RNIPP, pourraient-ils servir d'« univers de référence » ? Pas directement, car les informations concernant les sorties du champ (départ du territoire national pour les individus, fin d'activité pour une entreprise), ne sont pas connues ou le sont avec retard. En effet, le RNIPP ne contient pas l'adresse des personnes et il n'y a aucune obligation de signaler des départs hors de France. Dans le domaine des entreprises, la cessation officielle intervient souvent bien après l'arrêt effectif de l'activité économique.

¹ <https://www.insee.fr/fr/information/6675111>.



L'Insee a considéré qu'il fallait aller plus loin, en créant des répertoires statistiques situés en aval de ces répertoires administratifs.



L'Insee a donc considéré qu'il fallait aller plus loin, en créant en aval des répertoires administratifs, des répertoires « statistiques ». De tels répertoires permettent de mettre en œuvre des traitements pour des usages statistiques, sans conséquence sur les personnes ou entreprises concernées par les traitements : par exemple, un traitement statistique peut

conclure qu'une personne ne vit plus sur le territoire national, mais cela n'aura pas de conséquence sur son affiliation à un régime de retraite.

La statistique d'entreprises est en avance par rapport à la statistique démographique et sociale sur ce point. L'Insee a, en effet, mis en œuvre en 2012 le Système d'immatriculation au répertoire des unités statistiques (Sirus), un répertoire statistique des entreprises et d'établissements, enrichi de quelques informations collectées ou construites par la statistique publique – contours des groupes, niveau d'activité des entreprises, etc. (Hachid et Leclair, 2022).

En matière de statistiques démographiques et sociales, la situation était plus complexe : si on dispose d'une base de sondage construite à partir des fichiers fiscaux pour les enquêtes auprès des ménages, couvrant l'ensemble du champ des logements ordinaires², l'absence d'identifiant partagé et la non-exhaustivité des sources rendaient impossible de la considérer comme un « univers de référence » auquel comparer l'ensemble des sources administratives.

► **Faire face à de nouveaux défis et à de nouvelles opportunités : créer le répertoire statistique des individus et des logements**

Plusieurs facteurs ont convergé pour aller plus loin en matière de statistiques démographiques et sociales.

D'une part, des besoins croissants concernant l'exploitation des données administratives, assortis de nouvelles opportunités. Côté besoins, il était nécessaire de répondre à de nouvelles questions, ou plus rapidement à des questions récurrentes, ou encore de mieux rendre compte de la diversité des situations notamment territoriales, via la production de données plus fines que celles obtenues sur la base d'échantillons ; par exemple, le dispositif Filosofi³ apporte une information finement localisée sur les niveaux de vie, à partir de déclarations fiscales et de données sur les prestations sociales. Côté opportunités, il existe des sources plus nombreuses, plus accessibles, mieux structurées et documentées, de meilleure qualité, ainsi que des capacités de traitement informatique permettant de traiter de très gros volumes de données de manière sécurisée et dans des délais courts (exemple du PASRAU⁴). Comme la plupart des instituts nationaux de statistique, l'Insee investit donc fortement sur ce champ (**encadré 1**).

2 Un logement ordinaire est un logement défini par opposition à un logement en résidence offrant des services spécifiques (résidences pour personnes âgées, pour étudiants, de tourisme, à vocation sociale, pour personnes handicapées, etc.).

3 Fichiers LOcalisés SOciaux et FIScaux.

4 Le dispositif PASRAU (Prélèvement À la Source pour les Revenus AUTres) résulte de travaux de simplification et de rationalisation des déclarations sociales et de la nécessité de transmettre à la DGFiP des informations nécessaires au prélèvement à la source.

► Encadré 1. Regard sur les instituts nationaux de statistiques (INS) étrangers

Dans les instituts nationaux de statistiques étrangers, si on constate un besoin partagé d'une production de données statistiques plus riche et réactive, s'appuyant davantage sur des données administratives, on observe des réponses différentes, en fonction du contexte technique, organisationnel, culturel, ou encore juridique. On identifie plusieurs grands modèles :

- un système statistique très intégré, qui utilise de longue date des registres de population administratifs et un identifiant partagé

Les pays dotés de registres de population administratifs reposant sur une obligation de déclaration de changement de domicile et ceux qui s'appuient sur un identifiant personnel partagé par toutes les administrations ont pu construire un système statistique très intégré. L'utilisation courante de la formulation « registre de population » (« register » en anglais) induit une certaine confusion dans la mesure où il ne précise pas sa finalité (administrative ou statistique).

En Finlande par exemple, la production de données statistiques s'appuie à 95 % sur des données issues de registres ou de sources administratives. Le recensement de la population finlandais est, depuis 1980, entièrement fondé sur ce type de données. Le système d'information sur la population rassemble de nombreuses données caractérisant les personnes tout en permettant une interconnexion avec d'autres fichiers (voir le site internet de l'institut de statistique finlandais). Il s'inscrit dans un État et une société où l'interconnexion de fichiers individuels ne soulève pas de problèmes techniques ou organisationnels majeurs, ni de problème d'acceptation par la population. Il n'est ni envisageable ni envisagé que la France s'oriente vers un tel dispositif.

- aux Pays-Bas : un système très intégré avec registres, sources administratives et données d'enquêtes

L'institut de statistique hollandais (CBS**) s'inscrit dans une démarche générale d'utilisation des sources de données disponibles, le « System of social statistical datasets (SSD) » (Bakker et alii, 2014). C'est un système de registres et d'enquêtes interconnectés et normalisés. Il contient une mine

* <https://dvv.fi/en/personal-data>.

** <https://www.cbs.nl/nl-nl/>.

*** <https://www.cbs.nl/en-gb/our-services/customised-services-microdata/microdata-conducting-your-own-research>.

d'informations sur les personnes, les ménages, les emplois et les prestations, les pensions, l'éducation, les hospitalisations, les rapports de criminalité, les logements, les véhicules, etc.

Aux Pays-Bas, il s'agit de la plus importante source de statistiques sociales officielles, sous forme de résultats agrégés qui préservent la confidentialité des données utilisées. Les données individuelles très détaillées auxquelles accèdent les chercheurs restent dans l'environnement sécurisé géré par le CBS, qui vérifie systématiquement que les données ou résultats exportés ne présentent pas de risque de rupture de confidentialité***.

- en Nouvelle-Zélande : un système d'identification statistique des individus et des logements afin de faciliter les appariements, en particulier dans un objectif de recherche

Il n'existe ni identifiant individuel partagé ni répertoire administratif mobilisable (Bycroft et alii, 2022). Un premier répertoire statistique d'individus a été constitué à partir des données d'état civil (naissances et décès), des données aux frontières (immigration-émigration) et des données fiscales. Il permet de relier, à la demande, plusieurs sources de données administratives entre elles, à des finalités statistiques ou de recherche. Ce dispositif a été très utile pour le recensement en 2018, perturbé par des difficultés de collecte susceptibles d'entraîner des biais dans les données produites. Il a permis d'obtenir des résultats statistiques en combinant des données (administratives ou d'enquêtes) sur l'éducation, le marché du travail, les prestations, la justice, la santé et la sécurité, les migrations et les données commerciales.

Depuis 2021, une rénovation est en cours pour disposer d'un véritable répertoire statistique des individus et des logements plus intégré et plus facile à gérer.

