

# Une norme d'échange pour alimenter des référentiels et en assurer la qualité



Bertrand Dubrulle\*, Olivier Rosec\*\*, Christian Sureau\*\*\*

*Améliorer la qualité du service pour les assurés de la sécurité sociale, c'est-à-dire disposer d'une information fiable, faciliter les interactions ou servir les prestations à bon droit dans les meilleurs délais, a nécessité la création de grands référentiels et un échange accru de flux de données. Pouvoir garantir la qualité de ces flux et alimenter correctement les référentiels est essentiel. Ainsi, la Cnav a défini et appliqué les principes des "normes d'échange" pour à la fois fournir une description claire et détaillée de l'échange de données et contrôler automatiquement les flux, afin de garantir la qualité de l'alimentation des référentiels. La norme représente un apport essentiel pour le statisticien qui maîtrise ainsi les objets grâce à la documentation structurée et peut alors utiliser des données de meilleure qualité. Les principes d'une norme d'échange et sa structure sont présentés. Un outillage lui est associé pour disposer d'un système générique, cohérent et réactif dans un contexte de forte évolutivité des réglementations. Ces principes et le logiciel associé ont permis de créer un outil de conversion de flux, d'une structure à une autre, en bénéficiant des avantages susmentionnés. La normalisation des échanges ainsi outillée est aujourd'hui effective dans deux cas : la Déclaration sociale nominative et l'alimentation du Répertoire de gestion de carrières unique (RGCU). Avec un recul d'une douzaine d'années la question se pose d'un déploiement à plus grande échelle et de son positionnement dans la stratégie de management de la donnée.*

 *Improving the quality of service for those insured by social security, i.e. having access to reliable information, facilitating interactions or providing benefits as quickly as possible, has required the creation of major benchmarks and increased data exchange. It is essential to ensure the quality of these data flows and correctly feed the repositories. Thus, the Cnav has defined and applied the principles of "exchange standards" to both provide a clear and detailed description of the exchange and automatically control the flows in order to guarantee the quality of the input to the repositories. The exchange standard represents a main information for the statistician who has a better control over data thanks to the structured documentation. The principles of an exchange standard and its structure are presented. Cnav has developed tools to have a generic, coherent and reactive system in a context of frequent changes of regulations. These principles and this tool have made it possible to create another generic tool for converting a flow, from one structure to another, while benefiting from the advantages of the previous one. This standardization of exchanges is now effective in two cases: the Nominative Social Declaration and the feed of the Répertoire de Gestion des Carrières Unique (RGCU). With a decade of hindsight, the question arises of a larger-scale deployment and its positioning in the data management strategy.*

\* Adjoint au directeur de la Gestion de la Donnée (DGD), DSI Caisse nationale de l'assurance vieillesse (Cnav), [bertrand.dubrulle@cnav.fr](mailto:bertrand.dubrulle@cnav.fr)

\*\* Responsable du pôle Expertise Technique Saturne et EBX, DGD, DSI, Cnav, [olivier.rosec@cnav.fr](mailto:olivier.rosec@cnav.fr)

\*\*\* Directeur de la Gestion de la Donnée, DSI Cnav, [christian.sureau@cnav.fr](mailto:christian.sureau@cnav.fr)

Le domaine de la protection sociale se caractérise par des systèmes d'information complexes nécessitant de fréquents échanges d'information. Dès février 2016, les nombreux échanges de données volumineuses entre les régimes et organismes de la protection sociale étaient cités dans le rapport IGAS-IGF<sup>1</sup> (Gratieux et Le Gall, 2016). Par exemple, en 2015, la Caisse nationale



**1881 échanges  
avec 381 partenaires,  
correspondant à un total  
de 1,152 million  
de fichiers reçus ou émis.**



d'assurance vieillesse (Cnav) a comptabilisé, sur les 11 premiers mois, 1881 échanges avec 381 partenaires, correspondant à un total de 1,152 million de fichiers reçus ou émis. Assurer la qualité de ces échanges et la cohérence des informations devenait alors nécessaire pour alimenter correctement les référentiels.

Pour la branche retraite, plus de 20 millions d'actifs et 15 millions de retraités sont gérés, avec des flux de 50 millions d'éléments de carrière en provenance des entreprises, 30 millions en provenance de Pôle emploi, 10 millions en provenance de la branche maladie mais aussi plusieurs dizaines de millions en provenance ou à destination des autres organismes de retraite.

