

CONSEIL D'ORIENTATION DES RETRAITES  
Séance plénière du 19 décembre 2024 à 10h00  
« Le suivi statistique du système de retraite »

<b>Document n° 5</b>
<i>Document de travail, n'engage pas le Conseil</i>

**Le système d'information de la Direction Technique  
de l'Agirc-Arrco en 2024**

*Agirc-Arrco*



*Demandeur : COR*

*Rédacteurs : Stanislas Bourbon - Fadia El Kadiri - Cécile Simon*

## **Le système d'information de la Direction Technique de l'Agirc-Arrco en 2024**

La Direction Technique (DT) de la Fédération Agirc-Arrco a deux missions principales : la connaissance des populations du régime et le pilotage technique de l'Agirc-Arrco. Pour réaliser ses objectifs, la DT s'appuie sur un système d'information statistique qui a considérablement évolué au cours des dix dernières années. Trois facteurs majeurs ont permis sa transformation :

1. la mise en place du Système d'Information Décisionnel de la Retraite Complémentaire (SID-RC),
2. la volonté de disposer d'un modèle de micro-simulation spécifique à l'Agirc-Arrco (MisrAA) qui soit fiable et robuste,
3. l'ambition de positionner l'Agirc-Arrco comme un acteur de référence en matière de données.

L'objet de cette note est de présenter le système d'information statistique de l'Agirc-Arrco, puis le champ et les données disponibles. Enfin, dans une troisième partie, la valorisation des travaux basés sur les données issues du SI de l'Agirc-Arrco et leur mise à disposition en libre accès sont rappelées.

Nous précisons que la description du système d'information statistique est réalisée à la date de la rédaction de la note (automne 2024). Celle-ci ne tient pas compte du plan de transformation qui est prévu par la DSI<sup>1</sup> Agirc-Arrco sur la période 2025-2026 et qui consiste à basculer le système actuel sur une architecture Hadoop.

### 1. Le système d'information statistique de l'Agirc-Arrco

Pour répondre aux enjeux de la DT, il a été nécessaire de refondre le concept de traitement de la statistique Agirc-Arrco en collaboration étroite avec la Direction des Systèmes d'Information (DSI). Cette nouvelle conception « data » a permis à la DT d'acquérir une grande autonomie et une souplesse

---

<sup>1</sup> Direction des Systèmes d'Information

dans ses capacités à capturer, à traiter et à stocker l'information, puis à la rafraîchir à un rythme "suffisant" pour ses besoins. Elle a de ce fait impliqué une adaptation de ses méthodes de collecte, d'exploitation et d'analyse de la donnée. Durant cette période de construction, un des principes fondamentaux de la DT a été de se mettre en capacité d'aller "piocher" l'information utile, de manière systématique, à tous les niveaux, là où elle transite, avec comme objectif majeur de répondre le plus efficacement possible aux besoins du pilotage technique du régime et d'aider les partenaires sociaux dans leurs prises de décision.

Pour réaliser ses travaux, la DT Agirc-Arrco utilise d'une part et majoritairement les données issues du SID-RC dans lequel on retrouve les données de gestion de la retraite complémentaire Agirc-Arrco, et d'autre part, des données dites « exogènes », qu'elles soient issues des flux statistiques inter-régimes ou de l'open data.

### 1.1. Les données issues du système de gestion Agirc-Arrco

L'Agirc-Arrco a lancé il y a environ 10 ans, la construction d'un gigantesque entrepôt de données reprenant l'ensemble des informations nécessaires à la gestion du régime : le SID-RC. L'objectif principal visait à centraliser, puis à mettre à disposition des directions utilisatrices de la Fédération « toutes » les données fonctionnelles de l'Agirc-Arrco, afin d'en permettre le pilotage sous tous ses aspects. Il s'agissait notamment d'améliorer la connaissance de la population des ressortissants du régime, en mettant particulièrement l'accent sur leurs historiques professionnels et familiaux et ainsi permettre aux services, techniques singulièrement, de pouvoir effectuer tous les chiffrages nécessaires au pilotage du régime.

Les impacts de cette transformation ont été multiples : un accès à des tables individuelles, brutes, exhaustives, peu ou pas retravaillées et alimentées au fil des chargements effectués en gestion. Les données demeurent généralement structurées, avec une mise à disposition quasi immédiate pour les services. Le coût de récupération de la donnée a donc été très nettement réduit pour la DT, apportant un gain en termes d'autonomie et d'agilité pour exploiter un matériau qui s'est considérablement enrichi au fil du temps. Une des doctrines de la DT a été de maximiser la récupération d'information et de la rendre utilisable, pour aujourd'hui ou pour plus tard. Car certes, les données brutes déclaratives peuvent s'avérer parfois inutiles pour gérer un dossier retraite, mais ce qui peut être superflu en gestion, ne l'est pas forcément en statistique. Il appartient donc aux statisticiens, dont c'est le métier, de gérer les manques éventuels et la qualité des données en retraitant les informations nécessaires, tout en respectant leur conformité.

En corollaire de cet avènement du SID-RC, de nouveaux processus de traitement de la donnée ont dû être mis en place au sein de la DT amenant à refondre la collaboration entre DT, DSI et MOA Fédérale,

à repenser les méthodes de travail et à faire monter en compétence les équipes : approfondissement de la connaissance des données de gestion, adaptation à l'industrialisation de l'alimentation, optimisation des exploitations statistiques, rigueur dans la gestion de la volumétrie, renforcement des contrôles, mise en œuvre de nouvelles techniques de complétude, souci constant de l'auditabilité des productions...

Les données contenues dans le SID-RC portent sur l'ensemble des domaines fonctionnels de la retraite Agirc-Arrco : adhésion des entreprises, déclarations des salaires, encaissements des cotisations, traitement des comptes de points des actifs en emploi ou au chômage, gestion des processus de liquidations des droits des assurés et des prestations à servir. Toutes ces informations sont déversées dans l'entrepôt de données -le SID-RC-, généralement au fil de l'eau, permettant aux statisticiens de traiter ces données, puis de les mettre à disposition des « clients » du régime Agirc-Arrco.