L'INS néo-zélandais est très soucieux de la communication réalisée sur ces techniques et la mobilisation des données : des informations sont faites pour rassurer les utilisateurs sur la qualité et pour obtenir l'adhésion de toute la population face à ce nouveau type de collecte, en particulier la population autochtone maori.

D'autre part, la suppression de la taxe d'habitation (TH) sur les résidences principales a accéléré le processus et imposé de trouver de nouvelles solutions pour les statistiques construites à partir de celle-ci.

L'Insee s'est appuyé depuis de longues années sur un fichier issu de la gestion de la TH, avec des utilisations de plus en plus riches. Ce fichier a d'abord permis d'établir une liste de logements dans chaque commune, afin de préparer et de contrôler les enquêtes annuelles de recensement. Il a aussi permis de calculer la population des communes en actualisant



La suppression de la taxe d'habitation sur les résidences principales a imposé de trouver de nouvelles solutions pour les statistiques construites à partir de celle-ci.



les résultats des enquêtes de recensement. Il a ensuite été utilisé comme base de sondage pour les enquêtes auprès des ménages et servi d'ossature au Fichier Démographique sur les Logements et les Individus (Fidéli) et de référence pour constituer les contours des ménages, préalable essentiel pour le calcul des niveaux de vie, y compris à des niveaux géographiques fins (Lamarche et Lollivier, 2021).

L'objectif avec Résil est donc de construire un dispositif rendant a minima les mêmes services que le fichier issu de la TH, c'est-à-dire réaliser le recensement, échantillonner les enquêtes auprès des ménages, reconstituer les niveaux de vie et créer des fichiers composites d'étude. L'Insee en a profité pour aller plus loin, sur trois axes : renforcer la pérennité du dispositif (ne pas revivre l'épisode de la suppression de la TH), progresser sur le contrôle du champ couvert et rendre le « service d'univers de référence » pour les processus de collecte et d'exploitation de données administratives.

La cible est donc de construire un dispositif plus robuste, à plusieurs titres :

- créer un répertoire, avec des unités identifiées sans ambiguïté et stables dans le temps, donc de meilleure qualité ;
- étendre le champ aux logements non couverts par la taxe d'habitation, notamment les logements des communautés (Ehpad⁵, internats, etc.) ;
- mobiliser plusieurs sources pour alimenter et mettre à jour ce répertoire, pour une couverture plus complète du champ d'observation, en assurant la continuité en cas de changement sur les données. La suppression de la taxe d'habitation a fait prendre conscience à l'Insee que ce risque était réel.

Il est également possible, grâce au caractère central du répertoire, de développer les appariements de données (et donc de production multi-sources) et de contrôler la couverture des sources administratives, au regard de leur utilisation à des fins statistiques.



La cible est le répertoire statistique des individus et des logements (Résil), pour créer les univers de référence attendus pour la statistique démographique et sociale et faciliter les appariements.



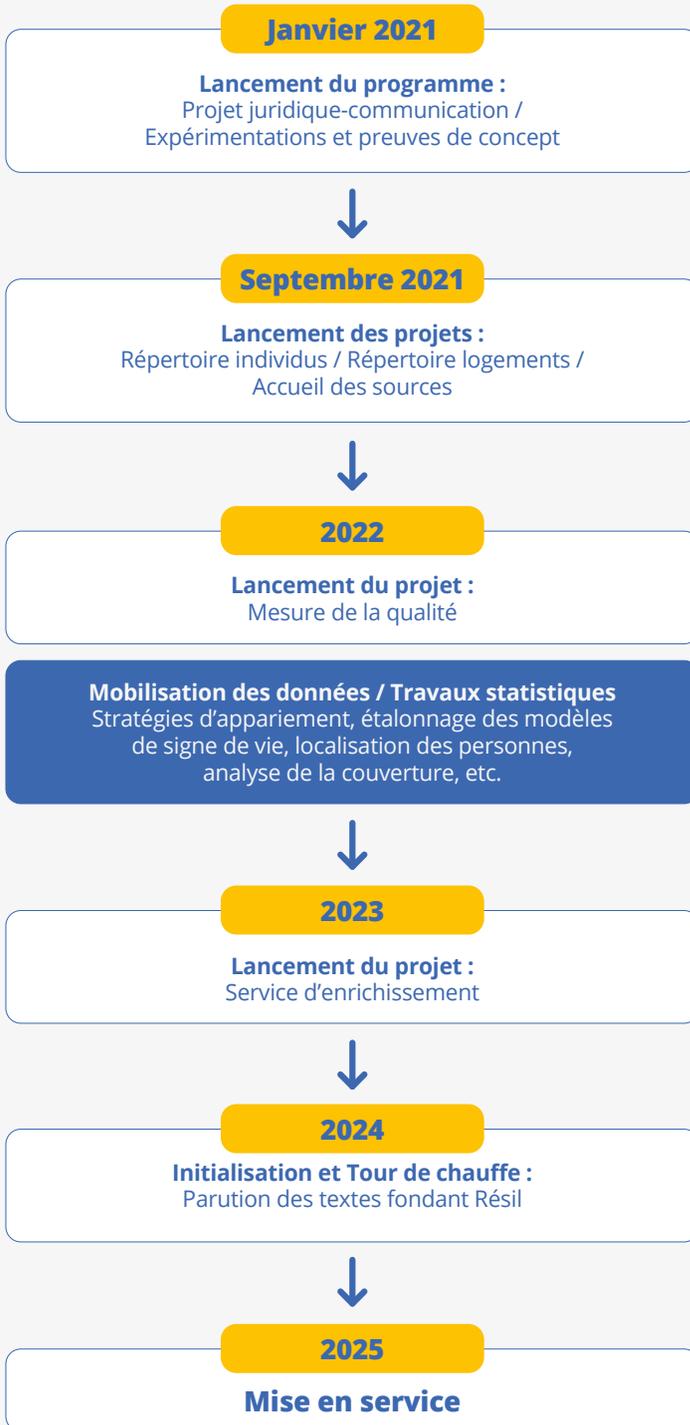
La cible est le répertoire statistique des individus et des logements (Résil), permettant de créer les univers de référence attendus pour la statistique démographique et sociale, mais aussi de faciliter les appariements de données administratives entre elles ou avec d'autres données, principalement d'enquêtes. L'objectif est d'y parvenir en 2025 (*figure 1*).

Selon les termes du décret créant Résil⁶, ce dernier « a pour finalité, en vue de contribuer au débat public ainsi qu'à l'élaboration et

⁵ Ehpad : Établissement d'hébergement pour personnes âgées dépendantes.

⁶ Voir les références juridiques en fin d'article.

► **Figure 1 - Résil : un projet qui se construit étape par étape**



à l'évaluation des politiques publiques, de renforcer la capacité de l'Institut national de la statistique et des études économiques et des services statistiques ministériels à produire des données et études statistiques, en permettant l'établissement d'un répertoire national de la population et des logements et en facilitant les appariements de données administratives avec d'autres sources de données ».

► Des usages variés, essentiels pour la construction des statistiques démographiques et sociales

Résil est donc une infrastructure de production permettant de répondre à plusieurs objectifs exclusivement statistiques.