Toutes ces informations sont utilisées pour calculer les droits et servir aux assurés des prestations exactes. Pour cela, des informations complètes et de qualité sont essentielles afin de disposer d'éléments certains pour exécuter des calculs corrects. Pour stocker et centraliser les données et en garantir la qualité, la création de référentiels s'est imposée.

Compte tenu du nombre d'organismes, du nombre d'échanges et du volume de données, définir les principes de normalisation des échanges et les appliquer est essentiel pour garantir la qualité d'alimentation des référentiels, en extraire toute la plus-value et offrir des services à valeur ajoutée à toute la sphère sociale.

Les aspects juridiques, d'infrastructure, la notion de norme d'échanges et ses principes sont présentés. Les différents composants de la structure en blocs de la norme sont décrits ainsi que les dimensions fonctionnelles et techniques de ces blocs et des données qui les composent (les « rubriques »). On explicite ensuite les différents types de contrôles automatiques sur les flux. Un ensemble d'outils, nommé Saturne, a été développé par la Cnav pour générer automatiquement ces contrôles et d'autres livrables. On explique cet outillage et son rôle majeur pour reproduire la démarche et disposer de la réactivité dans la prise en compte des évolutions. Pour conclure, les apports après 12 ans d'expérience sont détaillés ainsi que les perspectives.

## ► Des aspects juridiques jusqu'à la norme d'échanges : des conventions entre les acteurs...

Les échanges se présentent sous différentes dimensions. Pour identifier les acteurs et le niveau de service lié à l'échange (fréquence, volume, délai de traitement...), ils sont encadrés par une convention juridique qui précise l'infrastructure technique permettant de garantir la robustesse et la traçabilité et définit les éléments fonctionnels liés à la structure des données et à la qualité attendue.

<sup>1</sup> Inspection générale des affaires sociales et Inspection générale des finances.

Les données doivent être accessibles, échangées uniquement entre acteurs autorisés et seulement sur le périmètre correspondant à leur besoin d'utilisation. Les échanges sont encadrés par une convention de service fixant les objectifs et modalités (fondements



***Les données doivent être accessibles, échangées uniquement entre acteurs autorisés et seulement sur le périmètre correspondant à leur besoin d'utilisation.***



juridiques, nature, modalités de transmission et de conservation des données, sécurité et traçabilité des échanges, échéances à respecter, règles de protection des données, conditions financières et modalités de suivi de la convention).

Par exemple, les données des déclarations employeurs issues de la déclaration sociale nominative (DSN<sup>2</sup>) (Humbert-Bottin E., 2018) et les autres ressources (exemple : retraite)

sont transmises et partagées mensuellement pour le prélèvement à la source ou encore pour les prestations sociales quand elles sont soumises à conditions de ressources (allocation logement, pension de réversion...). Pour ces échanges, une convention régit les rapports entre le promoteur du service, le Groupement d'Intérêt Public de Modernisation des Déclarations Sociales (Gip-MDS), et les maîtrises d'œuvre qui y contribuent :

- dans le cas de la DSN d'une part la Cnav et d'autre part l'Urssaf<sup>3</sup> Caisse nationale ;
- dans le cas de Pasrau<sup>4</sup>, la DGFiP<sup>5</sup>.

## ► ... une infrastructure technique...

Historiquement, les régimes et organismes de la sphère sociale ont instauré des liens informatiques spécifiques pour chaque échange. Les difficultés sont d'en assurer la robustesse, la sécurité et la traçabilité.

Des infrastructures ont été mises en place pour véhiculer, piloter et superviser chaque grand projet d'échange. Pour mettre fin à cette prolifération opportuniste, l'article R. 114-31 du code de la sécurité sociale crée le DGE<sup>6</sup> devenu un vecteur incontournable de communication et d'échange de données au sein de la sphère sociale.

Ce dispositif de gestion des échanges est un outil permettant les échanges entre Organismes de protection sociale (OPS), d'informations relatives aux assurés/bénéficiaires tout en garantissant la robustesse, la sécurité et la traçabilité. Il permet le routage de données concernant un unique assuré ou des millions d'assurés en flux de masse de la part d'un ou plusieurs émetteurs vers un ou plusieurs destinataires.

<sup>2</sup> DSN : Déclaration Sociale Nominative.

<sup>3</sup> Union de recouvrement des cotisations de sécurité sociale et d'allocations familiales.