L'évolution du SID-RC, par construction, résulte des transformations opérées sur le Système d'Information Agirc-Arrco (le système de gestion de la retraite complémentaire). Toute modification effectuée en amont a potentiellement un impact pour le système d'information statistique, qu'elle soit exogène, comme par exemple la mise en place de la DSN à partir de 2017, ou qu'elle relève de l'interne Agirc-Arrco, comme par exemple le déploiement du nouveau Système Central de Liquidation (SCL) et du Référentiel National de Gestion des Droits (RNGD). À chacune de ses évolutions, le sujet doit être instruit collectivement avec la DSI et la Direction du Produit Retraite Agirc-Arrco (MOA Fédérale), pour poser un diagnostic et établir une feuille de route de bascule. Ces opérations doivent permettre d'identifier et de limiter les impacts pour la DT, qui procède alors aux adaptations nécessaires dans ses processus de traitement des données. Surtout, il s'agit d'éviter toute régression dans la production des indicateurs statistiques du régime Agirc-Arrco. Le SID-RC embarque également des données qui servent à la DT pour des travaux d'appui à la gestion tels que les prévisions de volumes d'activité à destination des GPS<sup>2</sup> (à partir de l'ancienne brique PRC : Plateforme de la Retraite Complémentaire et de la brique SCL : Système Central de Liquidation) ou encore à partir des données récemment embarquées dans le SID-RC et stockant les informations sur les interactions avec les affiliés (brique CRM : Customer Relation Management). Le SID-RC est régulièrement enrichi par des nouvelles données avec les évolutions du système d'information de l'Agirc-Arrco.

---

<sup>2</sup> Groupes de Protection Sociale

## 1.2. Les données issues de sources externes

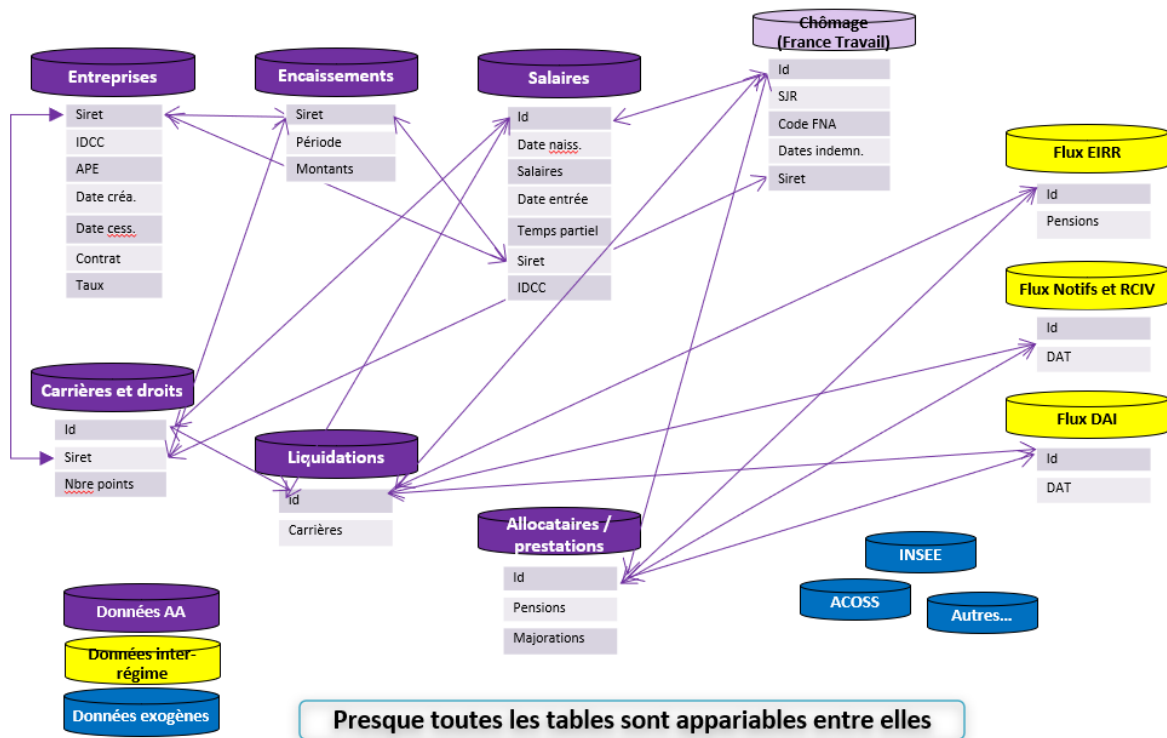
Dès que cela a été possible, la DT de l'Agirc-Arrco a cherché à enrichir son système d'information statistique des données inter-régimes. En décidant de tout mettre en œuvre pour récupérer l'essentiel de ces informations, à commencer par celles issues des flux du droit à l'information (DAI), la DT a répondu à cette doctrine de récupérer de manière systématique toutes les informations qui transitent par le SI Agirc-Arrco, dès lors que ces informations sont susceptibles d'améliorer la connaissance et le pilotage du régime. Ces données s'appuient généralement sur des processus inter-régimes existants, comme celui portant sur les campagnes de droit à l'information au titre duquel un processus de récupération des flux à des fins statistiques, piloté par le GIP Union Retraite, a été mis en place dès 2013. D'autres bases statistiques sont issues de flux de données qui sont alimentés par les régimes de bases (Cnav et MSA), permettant notamment de récupérer les différentes durées d'assurance tous régimes (via les notifications et les relevés individuels de carrière ventilés). Enfin, le dernier flux capté par l'Agirc-Arrco concerne le répertoire EIRR (Échanges Inter-Régimes de Retraites) : créé en 2009, il met à disposition de la DT des données individuelles sur les pensions tous régimes confondus des allocataires Agirc-Arrco depuis 2014.