Il offrira, uniquement au service statistique public, un service d'appariement de données avec des usages multiples⁷ et pouvant notamment servir (Dupont, 2023) à :

- alléger la collecte d'information par enquête en ne posant des questions que sur des aspects non couverts par les données administratives ;
- appairer des données administratives entre elles, afin de produire des statistiques à une échelle fine, impossible à produire sur échantillon ;
- enrichir un fichier par des variables complémentaires permettant d'approfondir les analyses (par exemple l'ajout d'informations sur le revenu dans l'enquête sur les ressources et les conditions de vie (SRCV)) ;
- éclairer des aspects méthodologiques particuliers : par exemple, appairer les fichiers de l'enquête Emploi et le fichier historique des demandeurs d'emploi afin de mesurer la différence entre les concepts de chômeur au sens du BIT⁸ et demandeur d'emploi inscrit à France Travail ;
- évaluer des politiques publiques (suivi de trajectoire de bénéficiaires d'aides particulières).

Ces appariements seront meilleurs et moins coûteux grâce à la présence d'identifiants communs ; on pourra par ailleurs en mesurer plus facilement la représentativité et la qualité.



Résil permettra également de mesurer la qualité des sources administratives qui constituent une des ressources principales de la statistique publique.



Résil permettra également de mesurer la qualité des sources administratives qui constituent une des ressources principales de la statistique publique. Il sera possible de comparer le champ effectivement couvert par une source statistique à la liste des individus ou logements présents dans Résil, et ainsi de détecter d'éventuels défauts de couverture.

Dans le prolongement des dispositifs existants, la base de sondage dans laquelle tirer des échantillons pour les enquêtes réalisées par le service statistique public auprès des ménages sera issue de Résil. La

⁷ Voir l'article de Koumarios, Lefebvre et Malherbe sur les appariements dans ce même numéro.

⁸ BIT : Bureau International du Travail.

couverture sera mieux assurée qu'auparavant, à la fois par la diversité des sources, l'ajout des personnes vivant dans les communautés et la prise en compte plus rapide des décès ; il sera également possible d'ajouter à la base de sondage des variables d'autres origines que les sources fiscales, ce qui rendra l'échantillonnage plus précis.

Résil fournira l'information permettant de préparer et réaliser les enquêtes annuelles de recensement, et d'en extrapoler les résultats, à l'instar de l'utilisation actuelle des fichiers issus de la taxe d'habitation (**encadré 2**). Résil pourra ainsi être mobilisé pour faciliter la production d'indicateurs démographiques plus précoces.

Enfin, Résil permettra de progresser sur la cohérence des traitements et des données produites, mais aussi sur l'efficacité de ces traitements, via l'utilisation d'outils partagés et performants. En facilitant le rapprochement de données, mais aussi leur confrontation, en unifiant les données de référence et les marges de calage⁹, il permet de « casser les silos » du système d'information actuel organisé par source.

► **Encadré 2. Résil et le Recensement de la population, un partenariat gagnant-gagnant**

Résil va produire, en remplacement de la taxe d'habitation, les données dont le recensement a besoin pour préparer, contrôler et extrapoler les enquêtes de recensement. Il pourrait permettre en particulier de fournir des estimations plus précoces, comme demandé par de nombreux utilisateurs et par Eurostat.

Inversement, le recensement permet d'évaluer la qualité des sources utilisées en entrée de Résil, qu'il s'agisse des individus ou des logements.

Par ailleurs, des méthodes statistiques dites d'estimation par système dual (Zhang et Dunne, 2017) devraient permettre à terme de mesurer la couverture de Résil d'une part, du recensement d'autre part, en confrontant ces deux sources et donc d'identifier des biais de couverture.

Comme dans d'autres pays, Résil devrait permettre de moderniser le système de recensement ; cependant, il est trop tôt pour dire comment se traduira cette modernisation et dans quel calendrier elle se fera.

► **Que contiendra Résil ?**

Concrètement, Résil sera constitué de deux sous-répertoires statistiques distincts, reliés l'un à l'autre : un répertoire des individus et un répertoire des logements. Ils seront mis à jour régulièrement sur les naissances et les décès et avec les sources fiscales et sociales (données sur les prestations sociales et familiales de la Cnaf¹⁰, DSN¹¹, PASRAU) ou d'autres sources sur un public spécifique, comme les fichiers d'inscription dans l'enseignement supérieur.

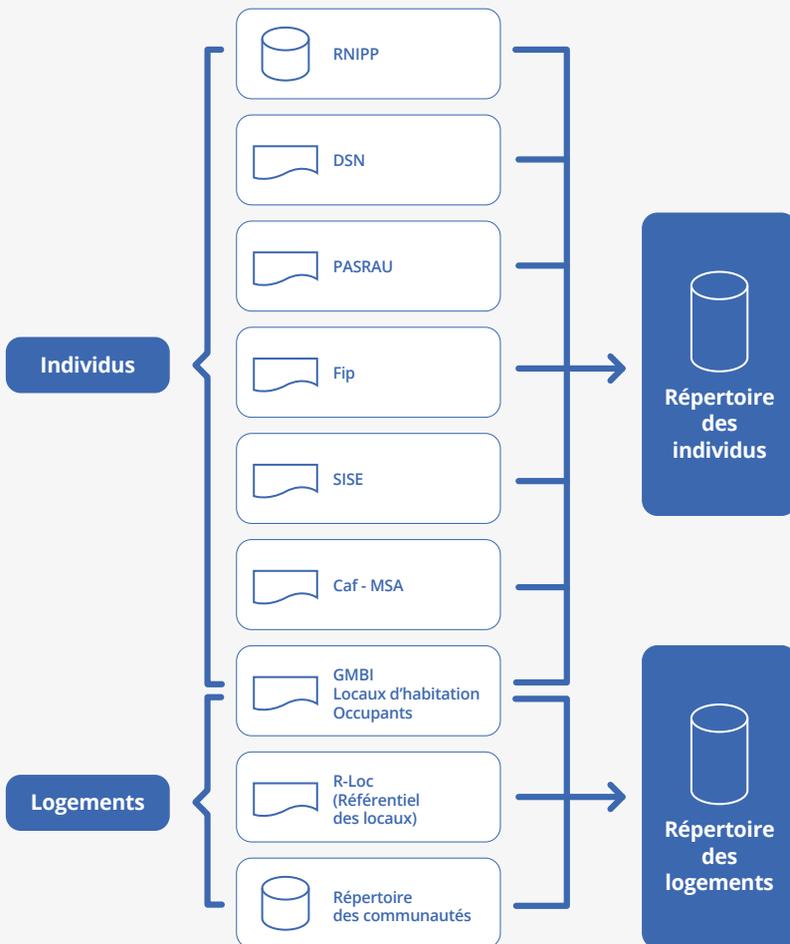
Un point essentiel : dans ces sources, on ne retient que des données d'identification des individus et des logements, des données d'adresse et des liens entre individus et logements (**figure 2**).

⁹ Le calage sur marges est une technique statistique visant à améliorer la précision des enquêtes par sondage. Elle consiste à modifier les poids de sondage des individus de l'échantillon afin que les totaux pondérés sur l'échantillon de certaines variables correspondent aux totaux connus pour ces variables sur l'ensemble du champ d'observation (la population, le parc de logements, les entreprises) (Deville et alii, 1992).

¹⁰ Cnaf : Caisse nationale d'allocations familiales.

¹¹ DSN : Déclaration sociale nominative.

► **Figure 2 - Résil et ses sources d'alimentation***



RNIPP : Répertoire National d'Immatriculation des Personnes Physiques
DSN : Déclaration sociale nominative
Dispositif PASRAU : Prélèvement À la Source pour les Revenus AUTres
Fip : Fichier d'imposition des personnes
SISE : Système d'information sur le suivi de l'étudiant
Caf : Caisses d'allocations familiales
MSA : Mutualité sociale agricole
GMBI : Gérer mes biens immobiliers

* Cette liste de sources peut évoluer dans le temps, sous réserve d'un avis favorable de la Cnil et du Cnis.