<sup>4</sup> Le dispositif PASRAU (Passage des revenus autres) résulte de travaux de simplification et de rationalisation des déclarations sociales. Il est le prolongement logique de la DSN (Déclaration Sociale Nominative) et a constitué ces dernières années une simplification majeure des procédures déclaratives concernant les salaires et les revenus versés par un employeur. Le dispositif PASRAU (fondé sur la norme NEORAU) complète donc la DSN (fondée sur la norme NEODES) pour les « revenus de remplacement ».

<sup>5</sup> Direction générale des Finances publiques.

<sup>6</sup> Dispositif de Gestion des Échanges adossé au RNCPS, Répertoire National Commun de la Protection Sociale fournissant pour chaque assuré social l'identité des organismes lui servant des prestations.

Il répond ainsi à des besoins d'échanges journaliers entre organismes. De plus, il expose un catalogue d'échanges très utile dans une démarche « Dites-le nous une fois<sup>7</sup> ».

## ► ... une normalisation du contenu : le concept de norme d'échange



***Définir la structure des données et les dispositifs de partage de la connaissance, garantir la qualité et la cohérence des données des différents flux.***



L'étape suivante consiste à définir la structure des données et les dispositifs de partage de la connaissance et à garantir la qualité et la cohérence des données des différents flux.

Pour y répondre, le principe de norme d'échange pour caractériser toutes les dimensions des flux de données est formalisé comme suit :

- une structure fonctionnelle et technique du message avec :
  - une arborescence de blocs de données ;
  - le typage de chaque donnée ;
  - la liste des contrôles entre rubriques ;
- une cinématique de l'échange, c'est-à-dire l'enchaînement des différentes étapes de l'échange.

Au-delà de la définition théorique, une norme se matérialise par des livrables :

- la documentation de la norme ;
- des services de contrôle automatique ;
- un outil d' « auto-contrôle » automatique pour l'émetteur ;
- une gestion des versions des différents livrables : ce point est essentiel dans un contexte où la norme bouge beaucoup, notamment, en raison de contraintes réglementaires.

Le besoin n'est pas récent. Différentes normes ont déjà été appliquées dans la sphère sociale :

- la Norme de Dématérialisation Des Déclarations de Données Sociales (N4DS) porte la Déclaration Automatisée des Données Sociales Unifiées (DADS-U) émise par les employeurs ou les concentrateurs de paie en direction du portail [www.net-entreprises.fr](http://www.net-entreprises.fr) géré par le Gip-MDS ;
- la norme A identifie les personnes physiques pour le bénéfice de ses partenaires par le Système National de Gestion des Identifiants (SNGI) (*Préveraud de Vaumas, 2022*).

<sup>7</sup> Le principe « Dites-le-nous une fois (DLNUF) », consiste à éviter aux usagers de fournir, lors de leurs démarches en ligne, des informations ou pièces justificatives déjà détenues par d'autres administrations, en s'appuyant sur le partage automatique de données entre organismes.

En dehors de la sphère sociale, d'autres normes d'échange de données existent. Par exemple, l'arrêté du 19 octobre 2018 approuvant le schéma national des données sur l'eau, les milieux aquatiques et les services publics d'eau et d'assainissement et l'article R. 131-34 du Code de l'environnement, ont mis en place le Sandre (Service National d'Administration des Données et Référentiels sur l'Eau).

## ► De la structure fonctionnelle des messages : arborescence en blocs de données...



***La logique de structuration des données est le premier élément à prendre en compte dans la norme.***



La logique de structuration des données est le premier élément à prendre en compte dans la norme pour clarifier la présentation et formaliser la signification des objets utilisés et des liens entre eux. Elle s'appuie sur un format standard pour les données élémentaires.

Pour décrire ces notions, deux exemples :

- la DSN (norme NEODeS) et Pasrau (norme NEORAU), mises au point par le Gip-MDS avec la collaboration de la Cnav et des organismes de la sphère sociale ;
- la norme R pour les flux d'alimentation du Répertoire de gestion de carrières unique (RGCU) (Sureau & Merlen, 2021), mise au point par la Cnav, avec les partenaires de la sphère sociale.

La description fonctionnelle est structurée avec des rubriques élémentaires (nom d'une personne, etc.) regroupées en blocs de données cohérents portant sur les mêmes concepts fonctionnels (bloc identité dans le cas présent avec le nom, le prénom, le numéro d'identification, etc.) et les liens entre les différents blocs (un « bloc » individu possède plusieurs « blocs » adresse par exemple). La représentation est cohérente avec le modèle conceptuel de l'ensemble des objets et de leurs relations, préalablement défini.