Globalement, il est prévu que ces traitements inter-régimes prennent fin lorsque les données de l'ensemble des régimes seront disponibles dans le RGCU (Répertoire de Gestion des Carrières Unique). La restitution -à l'Agirc-Arrco- des données massives du RGCU à des fins statistiques est bien prévue, mais sa mise en œuvre n'est pas encore effective. Une expression des besoins Agirc-Arrco a été transmise à l'opérateur RGCU au printemps 2024 dans le cadre du processus officiel couvrant les demandes d'extractions des régimes. Une mise à disposition des données est actuellement envisagée au début de l'année 2025. Un des rôles de la DT est de s'assurer qu'à chaque évolution portant sur les flux inter-régimes, les besoins du système d'information statistique Agirc-Arrco soient pris en compte, afin que ce système ne subisse pas de régression et qu'il soit enrichi, autant que faire se peut, au moyen de ces nouvelles sources de données. À ce titre, l'accès aux données issues du RGCU répliqué<sup>3</sup> devra permettre à la DT de compléter sa connaissance statistique des carrières. Aujourd'hui, l'Agirc-Arrco ne dispose en effet pas des parcours complets de ses assurés dans les autres régimes et donc ne connaît pas les durées d'assurance de tous ses assurés à date. Pour combler ce défaut de couverture, la DT procède à des opérations de complétion. La mise à disposition des données très riches du RGCU devrait permettre au système d'information statistique de l'Agirc-Arrco de faire un nouveau bond en avant en récupérant de façon plus simple et exhaustive le détail de la carrière des assurés Agirc-Arrco dans les autres régimes.

---

<sup>3</sup> Mise à disposition des données massives du RGCU sur le périmètre de l'Agirc-Arrco

**Graphique 1 : le système d'information statistique de la DT Agirc-Arrco, un accès à des données d'une grande richesse via des tables structurées par domaine qui peuvent être reliées entre elles**



## 2. Présentation des bases de données de l'Agirc-Arrco et de leurs utilisations

Dans cette seconde partie, les bases de données disponibles à la DT sont présentées de façon succincte en répondant aux principales questions suivantes : quelles sources ? quels champs ? quelles données ? quelles volumétries ? quelles utilisations ?

### 2.1. Les référentiels nationaux de l'Agirc-Arrco

Il existe quatre référentiels nationaux à l'Agirc-Arrco : le premier porte sur les individus (RNI), le deuxième sur les entreprises en France (RNE), le troisième sur les entreprises adhérentes à l'Agirc-Arrco (RNA) et le dernier sur les droits des assurés à l'Agirc-Arrco (RNGD). Dans cette partie, sont présentés les trois premiers, celui sur les droits Agirc-Arrco sera détaillé ci-après (partie 2.2.).

Ces référentiels nationaux mis à disposition de la DT en 2017 remplacent les différents « référentiels » existant auparavant dans chaque groupe de protection sociale. Ils permettent de centraliser les

données d'identification et d'avoir l'information la plus récente pour chaque individu et pour chaque entreprise. Ces référentiels sont rafraichis quotidiennement.

Le **référentiel « Individus »** porte sur près de 100 millions d'individus ; les personnes ayant quitté le régime définitivement (personnes décédées, allocataires ayant perçu un capital unique...) sont identifiées comme « sorties » mais demeurent dans les bases de données. Les principales données disponibles sont :

- Des identifiants individuels ou des identifiants « régime » (régime général, CCMSA, Monaco...);
- Les données démographiques (date et lieu de naissance, date et lieu de décès...);
- Les données pour contacter l'assuré (notamment l'adresse);
- Le « rôle » de l'individu (actif, retraité de droit direct, retraité de droit dérivé).

Ce référentiel « Individus » est peu utilisé seul dans le cadre des travaux de la DT, mais il est très largement mobilisé en croisement avec les autres bases de données (base allocataires, base droits...) pour connaître les caractéristiques des assurés. Il permet ainsi de répondre à toutes les demandes d'indicateurs par âge, sexe ou encore lieu de résidence par exemple.

Sa seule utilisation directe actuelle porte sur le suivi des décès des assurés Agirc-Arrco. Ces travaux ont été initiés lors de la crise sanitaire afin de répondre aux besoins de chiffrages sur les effets de la surmortalité liée au Covid pour le régime.

Enfin, toutes ces données individuelles sont un input du modèle de micro-simulation de l'Agirc-Arrco (MiSRAA). Elles permettent entre autres, d'établir les tables de mortalité.

Le **référentiel « Entreprises »** compte 5,3 millions d'établissements et 3,8 millions d'entreprises en 2023. Les principales informations présentes dans cette base sont :

- les identifiants entreprises (siret, siren);
- la raison sociale de l'établissement avec le siège social;
- le motif de création de l'entreprise (resp. de l'établissement) et la date associée;
- en cas de fermeture, le motif et la date;
- le code d'activité principale de l'entreprise (ou de l'établissement) et la date d'effet associée;
- la catégorie juridique de l'entreprise et la date d'effet associée;
- les codes de convention collective.

Le **référentiel « Adhésions »** compte également 5,3 millions d'établissements et 3,8 millions d'entreprises en 2023. Les principales informations disponibles (en plus des données contenues dans le RNE) sont :



- le code CPN (Catégorie Professionnelle Normalisée – ancienne norme CRC), l’identifiant CPRF (Catégorie Professionnelle référence) et la catégorie réglementaire permettant de définir si le taux de cotisation s’applique à un cadre, un non-cadre, un apprenti, un ouvrier... ;
- l’assiette de cotisations ;
- les taux de cotisation en vigueur et une information précisant si ce taux est obligatoire ou supplémentaire ;
- les parts patronales et salariales ;
- les dates de début et de fin des taux, du contrat et de la garantie.

Ces deux derniers référentiels croisés notamment avec les données DSN permettent la production d’indicateurs régime importants tels que les effectifs d’entreprises adhérentes à l’Agirc-Arrco, les taux moyens observés T1/T2, la répartition des entreprises selon les taux T1/T2 ou encore de déterminer les entreprises avec des adhésions particulières (différentes du cas standard).

## 2.2. Domaine « Carrières et Droits » des assurés actifs Agirc-Arrco

Les données sur les carrières dont dispose la DT sont multiples :

- Celles relatives aux personnes au chômage indemnisé ouvrant des droits à la retraite complémentaires Agirc-Arrco ;
- Celles portant sur les droits acquis à l’Agirc-Arrco via le Référentiel National de Gestion des Droits (RNGD) ;
- Et enfin celles correspondant aux droits acquis dans les autres régimes (trimestres et points) via les flux du Droit à l’Information (DAI) ;

Les **données relatives aux droits « chômage » des assurés Agirc-Arrco** sont présentes dans le fichier « Travaux de Fin d’Année » (TFA). Ce fichier est actuellement fourni annuellement par France Travail.

Chaque année, c’est à partir de ce fichier que la DT établit les montants à facturer aux organismes tiers (État via le FSV et Unédic) qui financent les points des assurés Agirc-Arrco acquis au titre des périodes de chômage.