À partir de ces répertoires seront produites annuellement des « photographies » composant l'univers de référence :

- la liste des individus présents sur le territoire national au 1^{er} janvier ;
- la liste des logements situés sur le territoire national au 1^{er} janvier, et leur statut (résidence principale, secondaire, logement vacant) ;
- la liste des ménages¹² et leur composition au 1^{er} janvier.

Les listes des individus et des logements serviront de référence pour la statistique démographique et sociale. Les listes de ménages sont indispensables pour construire des données telles que les niveaux de vie ou pour réaliser des enquêtes.

Les informations contenues dans les répertoires Résil seront essentiellement des clés d'identification pour assurer le rôle de liste de référence (éviter les oublis et les doublons) et pour réaliser les appariements :



Le numéro d'inscription au RNIPP (le NIR ou numéro de sécurité sociale) ne sera pas stocké dans Résil.



- des identifiants d'individus : le code statistique non signifiant (CSNS) (Espinasse et alii, 2023) créé par la loi pour une République numérique en 2016¹³ pour faciliter les appariements entre sources au sein du service statistique public, un identifiant spécifique et strictement interne à Résil, pérenne pour la gestion de l'historique, les identifiants des sources utilisées dans Résil (pour les individus et les foyers). Le numéro d'inscription au RNIPP (le NIR ou numéro de sécurité sociale) ne sera pas stocké dans Résil ;

- des identifiants de logements : un identifiant spécifique à Résil pérenne, les identifiants des sources utilisées pour Résil ;
- chaque logement aura un identifiant d'adresse issu du référentiel d'adresses de l'Insee du type « BZ140JD » (et non l'adresse en clair « 8 rue Zéphyrin Brioché à Gleux-lès-Lure, département de la Haute-Saône »), ce qui le rend inexploitable en dehors de l'Insee ;
- les données d'état civil : nom, prénom, date et lieu de naissance, le cas échéant date du décès ;
- des liens entre les individus et leur logement ou leur adresse¹⁴ d'habitation, avec, le cas échéant, plusieurs logements possibles pour un même individu selon les sources ; in fine une résidence principale sera déterminée pour chaque individu.

Quelques autres variables permettant la gestion du répertoire et la mesure de la qualité :

- dates de mise à jour des données ;
- dates d'effet (date de début et date de fin) pour certaines variables de Résil pour lesquelles on souhaite conserver un historique ;

¹² Un ménage regroupe l'ensemble des personnes partageant un même logement.

¹³ Voir les références juridiques en fin d'article.

¹⁴ Dans certains cas, il ne sera pas possible de faire la distinction entre plusieurs logements situés à la même adresse.

- indicateur de présence sur le territoire français ;
- présence de la personne ou du logement dans chaque source administrative (oui/non).

Résil ne contiendra aucune autre information. Les données telles que le revenu, l'état matrimonial, la profession, la surface des logements, etc. figureront dans des bases spécifiques indépendantes de Résil, et ne seront mobilisées qu'à la demande dans le cadre d'un traitement distinct. Résil ne sera donc pas une « méga-base » contenant tout ce que l'on sait sur chaque individu ou chaque logement.

Le répertoire Résil s'appuie sur quatre piliers majeurs : des données d'origines diverses, des traitements statistiques permettant de les transformer en un répertoire statistique de qualité, un fondement juridique solide et un « mandat social ».

► Quatre piliers pour Résil

Premier pilier : des sources d'informations diverses pour un résultat robuste

Résil utilisera plusieurs sources de données pour assurer :

- la meilleure couverture possible de la population (aucune source administrative n'est exhaustive, et aucune n'est parfaitement conforme aux concepts statistiques de population résidente¹⁵) ;
- une localisation plus précise des individus et une meilleure appréhension des résidences multiples ;
- la pérennité du dispositif au défaut d'une source, voire sa transformation ou sa disparition, pour ne pas subir de rupture de collecte (exemple de la suppression de la taxe d'habitation).

Les résultats des premières expérimentations confirment l'intérêt de mobiliser chacune de ces sources, en sus de la seule source fiscale, en ce qui concerne la couverture de la population. Si on prend l'Enquête Annuelle de Recensement (EAR) comme référence, le gain global de couverture¹⁶ est de l'ordre de 2 points chez les individus de plus de 18 ans, mais il s'élève à 10 points pour les 21-25 ans. La couverture par âge est ainsi plus homogène qu'avec la seule source fiscale.

Pour les personnes vivant en **communauté**¹⁷ (environ 1,3 million de personnes dans des maisons de retraites, cités universitaires, internats, foyers de travailleurs, établissements pénitentiaires, communautés religieuses, etc.), le taux de couverture progresse de 10 points, de 80 % à 90 %.

¹⁵ La population résidente comprend toutes les personnes résidant en France, quelle que soit leur nationalité et leur situation, à partir du moment où elles sont en France depuis au moins un an, ou, si elles viennent d'arriver, qu'elles ont l'intention d'y rester pour au moins un an. En revanche, les personnes de passage (touristes, travailleurs saisonniers ou étudiants étrangers venant pour une année scolaire de 9 mois) n'y figurent pas. Cette définition correspond aux règles internationales et permet ainsi des comparaisons entre pays.
Les personnes sans résidence habituelle dans un autre pays sont comptées dans la population résidente de la France si elles s'y trouvent à la date de référence du calcul de cette population.

¹⁶ Le taux de couverture est estimé par la part des personnes recensées retrouvées dans la source fiscale ou les autres sources.

¹⁷ Voir définition (au sens du recensement de la population) : <https://www.insee.fr/fr/metadonnees/definition/c1134>.

► **Tableau des sources utilisées dans le processus d'alimentation de Résil**

Source	Justification
Répertoire national d'identification des personnes physiques (RNIPP)	<ul style="list-style-type: none"> • Mise à jour de la liste des individus avec les naissances et décès. • Mise à jour des données d'état civil.
Fichiers fiscaux : <ul style="list-style-type: none"> • Fichier d'imposition des personnes (Fip) • Fichier permanent des occurrences de traitement des émissions (Pote) • Fichier de mise à jour des informations cadastrales (Majic) • R-Loc (Référéntiel des Locaux) • Fichier « Gérer mes biens immobiliers » (GMBI) 	<ul style="list-style-type: none"> • Connaissance du parc de logements (Référéntiel des Locaux, relais progressif des données du Cadastre). • Connaissance de leur usage (résidence principale ou secondaire) et de leurs occupants (via les données de GMBI pour l'occupant déclaré, Fip et Pote pour les autres occupants du logement). • Sont retraitées les données identifiantes et localisantes, à l'exclusion de toute information sur les bases fiscales, revenus fiscaux, impôts dus ou payés.
Déclaration Sociale Nominative (DSN)	La fréquence mensuelle de cette source permet une mise à jour fiable du répertoire d'individus (durée de présence sur le territoire, prise en compte de jeunes actifs en complément des données fiscales, prise en compte plus précoce des changements d'adresse).
Prélèvement à la source pour les revenus autres (PASRAU)	La fréquence mensuelle de cette source permet une mise à jour fiable du répertoire d'individus (durée de présence sur le territoire, prise en compte de bénéficiaires de revenus autres que salariaux, en complément des données fiscales).
Fichier de référence des allocataires Caf (Caisses d'allocations familiales)	Cette source apporte des éléments sur la composition des foyers d'allocataires, éléments indispensables à une information fiable sur les ménages (notamment leur composition).
Fichier d'allocataires de la MSA (Mutualité sociale agricole)	<ul style="list-style-type: none"> • Cette source apporte des éléments sur la composition des foyers d'allocataires, éléments indispensables à une information fiable sur les ménages (notamment leur composition). • Le champ de la MSA est complémentaire de celui des Caf.
Fichier annuel d'inscriptions dans l'enseignement supérieur	Ce fichier permet d'améliorer la couverture du répertoire sur les jeunes de 17 à 25 ans, parfois absents des sources fiscales ou localisés de manière ambiguë.
Répertoire des communautés constitué pour la réalisation du recensement de la population	Les communautés constituent la résidence principale d'environ 1,3 million de résidents sur le territoire. Il est donc essentiel de compléter le parc de logements ordinaires par celui des communautés. Le répertoire des communautés est un intrant du répertoire des logements de Résil.
Enquêtes statistiques de contrôle de la qualité du répertoire	L'Insee pourrait être amené à réaliser des enquêtes statistiques de contrôle de la couverture du répertoire, voire des enquêtes de complétude sur des territoires mal couverts par les sources administratives.