Cette structure est présentée sous la forme d'un arbre pour présenter les blocs et leurs rubriques ainsi que les dépendances entre les blocs. Les normes R, NEODeS ou NEORAU ont donc une structure hiérarchique.

Cette structuration est ensuite déclinée en modèles de message reprenant tout ou une partie des objets précédents pour répondre aux différents besoins d'échanges spécifiques (alimentations ou restitutions par exemple). Dans chaque modèle de message, un bloc se caractérise par une cardinalité : présence une fois et une seule, 1 à n fois, 0 à n fois (*figure 1*).

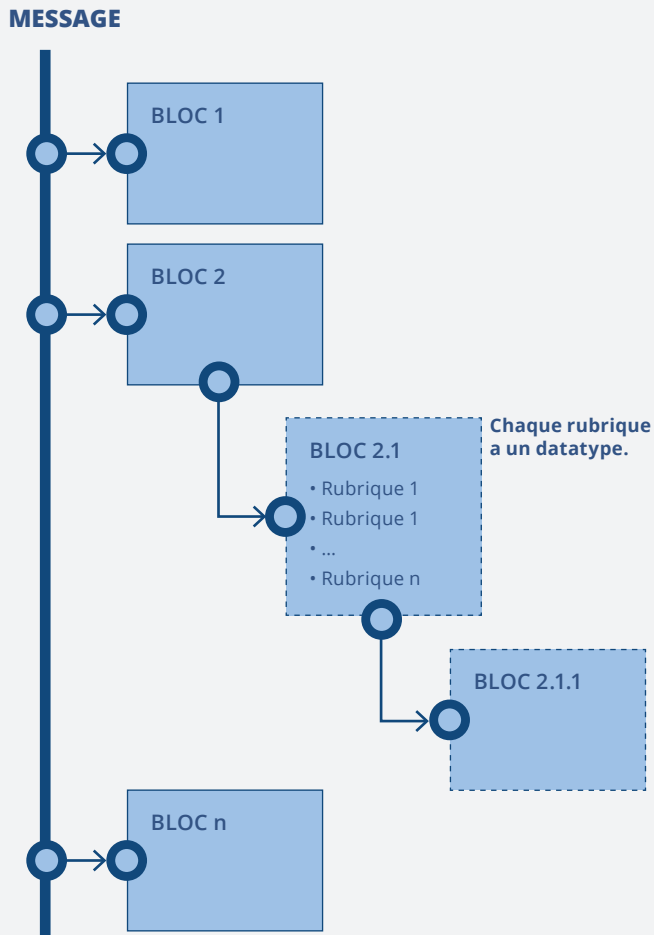
La norme comprend un en-tête avec les informations concernant l'émetteur et la nature et plusieurs messages contenant les données fonctionnelles.

Par exemple, pour la DSN, quelques blocs et les rubriques identifiantes dans chacun d'eux :

- Bloc « Entreprise » ;
  - SIREN (Système d'identification du répertoire des entreprises)
- Bloc « Établissement » ;
  - NIC (Numéro interne de classement : 5 derniers chiffres du code SIRET)

- Bloc « Individu » ;
  - Numéro d’inscription au répertoire (NIR)  
et/ou Numéro technique temporaire (NTT)
  - Nom de famille
  - Prénoms
  - Date de naissance
  - Bloc « Contrat (contrat de travail, convention, mandat) »
  - Numéro de contrat
  - Date de début de contrat
  - Identifiant du contrat d’engagement  
(pour rattacher les lignes de service au contrat de travail).

► **Figure 1 - Structure de la norme**



## ► ... à la structure technique des messages

La structure technique des normes d'échange évolue dans le temps ; les plus anciennes étaient généralement des lignes à champ de longueur fixe ou du texte avec séparateur<sup>8</sup> :

- Format « Texte longueur fixe »

Le format "Texte longueur fixe" a pour caractéristique principale d'utiliser des champs de taille fixe. Dans chaque colonne de données, toutes les valeurs ont le même nombre de caractères. Comme il est impossible que des noms propres ou des montants aient tous la même longueur, des caractères de remplissage sont utilisés pour combler les "trous" ;

**Un changement de modèle des données du référentiel en lien avec la norme peut avoir un fort impact.**

- Format CSV

Le format CSV est le plus universel. Il sépare les champs d'un enregistrement par une virgule ou un point-virgule et délimite les enregistrements par des *quotes*.