Le champ de cette base de données porte sur l’ensemble des chômeurs indemnisés affiliés à l’Agirc-Arrco (hors convention de gestion), elle est exhaustive depuis l’exercice 2012. Pour l’année 2022, la base de données porte sur 5,3 millions d’individus, soit environ 10,5 millions de lignes (un individu peut percevoir plusieurs allocations la même année).

Les principales informations disponibles dans cette base sont :

- des données individuelles (date de naissance, sexe, code postal de résidence) ;
- des données relatives au dernier contrat de travail (date de fin de contrat, catégorie professionnelle, siret et raison sociale de l'entreprise...) ;
- des données sur l'indemnisation chômage (type d'allocation perçue, dates de début et de fin d'indemnisation, nombre de jours indemnisés, montant de l'allocation journalière...).

Cette base de données est également utile dans le cadre des autres travaux de la DT, que ce soit pour des demandes récurrentes ou ponctuelles. Elle était par exemple utilisée pour identifier les bénéficiaires de l'ASS, qui étaient exonérés des coefficients de solidarité Agirc-Arrco au moment de leur départ en retraite. Elle a également permis d'identifier les personnes au chômage avant leur retraite dans le cadre des travaux sur « la situation des assurés Agirc-Arrco avant la retraite »<sup>4</sup>. Enfin, elle permet de répondre aux demandes des partenaires sociaux, comme par exemple le chiffrage des impacts des réformes chômage sur les droits acquis au titre de la retraite complémentaire.

Avec le passage au RGCU, cette base de données ne sera plus alimentée à l'Agirc-Arrco par le fichier TFA de France Travail, les données nécessaires au calcul des factures aux tiers seront reprises directement dans le répertoire inter-régime. Le dernier exercice couvert par la base TFA de Pôle emploi devrait être l'exercice 2026<sup>5</sup> (transmis début 2027).

Pour les études statistiques, le **Référentiel National de Gestion des Droits (RNGD)** est la source de données sur les carrières et les droits des assurés Agirc-Arrco<sup>6</sup>. Les travaux ont été initiés à compter de 2017 et le référentiel est pleinement opérationnel à la DT depuis 2020. Le RNGD est alimenté par la DSN pour les droits au titre de l'emploi et de la maladie et par le fichier TFA (cf. paragraphe précédent) pour les droits au titre du chômage. Auparavant, les données de carrière étaient gérées dans les systèmes d'information propres des groupes de protection sociale puis agrégés dans une base centrale au niveau Agirc-Arrco : la base fédérale droit.

À partir des 46 tables stockées dans l'entrepôt de données du SID RC, la DT construit 2 datamarts à partir desquels la majorité des études peuvent être produites. Pour des demandes plus spécifiques, des extractions peuvent être réalisées sur les données sources.

---

<sup>4</sup> « La situation des assurés Agirc-Arrco avant la retraite en 2020 », Cahiers Etudes et Statistiques n°6  
<https://www.calameo.com/read/002711729b70af55b53e7>

<sup>5</sup> Cet exercice sera fait en « double commande » afin de s'assurer que le passage au RGCU s'est bien opéré et que le calcul de la facture est identique à partir des 2 sources.

<sup>6</sup> En effet, à fin 2024, le service statistique de l'Agirc-Arrco ne dispose pas des données du RGCU. En revanche, s'agissant des activités de gestion, le RNGD est le référentiel des droits, alors que le référentiel carrière est le RGCU.

Ces deux datamarts ont trois dimensions : individu x année x régime. Même si le régime Agirc-Arrco a été unifié à compter de 2019, en historique, toutes les données sont disponibles pour l'ex-Arrco d'une part et le cas échéant, dans une seconde « ligne » pour les cadres cotisant à l'ex-Agirc. Dans ces bases, on dénombre 44,9 millions d'individus ayant acquis des droits à l'Agirc-Arrco au cours de leur carrière et n'étant pas encore retraités de droit direct fin 2022 dont 26,8 millions d'individus ayant des droits inscrits au titre de l'année 2022. Cela se traduit par environ 1,1 milliard de lignes dans chaque datamart.

Le premier datamart porte sur les « droits ». Les principales données disponibles par exercice et par régime sont :

- le nombre de points acquis
  - par tranche (T1 et T2 pour l'Arcco, TB et TC pour l'Agirc jusqu'à fin 2018 et T1 et T2 pour l'Agirc-Arrco à compter de 2019),
  - par type de points (cotisés, au titre du chômage, de la maladie...);
- le nombre de jours par type de droits ;
- les dates de début et de fin par type de droits.

Le second datamart porte sur les « salaires », les principales informations sont les montants des assiettes salariales par exercice et par régime.

À partir de cette base, des indicateurs récurrents sont produits pour le compte du régime : effectifs annuels par sexe, points moyens acquis par sexe et générations aux âges quinquennaux (20 ans, 25 ans...) par type de droits (totaux ou emploi). Ces indicateurs alimentent les différentes publications de l'Agirc-Arrco. De même, les données issues du RNGD permettent d'établir les compensations financières entre l'Agirc-Arrco et d'autres régimes de retraite complémentaire ou de grandes entreprises (Ircantec, CPRPSNCF, La Poste).

Des demandes ponctuelles sont également réalisées à partir de ces données, que ce soit pour d'autres directions de la Fédération, les partenaires sociaux ou encore des organismes tels que la Cour des Comptes ou les assemblées parlementaires.

Enfin, les données issues du RNGD permettent d'alimenter l'Echantillon Inter-régimes de Cotisants (EIC) pour le champ Agirc-Arrco tous les 4 ans.

Le RGCU est alimenté par les données du RNGD pour la partie droits Agirc-Arrco.