Pour chacune des sources, le choix des données retenues est sélectif.

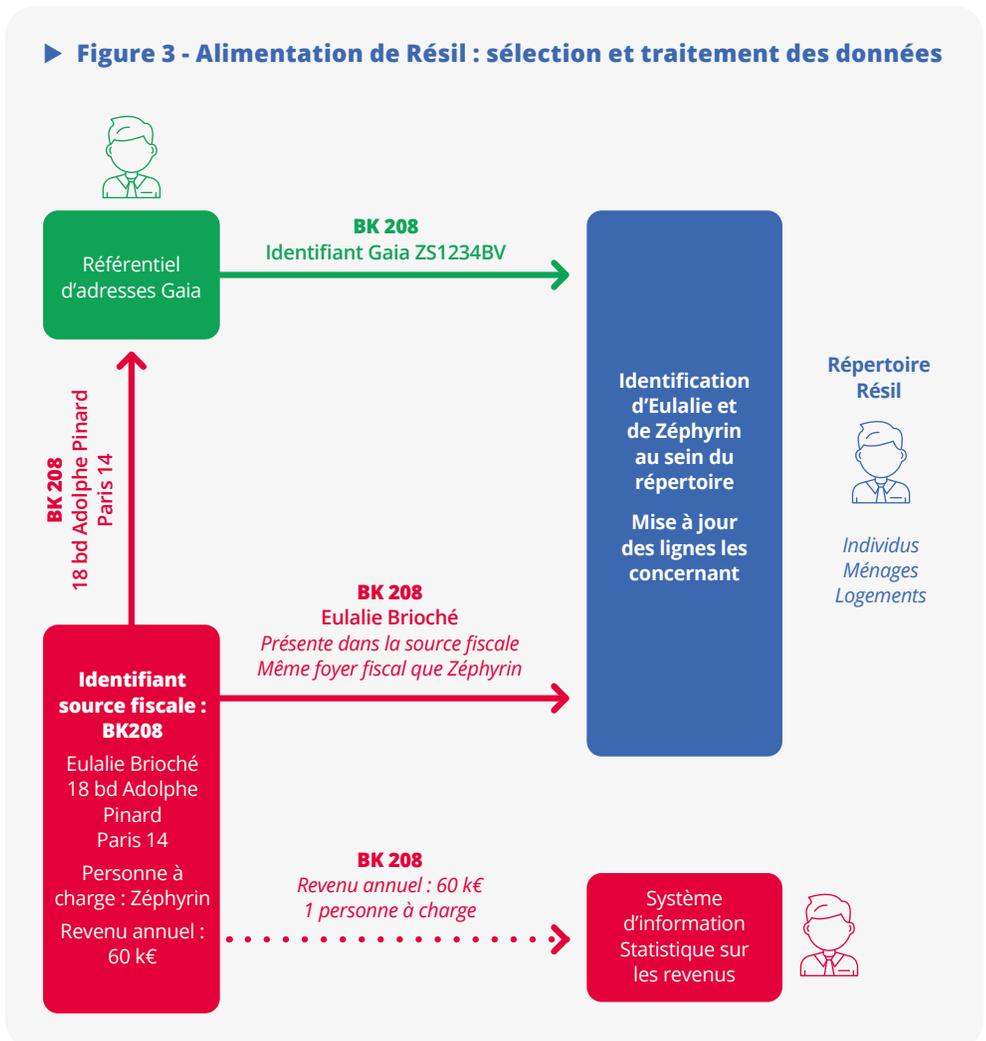


Pour chacune des sources (**voir Tableau des sources**), le choix des données retenues est sélectif. Les données seront sélectionnées et orientées en entrée du système d'information de l'Insee pour n'alimenter que les processus qui en ont besoin. Le dispositif d'accueil fonctionne comme une gare de triage. Lors de cette étape, les NIR présents dans les sources seront remplacés par les codes statistiques non significatifs

(CSNS) correspondants : les données d'identité seront orientées vers Résil, soit pour mise à jour, soit pour garantir la qualité de l'identification. Les données d'adresses seront également traitées à la source pour les remplacer par un identifiant non significatif provenant du répertoire d'adresses de l'Insee.

Les données « métier », accompagnées des identifiants non significatifs de personnes et d'adresses, seront intégrées dans les systèmes d'information pour produire des données statistiques (**figure 3**).

► **Figure 3 - Alimentation de Résil : sélection et traitement des données**



Résil s'inscrit dans un principe de minimisation des données traitées : il comportera très peu de variables et permettra de supprimer les données directement identifiantes des autres systèmes d'information de l'Insee pour les remplacer par des pseudonymes. Par ailleurs, Résil ne contiendra aucune donnée statistique permettant de caractériser les individus et logements, celles-ci étant traitées uniquement par les applications destinées à produire des données statistiques.

Pour alimenter le répertoire, un outil modernisé de mise à disposition des données administratives utilisées à des fins statistiques a été développé.



L'arrivée de Résil est une opportunité pour rationaliser le dispositif d'accueil des données administratives.



Pour la plupart des sources mobilisées, le répertoire Résil est un nouvel « utilisateur » en complément des producteurs de données statistiques sur l'emploi, les revenus ou le logement. Mais l'arrivée de Résil est une opportunité pour rationaliser le dispositif d'accueil des données administratives, constitué actuellement de plusieurs dispositifs juxtaposés, liés à la fois aux sources et aux usages. Outre le projet de construction des répertoires, le projet Résil s'accompagne d'un projet de modernisation et d'unification du dispositif d'accueil et de