- Format XML

Ces problèmes sont résolus avec les nouvelles syntaxes issues des protocoles développés avec le web. Le XML (*eXtensible Markup Language*) est le standard le plus communément utilisé. Il permet de créer ses propres balises (adaptation au contexte de l'information représentée) en séparant les données (contenu et structure) de la forme (présentation). Le XML apporte les notions d'extensibilité (nouvelles balises), de lisibilité (structuration de balises) et de portabilité (interopérabilité entre les environnements d'exploitation des fichiers XML).

## ► De la structure fonctionnelle des rubriques...

Une démarche identique à la précédente est ensuite appliquée pour les rubriques. On descend d'un cran en passant de la formalisation du message à celle des rubriques qui la composent, tout d'abord dans une démarche fonctionnelle qui décrit précisément chaque rubrique, puis dans une approche plus « technique » avec la liste des caractères autorisés. Une définition précise de chacune est fournie ainsi que son format et la plage de valeurs qu'elle peut prendre. Le format précise la typologie du champ (numérique, chaîne de caractères) et sa longueur.

Outre le format, la liste des valeurs que peut prendre la rubrique est précisée. Selon la nature de celle-ci, la typologie des valeurs est expliquée :

- intervalle de valeurs pour les rubriques numériques (la rubrique peut par exemple prendre une valeur de 0 à 100) ;
- gamme de valeurs basées sur une nomenclature (exemple : la rubrique doit prendre la valeur 1, 2 ou 3) ;

<sup>8</sup> Voir l'article d'Alexis Dondon et Pierre Lamarche sur les formats de données dans ce même numéro.

- De cette façon, dans la norme, est associé à chaque rubrique ce que l'on appelle un « type » ou « *datatype* », caractérisé de façon précise par la nature de la rubrique (numérique, alphanumérique...), la liste de valeurs possibles (s'il y a lieu), l'expression régulière associée (s'il y a lieu), et les longueurs minimum et maximum (**figure 2**).

| <div>Type</div> <div>Prime.Type</div> <div>S21.G00.S2.001</div>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| <div>  <p>Motif définissant le type de la prime, gratification, supplément ou indemnité.</p> <p>Parmi les valeurs de cette rubrique, certaines relèvent du champ fiscal : Indemnité d'expatriation, Indemnité d'impatriation.</p> </div> <div>  <div> <div> <div></div> <div>x</div> <div></div> </div> <div>[3,3]</div> </div> </div>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |  |
| <div>  <div> <div>001 - Indemnité spécifique de rupture conventionnelle</div> <div>002 - Indemnité versée à l'occasion de la cessation forcée des fonctions des mandataires sociaux</div> <div>003 - Indemnité légale de mise à la retraite par l'employeur</div> <div>004 - Indemnité conventionnelle de mise à la retraite par l'employeur</div> <div>005 - Indemnité légale de départ à la retraite du salarié</div> <div>006 - Indemnité conventionnelle de départ à la retraite du salarié</div> <div>007 - Indemnité légale de licenciement</div> <div>008 - Indemnité légale supplémentaire de licenciement</div> <div>009 - Indemnité légale spéciale de licenciement</div> <div>010 - Indemnité légale spécifique de licenciement</div> <div>011 - Indemnité légale de fin de CDD</div> <div>012 - Indemnité légale de fin de mission</div> <div>013 - Indemnité légale due aux journalistes</div> <div>014 - Indemnité légale de clientèle</div> </div> </div> |  |

Compte-tenu de la diversité des types de caractère, de la quantité et du volume des flux, il convient d'être très rigoureux sur la structure technique de chaque rubrique, jusqu'au format des caractères. La gamme des caractères autorisés est aussi normalisée. À savoir : on autorise de 0 à 9, de a à z (majuscule et minuscule) ainsi que quelques caractères spéciaux comme & ' ( ) ... et on interdit les caractères spéciaux comme ! # \$ % < > ... (**figure 3**).