Les **données portant sur les droits des assurés acquis dans les autres régimes de retraite** sont collectées dans le cadre des campagnes de droits à l'Information (DAI) via les RIS<sup>7</sup> et les EIG<sup>8</sup>. Un processus de constitution et de mise à disposition de fichiers statistiques a été créé par le GIP Union Retraite en 2013 pour les régimes. Depuis le début des campagnes, il y a eu neuf collectes de données à des fins statistiques (2013-2021). Les fichiers de données en format CSV sont mis à disposition de la DT Agirc-Arrco une fois par an. Ces données concernent les cotisants de l'Agirc-Arrco âgés de 35 ans et plus, par tranche d'âge quinquennale. Les fichiers (3 par campagne : synthèse, trimestres, points) contiennent les données suivantes :

- **Fichier Synthèse** (niveau individu-régime) : données sociodémographiques, durée d'assurances en trimestres (DAT), total des trimestres et des points.
- **Fichier Trimestres** (niveau individu-année-régime) : trimestres validés, cotisés, assimilés (chômage, maladie, militaire, invalidité et total des assimilés), liquidés partiellement et trimestres AVPF.
- **Fichier Points** (niveau individu-année-régime) : points totaux, cotisés et « dits » gratuits (chômage, maladie...)

À partir de ces fichiers inter-régimes, des données pour chaque individu et pour chaque année d'activité sont consolidées avec le nombre de trimestres acquis dans chaque régime par type de trimestres. Ces données servent ainsi à établir des indicateurs sur les trimestres validés, cotisés ou assimilés par sexe, âges quinquennaux et par régime, entre 35 et 60 ans. Elles permettent de mieux connaître les carrières inter-régimes et ainsi de comprendre la situation des assurés Agirc-Arrco avant leur départ à la retraite. Ces informations sont mobilisées en input du modèle de micro-simulation Agirc-Arrco (MiSRAA) permettant de projeter les futures durées d'assurance, points et trimestres obtenus dans les autres régimes (de base singulièrement).

À partir de 2025, ces éléments devraient être transmis à partir du RGCU. Une phase pilote se déroule en 2024 avec l'opérateur, afin de tester le processus aboutissant à la fourniture effective des données aux régimes, et donc à l'Agirc-Arrco. Cela comprend, d'une part, un volet sur la contractualisation avec les régimes, avec pour objectif la clarification du processus, et d'autre part, un volet sur le déploiement afin de tester l'acheminement sécurisé des données.

La mise à disposition d'un système d'information répliqué alimenté par ce répertoire inter-régime devrait permettre de fournir à la DT des informations beaucoup plus complètes et via un processus

---

<sup>7</sup> Relevé de Situation Individuelle

<sup>8</sup> Estimation Indicative Globale

simplifié sur les durées d'assurance et les droits acquis par l'ensemble des assurés Agirc-Arrco dans les autres régimes (tout au long de leur carrière).

### 2.3. Domaine « Allocataires »

En 2019, la DT a procédé à une transformation majeure de l'alimentation de son système d'information statistique portant sur les données Allocataires. Elle dispose dorénavant des données brutes directement issues du système de gestion des allocations du régime Agirc-Arrco (Brique Allure) et mis à disposition à la DT dans des entrepôts de données par la DSI. Ces entrepôts de données sont alimentés et mis à jour avec une fréquence quotidienne. A partir de juin 2024 et avec une montée en charge progressive, un nouvel écosystème applicatif (Celia : CEntraLlisation Allocataires) se substituera à la brique Allure. Cet écosystème a pour objectif de construire et déployer le nouveau système de paiement des allocataires et répondre à leurs besoins. La DT a déjà commencé à travailler sur les données issues de Celia afin d'assurer une continuité des processus de production sur les allocataires.

À partir de ces données brutes, la DT crée, selon un processus bien structuré, une base individuelle allocataires par année de liquidation avec une mise à jour mensuelle. Cette base est la principale source d'informations de la DT sur les retraités de l'Agirc-Arrco. À partir de l'exercice 2017, il existe une rupture de série avec l'ancien système<sup>9</sup>. Pour les années 2017 et 2018, des datamarts ont été réalisés en prenant en compte de façon distincte les ex-régimes Arrco et Agirc et à partir de 2019, l'intégration du régime fusionné a été effective sur tout le système d'information statistique allocataires de la DT permettant de reconstituer un historique Agirc-Arrco consolidé depuis 2009.

Des informations relatives aux dossiers et pensions des allocataires, ainsi que des données relatives aux individus et à leur situation règlementaire (taux CSG, CRDS, CASA...) sont récupérées, recalculées ou formatées dans cette base.

La liste suivante contient les principaux éléments disponibles dans cette base individuelle :

- **Type de retraite** : droits directs/droits de réversion ;
- **Informations sur les dossiers** : institution gestionnaire de liquidation, date d'effet de la liquidation ;

---

<sup>9</sup> Avant 2019, l'ensemble des IRC (Institutions de Retraite Complémentaire) étaient amenées à transmettre les éléments statistiques via une brique statistique appelée « Oscar ». Pour l'exercice 2017, il existe une base allocataires DT Agirc-Arrco constituée à partir de l'ancien système issu d'Oscar et une autre à partir du nouveau système.

- **Informations sur les allocataires** : NIR anonymisé, date de naissance, sexe, date de décès, lieu de résidence (ces informations individuelles sont récupérées du référentiel RNI, cf. partie 2.1) ;
- **Nature de la carrière (Motif de départ à la retraite)** : carrière longue, invalidité, handicap, retraite progressive ;
- **Nombre de points** : points servis, points inscrits, points majorations familiales... ;
- **Taux utilisés pour le calcul des droits** : taux d'abattement, taux retraite progressive, coefficients de solidarité, taux de réversion... ;
- **Situation réglementaire** : régime général ou local, taux de CSG, taux CRDS, CASA...

Au titre de l'année 2022, la base individuelle allocataires Agirc et Arrco compte 20,7 millions de dossiers et 13,6 millions d'allocataires distincts en droits directs y compris les versements forfaitaires uniques servis dans l'année et les réversions.

Les statistiques sur les données allocataires portent essentiellement sur :

- le stock des allocataires vivants au 31/12/N
- le flux des allocataires y compris les versements uniques (VU)
- les décédés de l'année N (au 31/12/N)

Elles permettent de :

1. Produire des indicateurs annuels et infra-annuels portant sur des suivis récurrents en termes d'effectifs, de nombre de points ou des montants de pensions par type de droits (droits directs, droits de réversion) et nature de droits (versements forfaitaires uniques, majorations familiales...) ;
2. Alimenter le modèle de micro-simulation Agirc-Arrco (MiSRAA) ;
3. Répondre à des demandes statistiques annuelles ou en infra-annuelles pour les directions fédérales, pour les instances Agirc-Arrco (Commission Technique et Financière, partenaires sociaux...) ou encore dans le cadre de demandes externes adressées par la DSS, la Cour des Comptes, le COR, la Drees, l'Apec....