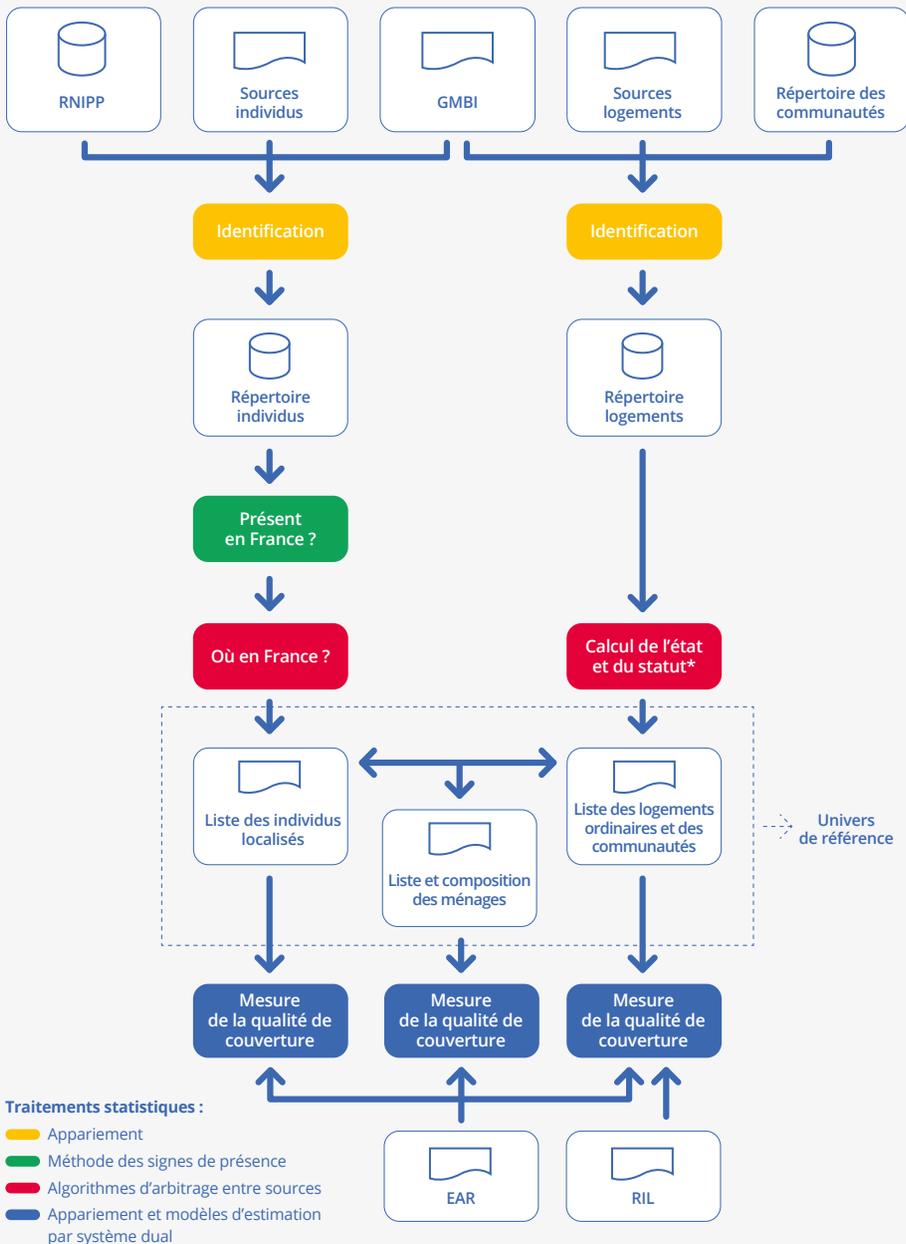
structuration des données administratives en données statistiques brutes facilement exploitables pour produire des chiffres relatifs aux unités statistiques (individus, logements et ménages). Un tel dispositif repose sur l'outil Accueil-Réception-Contrôle, dit ARC¹⁸. Le principe est d'utiliser le même outil pour l'accueil des différentes sources et leur mise à disposition vers les différents utilisateurs au sein du service statistique public. Cette rationalisation permet d'investir sur les performances et l'enrichissement fonctionnel d'un tel outil, sans sacrifier sa sécurité ni sa capacité d'adaptation rapide à des changements dans les sources, voire l'accueil de nouvelles sources. Compte tenu du nombre de sources à accueillir, de leur volume et de leur fréquence (au moins trois sources, parmi les plus volumineuses, sont mensuelles), mais aussi de la nécessité d'une alimentation rapide des systèmes de production d'indicateurs statistiques, la robustesse et les performances de traitement d'un tel outil sont cruciales. Sa capacité d'adaptation rapide aux transformations des sources administratives (dictées par les politiques qu'elles mettent en œuvre et non pas par les statistiques qu'elles permettent de produire) et à l'apparition de nouvelles sources est également essentielle.

Second pilier : pour assurer la qualité du répertoire, des traitements statistiques performants et innovants, inspirés par des pratiques internationales

Disposer des données est indispensable mais ne suffit pas. Il faut également des outils performants pour transformer ces données en un répertoire puis en un univers de référence (liste des individus effectivement résidents à la date du 1^{er} janvier, liste des logements habitables, liste des ménages), qui soit d'une qualité propre à son utilisation statistique.

¹⁸ Voir l'article de Lefebvre, Soulier et Tortosa sur l'accueil des données administratives dans ce même numéro.

► **Figure 4 - Des sources à l'univers de référence : une succession de traitements statistiques**



* Calcul de l'état et du statut : il s'agit de déterminer si le logement est habité ou non, s'il s'agit d'une résidence principale, secondaire, ou d'un logement vacant.

Il s'agit d'abord d'identifier les individus, pour « mettre à jour les bonnes lignes du fichier », sur la base d'informations parfois incomplètes ou entachées d'erreurs ; c'est tout l'enjeu des moteurs d'identification (*figure 4*). Résil s'appuiera sur un processus plus discriminant que celui utilisé pour le CSNS, dans la mesure où il pourra utiliser, pour les cas les plus douteux, des informations complémentaires telles que la composition du ménage ou l'adresse de la personne.

Ensuite, il faut distinguer, parmi les individus présents dans le répertoire, ceux qui résident effectivement sur le territoire national et pour chacun d'eux déterminer leur résidence principale. Cela permet de construire l'univers de référence.

La personne vit-elle toujours sur le territoire national ? La méthode des « signes de présence » (ou « signes de vie » dans la littérature académique et professionnelle) consiste à mobiliser, en complément des données d'état civil incontestables (naissances, décès), l'information relative à la présence des individus dans telle ou telle source administrative, avec une adresse située sur le territoire. Si un individu non décédé dans le RNIPP est absent dans toutes les sources administratives, il y a une forte probabilité qu'il ne réside plus sur le territoire. S'il ne figure que dans une partie des sources dans lesquelles on devrait normalement le retrouver et possède une adresse à l'étranger dans au moins une source, il y a une probabilité non négligeable qu'il ait quitté le territoire. Chacun des signes de présence dans les sources peut être pondéré par la qualité de l'identification de l'individu¹⁹ et sa pertinence au regard des individus concernés. Par exemple, si on utilise le fichier des étudiants, il sera pertinent pour les 18-25 ans. Cette méthode est pratiquée dans plusieurs pays, tels que l'Estonie, l'Irlande, l'Italie ou l'Australie. Elle est encouragée par Eurostat²⁰, en lien avec le développement de l'usage des données administratives.

Des règles de décision en cas d'adresses multiples

Quand une personne a des adresses différentes d'un fichier à l'autre, il convient également de déterminer quelle est l'adresse de sa résidence principale. La divergence peut résulter d'un décalage dans la mise à jour des fichiers administratifs (la personne a déménagé mais l'information n'a pas encore été prise en compte) ou d'une multiple résidence (les « célibataires géographiques », les étudiants logés sur leur lieu d'études mais encore attachés au foyer fiscal de leurs parents, les enfants en garde alternée, etc.). Les règles de décision à retenir doivent permettre de localiser les personnes à leur résidence principale, selon les concepts prescrits pour les comparaisons internationales et mis en œuvre pour le recensement de la population.