Certaines demandes externes sont récurrentes :

- Enquête annuelle auprès des caisses de retraites (EACR) destinée à la Drees ;
- Echantillons Inter-régimes de Retraités (EIR) tous les 4 ans pour la Drees ;
- Statistiques annuelles sur les allocataires résidents à l'étranger pour le Cleiss.

Les demandes ponctuelles portent sur des productions statistiques et des simulations très diverses et peuvent être très précises : suivi des majorations familiales et pour ancienneté, étude d'impact sur les dispositifs spécifiques concernant les allocataires (Coefficients de solidarité, Retraite progressive,

Cumul emploi retraite...). Toutes ces informations peuvent être déclinées selon plusieurs dimensions, par génération, par âge, par année de liquidation, par découpage géographique...

En plus des bases individuelles issues et constituées via les sources internes de l'Agirc-Arrco, la DT dispose d'autres données complémentaires sur les allocataires provenant de sources externes. Il s'agit de :

- L'**EIRR** (Échanges Inter-Régimes de Retraites), qui porte sur les prestations reçues par les assurés de 55 ans et plus, pour tous les types de droits perçus (droit direct et droit de réversion) et tous les régimes français et étrangers auprès desquels un assuré Agirc-Arrco perçoit ou peut percevoir une pension. Ce fichier est reçu une fois par trimestre. La jointure avec la base allocataire et cette source de données, permet à la DT de reconnaître les allocataires de l'Agirc-Arrco multi-pensionnés ainsi que de calculer les pensions moyennes tous régimes et par conséquent de quantifier la part de la pension moyenne versée par l'Agirc-Arrco.

Le **RCIV** (Relevé de Compte Individuel Ventilé) et les **notifications Cnav et MSA**, qui permettent de connaître, d'une part les carrières inter-régimes des allocataires Agirc-Arrco, et d'autre part, le nombre de trimestres acquis dans les régimes de bases et par conséquent la durée d'assurance totale (DAT tous régimes confondus) des allocataires. Avec le projet de la mise en place du RGCU et la bascule de certaines données dans le cadre de ce projet, 2021 était le dernier exercice reçu avec les données du RCIV. En attente de l'accès aux données du RGCU, la DT récupère les DAT de la brique SCL (cf. supra).

#### 2.4. Domaine « Salaires-Cotisans-Entreprises »

Depuis plus de 15 ans<sup>10</sup>, l'Agirc-Arrco constitue annuellement une base statistique permettant d'avoir une bonne connaissance de la population salariée du secteur privé. La mise en place de cette démarche répondait à une demande ancienne des partenaires sociaux : permettre d'obtenir une mesure homogène et précise du salaire moyen des ressortissants du régime. Les données de la base salaires sont historiquement constituées à partir de l'exploitation des déclarations des entreprises adhérentes à l'Agirc-Arrco (DADS jusqu'en 2018 et DSN à compter de 2017). La première base salaires complète sur le périmètre Agirc-Arrco remonte à l'exercice 2005, avec un taux de couverture au début incomplet et qui s'est amélioré jusqu'à devenir aujourd'hui quasi exhaustif. La population prise en compte correspond à l'ensemble des salariés cotisants du régime (de l'industrie, du commerce, des services et de l'agriculture), quel que soit le type de contrat ou la quotité de travail. Ce salaire brut comprend tous

---

<sup>10</sup> Et même à partir de 1994 pour le périmètre de l'ancien régime Agirc.

les montants assujettis aux cotisations Agirc-Arrco, les assiettes situées au-delà de 8 plafonds étant également remontées dans le système d'information.

Depuis le 1<sup>er</sup> janvier 2017, les entreprises sont amenées à effectuer leurs déclarations de salaires au rythme mensuel via la DSN (Déclaration Sociale Nominative).

L'obligation de la généralisation de la DSN au 1<sup>er</sup> janvier 2017 a pu toutefois faire l'objet d'un aménagement pour certains employeurs. C'est pourquoi l'exercice 2017 (puis 2018 dans une moindre mesure) a vu la coexistence de 2 normes : NEODES (norme de la DSN) et DADS et a nécessité un important travail de rapprochement des déclarations effectuées entre les deux normes. 2019 est le premier exercice complet exprimé sous la norme DSN. En 2022, environ 96 % du périmètre Agirc-Arrco<sup>11</sup> est déclaré en DSN, 2 % via une norme spécifique pour les salariés de particuliers-employeurs qui ont vocation en cible à basculer en DSN et le reste (2 %) via d'autres normes déclaratives et correspondant à des populations spécifiques (Monaco, Com, Guso...).

A partir de ces déclarations, la DT constitue 2 bases Salaires-Cotisants-Entreprises individuelles utilisées pour des besoins différents :

- Une première base avec les informations à la maille mensuelle. En 2022, cette base individuelle concerne 36,3 millions de postes (cotisants-établissements) avec 25,5 millions de cotisants dans 2 millions d'entreprises,
- Une 2<sup>ème</sup> base plus agrégée à la maille annuelle et enrichie avec des données sur les entreprises issues de RNE et RNA (taux de cotisations, NAF...) et des informations sur les individus, issues de RNI (sexe, code postal, pays, ...). Cette base intègre des recalculs sur les tranches d'assiettes nécessitant d'avoir une connaissance complète sur chaque exercice.

Plusieurs informations sont disponibles dans ces bases (des informations issues directement de la source et d'autres calculées) sur :

- **L'individu** : le sexe, la date de naissance, l'âge.
- **L'établissement** qui emploie le salarié : le siren, le siret.
- **Le salaire** : les rémunérations effectivement perçues au cours de l'année, les tranches T1 et T2, les indemnités du chômage partiel (indemnités employeurs), les cotisations par tranches.
- **L'activité salariée** exercée par l'individu : le statut (cadre ou non-cadre), la CSP, le type de contrat de travail, le temps de travail (temps plein/temps partiel).
- **La période travaillée** : l'exercice, les dates de début et de fin d'activité, le motif de début et de fin d'activité, la durée de période maladie.

---

<sup>11</sup> Calculé via la masse salariale



Au-delà du calcul du salaire moyen Agirc-Arrco qui se décline en différentes notions (salaire équivalent temps plein, SMPT...), ces bases sont également utilisées pour des demandes récurrentes et ponctuelles, internes et externes à l'Agirc-Arrco. En effet, la DT est sollicitée sur l'exploitation de ses bases pour des demandes très différentes effectuées :

- par les instances de l'Agirc-Arrco dans le cadre des commissions techniques et financière, des commissions paritaires et des négociations paritaires ;
- Par l'unité en charge des projections à la DT notamment en input du modèle MiSRAA,
- Par les différentes directions fédérales : Direction Générale, Direction Financière, Secrétariat général, Direction Produit Retraite... ;
- Par les organismes extérieurs : pour des demandes sur les salariés du secteur privé concernant les effectifs de cotisants, les masses salariales, les montants de cotisations ou des statistiques pour des populations spécifiques (cadres cotisants à l'Apec, résidents dans les outre-mer, statistiques sur des entreprises par convention collective ou encore des chiffrages pour les compensations financières avec d'autres caisses de retraite ...);
- Par les organismes de contrôle (Cour des Comptes, Commissaires aux comptes, ...).

Mais surtout, ces données permettent à la DT de produire de nombreux indicateurs statistiques avec des historiques importants sur les cotisants et les entreprises : effectifs, salaires, ventilés par taille d'entreprise, par secteurs d'activité, par région ou encore par taux de cotisations.

La DSI met également à disposition de la DT des données portant sur les encaissements du régime (cotisations, majorations...). Ces données sont stockées dans la brique "Gestion du Recouvrement des Cotisations Collectives" (Grecco). Elles contiennent notamment des informations sur les calendriers (dates de paiements et dates d'afférence du paiement), les documents d'appels, le montant des créances et les montants des paiements réellement perçus. Cette source de données ne couvre pas l'ensemble du champ des encaissements de l'Agirc-Arrco. En effet, il n'y a pas de remontées pour les entreprises couvertes par l'Alliance Professionnelle<sup>12</sup> (soit un peu plus de 15% du champ Agirc-Arrco). Les données sont actualisées quotidiennement et permettent notamment de faire des suivis des encaissements et des créances des cotisations. Ces données sont également utilisées pour le calcul des compensations financières avec les caisses des autres régimes.

A compter de 2026, une nouvelle brique sur les encaissements devrait remplacer Grecco et être disponible pour la DT. Contrairement à ce qui est le cas actuellement, elle portera sur l'ensemble du champ Agirc-Arrco.

---

<sup>12</sup> Ce périmètre couvre les IRC suivantes : Pro-BTP, IRP Auto, Audiens, Agrica, B2V, Lourmel

### 3. Diffusion des données de l'Agirc-Arrco

Depuis plusieurs années, de nombreux organismes se sont engagés dans une démarche globale d'ouverture de leurs données. L'Agirc-Arrco participe également à ce mouvement en alimentant la rubrique « Nous Connaitre / Nos études et publications »<sup>13</sup> du site internet de l'Agirc-Arrco avec des études, un recueil annuel de données statistiques et un module de datavisualisation. Les données disponibles en libre accès sont des séries agrégées portant sur le régime.

L'Agirc-Arrco peut également mettre à disposition certaines de ses données, de façon récurrente ou ponctuelle, à des organismes publics (centres de recherche, services statistiques ministériels, etc.), dans des conditions encadrées via la signature de conventions de partenariat. A ce titre, la DT fournit des données tous les 4 ans à la Drees dans le cadre de la constitution des Echantillons Inter-régimes de Cotisants et de Retraités (EIC et EIR). De même, tous les ans, l'Agirc-Arrco participe à l'Enquête annuelle auprès des caisses de retraites (EACR), également coordonnée par la Drees.

Enfin, la DT a récemment engagé auprès de l'Autorité de la statistique publique une démarche de labellisation d'une partie des séries mises à disposition sur le site internet de l'Agirc-Arrco.

#### 3.1. Données

L'Agirc-Arrco diffuse tous les ans, au printemps, un recueil de données « **Le Chiffre Agirc-Arrco** ». Ce document au format pdf est téléchargeable sur le site internet<sup>14</sup>. Il regroupe l'ensemble des statistiques du régime. Les données mises en ligne sont disponibles à compter de l'exercice 2001. Selon les sujets, un historique plus profond peut-être demandé à la DT par e-mail<sup>15</sup>. Cette publication se décompose en sept thèmes :

- les paramètres (taux de cotisation, valeurs d'achat et de service du point, rendement) ;
- les cotisants et salaires (effectifs de cotisants, répartition par âge, par lieu de travail, salaires ...);
- les entreprises et assiettes salariales (effectifs d'entreprises cotisantes, répartition par secteur d'activité, par région, salaire moyen par tête...);
- les droits acquis à l'Agirc-Arrco (effectifs assurés, nouveaux entrants, points acquis par âge...);

---

<sup>13</sup> <https://www.agirc-arrco.fr/nous-connaitre/nos-etudes-et-publications/>

<sup>14</sup> <https://www.agirc-arrco.fr/nous-connaitre/nos-etudes-et-publications/analyses-statistiques/chiffre-agirc-arrco/>

<sup>15</sup> Mail de contact : [Etudes-statistiques\\_Agirc-Arrco@agirc-arrco.fr](mailto:Etudes-statistiques_Agirc-Arrco@agirc-arrco.fr)

- les retraités et pensions (effectifs de retraités, par type de droit, âge moyen de départ en retraite, montant annuel moyen des pensions, bénéficiaires d'une majoration...);
- les populations du régime (répartition par âge des différentes populations d'assurés, rapport démographique, nombre de décès...);
- les factures au tiers au titre du chômage indemnisé (Unédic, Etat).

Les principales statistiques du recueil annuel sont également mises en avant dans la rubrique « **Données chiffrées** » du site internet<sup>16</sup>. La DT propose des analyses courtes avec les principales caractéristiques et quelques illustrations sur cinq thèmes différents (retraités, nouveaux retraités, cotisants et salaires, droits, entreprises). Les données peuvent être téléchargées au format excel. Ces rubriques sont actualisées au minimum deux fois par an : début avril suite à la diffusion du Chiffre Agirc-Arrco, et éventuellement à l'automne, avec des données définitives.