Pour travailler sur des données stables, il est important d'avoir une photo donnant la situation au 1^{er} janvier.

L'univers de référence provenant de Résil doit correspondre à une situation stable pour permettre de partager son utilisation. Or le répertoire est vivant, ses mises à jour sont régulières. Il faut donc « prendre une photo » du répertoire reflétant la situation à une date donnée, celle du 1^{er} janvier par convention.

¹⁹ Si les traits d'identité de la source ne permettent pas d'identifier de façon sûre un individu du répertoire, on donnera un poids moins important au signe de présence dans cette source, car il pourrait s'agir d'une erreur d'appariement.

²⁰ Eurostat est l'Office statistique de l'Union européenne.



Pour travailler sur des données stables, il est important d'avoir une photo donnant la situation au 1^{er} janvier.



Résil sera mis à jour à partir de sources diverses qui n'arriveront pas toutes au même moment. Plutôt que d'attendre la dernière source pour développer la photo, il a été décidé de produire trois versions de chaque photo, au fur et à mesure de l'arrivée des données.

Par exemple, l'univers de référence provisoire au 1^{er} janvier 2025 sera produit à l'été 2025, la version semi-définitive en janvier ou février 2026, la version définitive au début de l'été 2026.

Chacun des utilisateurs pourra donc procéder à son propre arbitrage entre fraîcheur et exhaustivité en fonction de ses utilisations de l'univers de référence.

Il est également nécessaire de mesurer la qualité de couverture, ce que tout répertoire doit pouvoir faire. Cela se fait en confrontant avec les enquêtes annuelles de recensement et en se fondant sur la méthode d'estimation par système dual.

Le principe de cette estimation est de confronter les deux collectes, Résil d'une part et l'enquête de recensement d'autre part, de décompter les personnes présentes dans les deux sources et celles présentes dans l'une ou l'autre des sources afin d'en déduire le nombre de personnes absentes des deux sources, donc la taille de la population totale, sous l'hypothèse notamment d'indépendance des deux collectes et d'absence de personnes comptées à tort dans chacune des deux collectes.

Cette méthode est utilisée dans plusieurs pays pour estimer la couverture des recensements exhaustifs. Elle se développe actuellement pour mesurer la couverture des répertoires, par exemple en Italie pour le répertoire rassemblant les registres municipaux de population.

Troisième pilier : un fondement juridique clair et solide, qui autorise le traitement et protège les données qui en sont issues

Résil s'inscrit dans le cadre juridique national et européen relatif à la production de statistiques publiques et à la protection des données individuelles (loi du 7 juin 1951 sur l'obligation, la coordination et le secret en matière de statistiques, Règlement Général pour la Protection des Données (RGPD) et Loi Informatique et Libertés²¹). Aucune disposition de nature législative n'est nécessaire ; l'environnement cité ci-dessus garantit l'accès aux données nécessaires et les obligations relatives à leur protection.

L'Insee a considéré que le niveau juridique approprié pour fonder un traitement de cette nature est celui d'un décret en Conseil d'État. Un arrêté du responsable de traitement, en conformité avec le RGPD, aurait pu suffire s'il s'était agi d'un texte purement technique, mais les interrogations de nature plus politique ou sociétale que peut susciter le dispositif justifiaient un texte de ce niveau juridique, en offrant

²¹ Voir les références juridiques en fin d'article.

un examen critique préalable, par la Cnil²² et par le Conseil d'État, renforçant ainsi la légitimité du dispositif.

Le décret en Conseil d'État créant Résil a été publié au Journal Officiel le 7 janvier 2024. Il comporte des dispositions prises dans la plupart des textes créant des traitements de données personnelles : la création du traitement et sa finalité, la liste des variables et leur durée de conservation, les utilisateurs du répertoire, les destinataires des données qu'il gère et des fichiers qu'il produira, et les dispositions relatives à la sécurité du système d'information. Il comprend également plusieurs dispositions plus spécifiques à Résil : la définition des appariements, les conditions d'évolution de la liste des sources, la référence expresse aux exigences déontologiques propres au métier de statisticien. Un arrêté²³ établissant la liste des sources utilisées pour construire et mettre à jour Résil, pris en application de ce décret, a également été publié au Journal Officiel le 7 janvier 2024.

Quatrième pilier : un mandat social à conforter en permanence

Tout ce qui précède vise à conférer à l'Insee la capacité, technique ou juridique, à construire et gérer le répertoire. Cette capacité doit s'accompagner d'une légitimité, ou d'un mandat social. Les corps constitués et autorités compétentes confèrent une partie de cette légitimité, à travers les textes qui encadrent le fonctionnement de la statistique publique et le traitement Résil en particulier. Cependant, cette légitimité ne serait pas suffisante sans un « mandat social », preuve que ce traitement est reconnu et accepté par la population, qui fait confiance à l'Insee pour le mettre en œuvre.



L'Insee a souhaité associer des représentants de la société civile en amont du projet sous la forme d'une concertation approfondie avec différents acteurs.



L'Insee a souhaité associer des représentants de la société civile en amont du projet sous la forme d'une concertation approfondie avec différents acteurs afin de tenir compte des points de vue exprimés dans la construction du répertoire statistique tout comme dans ses usages.

La concertation²⁴ menée en 2022 a reposé sur deux grandes actions : d'une part une **rencontre**²⁵ du Conseil national de l'information statistique (Cnis)²⁶ le 28 janvier 2022, d'autre part la création d'un groupe de concertation,

placé sous l'égide du Cnis, qui a fonctionné de mai à septembre et dont le **rapport**²⁷ est publié sur son site.

²² Cnil : Commission nationale de l'informatique et des libertés.

²³ Voir les références juridiques en fin d'article.

²⁴ Voir l'article de Dupont, Dussart et Guillaumat-Tailliet sur les enjeux éthiques de Résil dans ce même numéro.

²⁵ <https://www.cnis.fr/evenements/appariements-de-donnees-individuelles-entre-richeesse-de-linformation-statistique-et-respect-de-la-vie-privee/>.

²⁶ Le Conseil national de l'information statistique (Cnis) assure la concertation entre les producteurs et les utilisateurs de la statistique publique.

²⁷ <https://www.cnis.fr/wp-content/uploads/2022/11/rapport-version-dfinitive.pdf>.

La rencontre a permis à l'Insee et à divers représentants de la statistique publique de présenter les pratiques, les usages et les techniques d'appariement ainsi que de montrer les apports pour la connaissance et l'action publique (mesurer l'insertion professionnelle des jeunes, étudier le devenir de bénéficiaires de minima sociaux, comprendre les écarts entre deux sources de données, etc.). Elle a permis l'expression d'interrogations, voire de craintes sur ce qu'un dispositif comme Résil pourrait permettre s'il était mal utilisé, mais elle a aussi fait émerger le souhait que l'effort de communication et de transparence mené à l'occasion de cette rencontre se poursuive par une concertation plus approfondie sur le projet.