Figure 2 : Exemple pour la sous-rubrique « Cotisants et salaires »



Enfin, une grande majorité des données du Chiffre Agirc-Arrco sont reprises dans les deux modules de **data visualisation**<sup>17</sup>. Le premier module porte sur les actifs et le deuxième sur les retraités. Un

<sup>16</sup> <https://www.agirc-arrco.fr/nous-connaître/nos-etudes-et-publications/donnees-chiffrees/>

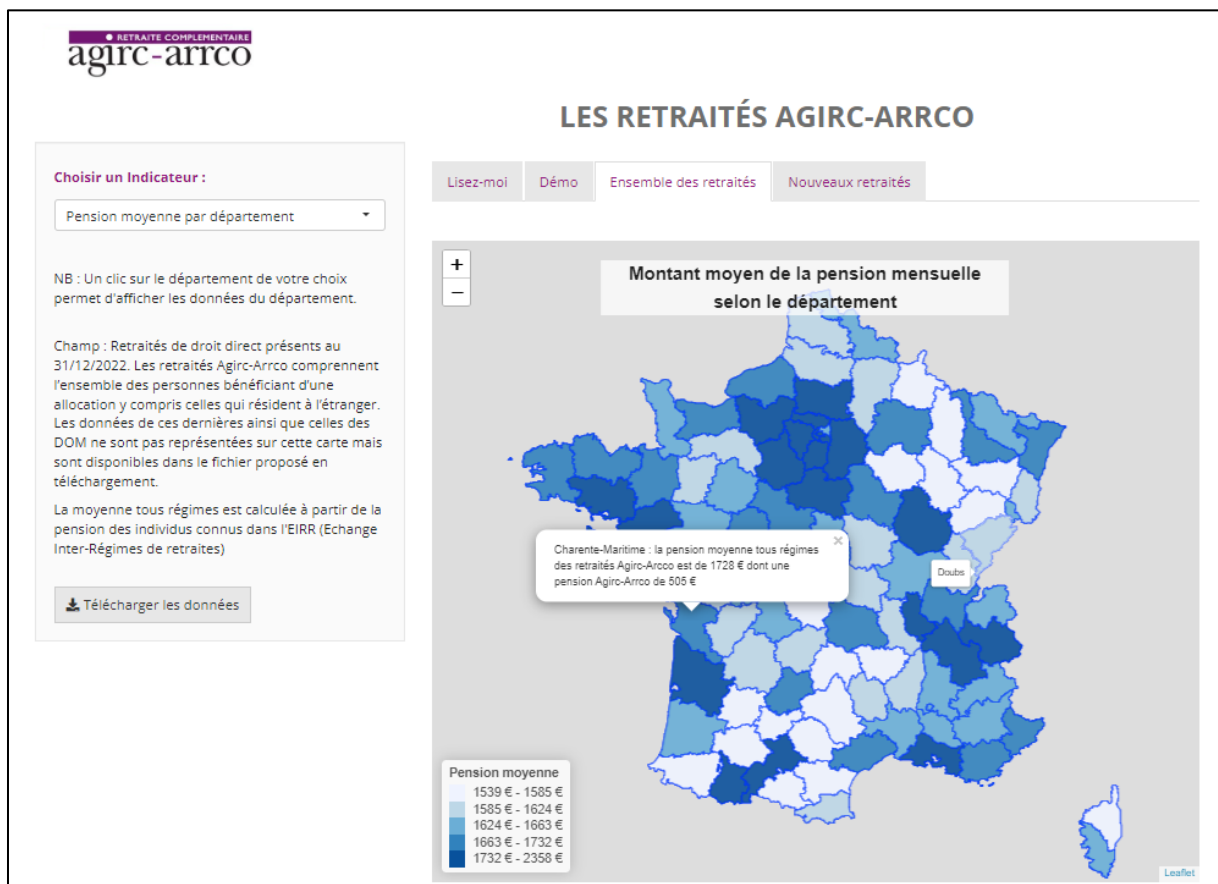
<sup>17</sup> <https://www.agirc-arrco.fr/nous-connaître/nos-etudes-et-publications/data-visualisation/>

troisième module sur la démographie Agirc-Arrco est en cours de développement et devrait être mis en ligne fin 2024.

L'internaute peut sélectionner les indicateurs qu'il souhaite visualiser et définir un périmètre d'analyse, avant de se voir proposer le graphique correspondant à ses filtres. Pour les différents indicateurs (effectifs, salaires moyen et médian, pensions de retraite...), il est possible de disposer des chiffres selon le sexe, la génération ou encore l'exercice. Toutes les données disponibles à l'écran sont téléchargeables au format excel.

La figure ci-dessous propose un exemple d'une page du module data visualisation consacrée aux « retraités ».

Figure 3 : Exemple pour le module de data visualisation



### 3.2. Etudes

Deux formats de publications sont également disponibles en téléchargement.

Les « **Cahiers Etudes et Statistiques** » sont des études de 8 pages réalisées à partir de données du SI statistique de l'Agirc-Arrco sur différents thèmes en lien avec l'actualité ou encore des présentations

des populations du régime. Ils sont téléchargeables sur le site internet<sup>18</sup>. A titre d'exemple, voici quelques sujets abordés : les salaires des cotisants à l'Agirc-Arrco, les caractéristiques des assurés Agirc-Arrco, le financement des droits à retraite complémentaire au titre du chômage, la situation des assurés Agirc-Arrco avant la retraite.

À compter de 2022, un nouveau format d'étude a été mis en ligne : les « **Focus thématiques** ». Ce sont des publications ponctuelles plus courtes (3/4 pages) sur des sujets d'actualité, elles sont également diffusées sur le site internet<sup>19</sup>. Les premiers numéros ont porté sur : les inégalités salariales femmes/hommes, la surmortalité des assurés lors de la crise sanitaire 2020-2021, les profils des nouveaux cotisants Agirc-Arrco, la mobilité en entreprise dans le secteur privé entre 2019 et 2022 ou encore les dispositifs permettant de cumuler emploi et retraite.

---

<sup>18</sup> <https://www.agirc-arrco.fr/nous-connaître/nos-etudes-et-publications/analyses-statistiques/cahiers-etudes-et-statistiques/>

<sup>19</sup><https://www.agirc-arrco.fr/nous-connaître/nos-etudes-et-publications/analyses-statistiques/focus-thematiques/>