Le groupe de concertation a permis de rassembler des expertises très variées (protection des libertés fondamentales, protection des données sur les plans juridique et informatique, transformation numérique, éthique, recherche, etc.) pour dresser la liste des interrogations suscitées par le projet Résil et échanger sur les réponses apportées par l'Insee. Ce groupe a conclu que le projet était légitime et conforme au principe de proportionnalité du traitement, à condition de ne pas utiliser certaines des sources initialement envisagées par l'Insee. Il a considéré également qu'il était nécessaire de bénéficier de regards extérieurs (Conseil d'État, Cnil, Cnis, Autorité de la Statistique Publique, Agence Nationale pour la Sécurité des Systèmes d'Information) lors de la construction et de l'utilisation de Résil. Il insiste également sur la nécessaire transparence sur le répertoire et ses usages, permettant un autre regard extérieur : celui des personnes concernées par le traitement.

L'Insee suit cette voie, tant au niveau du contenu du [décret](#)²⁸ que des modalités de construction du système d'information, mais aussi dans la communication sur le répertoire et ses usages. Une [page consacrée à Résil](#)²⁹ a ainsi été mise en place sur le site internet de l'Insee. Le Cnis jouera un rôle important, selon des modalités à construire, quant à la poursuite de la concertation sur Résil et sur les services qu'il rendra, notamment la construction de fichiers enrichis par appariements.

Dans les deux années qui viennent, les développements du système d'information seront finalisés, le répertoire sera initialisé et les traitements statistiques destinés à en assurer la qualité seront réalisés. Les premiers univers de référence (au 1^{er} janvier 2025) seront produits progressivement entre mi-2025 (le provisoire) et mi-2026 (le définitif) et les premiers services d'appariements seront rendus début 2026.

²⁸ <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000048866207>.

²⁹ <https://www.insee.fr/fr/information/7748883>.

► Fondements juridiques

- Loi n° 51-711 du 7 juin 1951 sur l'obligation, la coordination et le secret en matière de statistiques. In : *site de Légifrance*. [en ligne]. Mise à jour le 25 mars 2019. [Consulté le 5 mars 2024]. Disponible à l'adresse : <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000888573>.
- Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés. In : *site de Légifrance*. [en ligne]. Mise à jour le 21 février 2024. [Consulté le 30 avril 2024]. Disponible à l'adresse : <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000886460>.
- Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique. In : *site de Légifrance*. [en ligne]. [Consulté le 5 mars 2024]. Disponible à l'adresse : <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000033202746>.
- Décret n° 2024-12 du 5 janvier 2024 portant création d'un traitement automatisé de données à caractère personnel dénommé « répertoire statistique des individus et des logements » (Résil). In : *site de Légifrance*. [en ligne]. [Consulté le 5 mars 2024]. Disponible à l'adresse : <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000048866207>.
- Arrêté du 5 janvier 2024 pris en application de l'article 2 du décret n° 2024-12 du 5 janvier 2024 portant création d'un traitement automatisé de données à caractère personnel dénommé « répertoire statistique des individus et des logements » (Résil). In : *site de Légifrance*. [en ligne]. [Consulté le 5 mars 2024]. Disponible à l'adresse : <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000048866233>.
- Délibération n° 2023-080 du 20 juillet 2023 portant avis sur un projet de décret en Conseil d'État relatif à la création du traitement automatisé de données à caractère personnel permettant la gestion du répertoire statistique d'individus et de logements, et sur l'arrêté y afférent. In : *site de Légifrance*. [en ligne]. [Consulté le 4 avril 2024]. Disponible à l'adresse : https://www.legifrance.gouv.fr/cnil/id/CNILTEXT000048934586?page=1&pageSize=100&searchField=ALL&searchType=ALL&sortValue=DATE_DECISION_DESC&tab_selection=cnil&timeInterval=01%2F07%2F2023+%3E+31%2F08%2F2023&typePaging=DEFAULT.

► Bibliographie

- BAKKER, Bart F.M., VAN ROOIJEN, Johan, VAN TOOR, Leo, 2014. The System of social statistical datasets of Statistics Netherlands: An integral approach to the production of register-based social statistics. In : *Statistical Journal of the IAOS*. Volume 30, n° 4, pp. 411-424. [en ligne]. [Consulté le 5 mars 2024]. Disponible à l'adresse : <https://www.cbs.nl/-/media/imported/documents/2016/53/system-of-social-statistical-datasets.pdf>.
- 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 5 mars 2024]. Disponible à l'adresse : <https://www.insee.fr/fr/information/7635825?sommaire=7635842>.
- BYCROFT, Christine, EATHERLEY, Clara, PAGE, Mathew et TA'ALA, Shane, 2022. A statistical person register in New Zealand: Progress and challenges. In : *Statistical Journal of the IAOS*. [en ligne]. 21 mars 2022. Volume 38, n° 1, pp. 225-230. [Consulté le 5 mars 2024]. Disponible à l'adresse : <https://content.iospress.com/articles/statistical-journal-of-the-iaos/sji210922>.
- DEVILLE, Jean-Claude et SÄRNDAL, Carl-Erik, 1992, Calibration Estimators in Survey Sampling. In : *Journal of the American Statistical Association*. [en ligne]. Juin 1992. Vol. 87, N° 418, pp. 376-382. [Consulté le 23 mai 2024]. Disponible à l'adresse : <https://www.jstor.org/stable/2290268>.
- DEMOTES-MAINARD Magali, 2019. Elire, un projet ambitieux au service du Répertoire électoral unique. In : *Courrier des statistiques*. [en ligne]. Juin 2019. Insee. N° N2, pp. 58-71. [Consulté le 23 mai 2024]. Disponible à l'adresse : <https://www.insee.fr/fr/information/4168399?sommaire=4168411>.
- DUPONT, Françoise, 2023. Quels types de sources l'Insee utilise-t-il pour construire ses statistiques ? In : *Le blog de l'Insee*. [en ligne]. 16 mai 2023. [Consulté le 5 mars 2024]. Disponible à l'adresse : <https://blog.insee.fr/quels-types-de-sources-l-insee-utilise-t-il/>.
- DUPONT, Françoise, 2023. Les appariements de données de la statistique publique : des analyses enrichies, un cadre juridique protecteur. In : *Le blog de l'Insee*. [en ligne]. 1^{er} septembre 2023. [Consulté le 5 mars 2024]. Disponible à l'adresse : <https://blog.insee.fr/appariements-de-donnees-de-la-statistique-publique/>.
- ESPINASSE, Lionel et ROUX, Valérie, 2022. Le Répertoire national d'identification des personnes physiques (RNIPP) au cœur de la vie administrative française. In : *Courrier des statistiques*. [en ligne]. 29 novembre 2022. Insee. N° N8, pp. 72-92. [Consulté le 5 mars 2024]. Disponible à l'adresse : <https://www.insee.fr/fr/information/6665188?sommaire=6665196>.
- HACHID, Ali et LECLAIR, Marie, 2022. Sirius, le répertoire d'entreprises au service du statisticien. In : *Courrier des statistiques*. [en ligne]. 29 novembre 2022. Insee. N° N8, pp. 115-130. [Consulté le 5 mars 2024]. Disponible à l'adresse : <https://www.insee.fr/fr/information/6665192?sommaire=6665196>.

- 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 5 mars 2024]. Disponible à l'adresse : <https://www.insee.fr/fr/information/5398683?sommaire=5398695>.
- 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 5 mars 2024]. Disponible à l'adresse : <https://www.insee.fr/fr/information/6665186?sommaire=6665196>
- ZHANG, Li-Chun et DUNNE, John, 2017. Trimmed Dual System Estimation. In : *Capture-Recapture Methods for the Social and Medical Sciences*, pp. 239-259. Éditions Chapman and Hall/CRC. ISBN 978-1-49-874531-4.