133

► **Contrôle du contenu des messages :  
les contrôles syntaxiques induits par le modèle...**

La description de la structure du message et des rubriques conduit mécaniquement à une série de contrôles à effectuer. Ils consistent à vérifier que le message respecte le « modèle de message » : par exemple, les blocs de données apparaissent dans l'ordre prévu, avec la bonne cardinalité (nombre minimum et maximum), les rubriques au sein de chaque bloc s'enchaînent dans le bon ordre, elles respectent le « type » prévu (par exemple, l'appartenance à une liste de valeurs prédéterminée ou la conformité à une expression régulière). Ainsi, on récupère les données attendues « dans les bonnes cases ».

Les contrôles fondés sur la structure sont appelés *contrôles syntaxiques*, puisqu'ils s'appuient sur la description d'une syntaxe, à la fois pour le message et pour chaque rubrique prise isolément.

► **... les contrôles sémantiques : cohérence interrubriques**

**Il faut assurer la cohérence interne du message.**

Cependant, les contrôles syntaxiques ne suffisent pas. Il faut assurer la cohérence interne du message. Ainsi, il existe au sein de chaque norme<sup>10</sup> d'autres contrôles, qui matérialisent ce besoin de cohérence entre rubriques au sein d'un même message.

► **Figure 3 - Table des caractères autorisés ou non pour une rubrique**

La table des caractères autorisés pour la valorisation des rubriques est un sous-ensemble de la table référencée ISO/IEC 8859-1. Les caractères interdits apparaissent sur fond grisé.

| ISO/CEI 8859-1 |                                                  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|----------------|--------------------------------------------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
|                | x0                                               | x1 | x2 | x3 | x4 | x5 | x6 | x7 | x8 | x9 | xA | xB | xC | xD | xE | xF |
| 0x             | caractères de contrôle et divers non imprimables |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 1x             |                                                  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 2x             |                                                  | !  | "  | #  | \$ | %  | &  | '  | (  | )  | *  | +  | ,  | -  | .  | /  |
| 3x             | 0                                                | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | :  | ;  | <  | =  | >  | ?  |
| 4x             | @                                                | A  | B  | C  | D  | E  | F  | G  | H  | I  | J  | K  | L  | M  | N  | O  |
| 5x             | P                                                | Q  | R  | S  | T  | U  | V  | W  | X  | Y  | Z  | [  | \  | ]  | ^  | _  |
| 6x             | `                                                | a  | b  | c  | d  | e  | f  | g  | h  | i  | j  | k  | l  | m  | n  | o  |
| 7x             | p                                                | q  | r  | s  | t  | u  | v  | w  | x  | y  | z  | {  |    | }  | ~  |    |

<sup>10</sup> Ou du moins chaque norme telle que ce concept est utilisé à la Cnav.

Il s'agit souvent de contrôles conditionnant la présence / absence d'une rubrique à la présence / absence d'une autre. Par exemple, dans le bloc Établissement de la DSN, il existe un contrôle exprimé en français : « Si le code postal est présent alors le code pays et le code de distribution à l'étranger sont absents et réciproquement. ». Ou, dans le cas de la norme R : « Si un élément de carrière est de type « Période assimilée maladie », on ne va pas trouver de revenu pour le régime général. ».

Certains contrôles inter-rubriques sont de simples inégalités : « La date de fin prévisionnelle de contrat doit être supérieure ou égale à la date de début du contrat. »

Il existe aussi des contrôles plus complexes sur le plan logique. Par exemple :

« Pour une déclaration donnée, le numéro de contrat doit être unique pour un établissement et un individu ».

Ces contrôles inter-rubriques, appelés contrôles sémantiques, sont écrits en langage naturel, afin qu'ils soient lisibles, et dans le même temps ils sont décrits en langage mathématique. Chaque contrôle sémantique est ainsi exprimé sous forme de formule logique, respectant une grammaire formelle prédéfinie.

La principale différence entre les contrôles syntaxiques et sémantiques (**encadré**) est la suivante : les seconds sont explicites (on écrit la règle de contrôle inter-rubriques), alors que les premiers sont implicites<sup>11</sup> (on décrit le type de la rubrique, ou la structure du message, et implicitement on dit que ce type ou cette structure doivent être respectés).

Quelques chiffres : en 2022, la Déclaration Sociale Nominative (DSN, norme NEODeS) comporte 7 modèles de messages, 61 blocs, 599 rubriques, 284 types. La norme R en 2023 comporte 26 modèles de messages, 163 blocs, 1174 rubriques, 282 types, 1163 contrôles sémantiques.

## ► La mise en œuvre *via* une cinématique d'ensemble des contrôles



***L'ensemble de ces règles de contrôle, syntaxiques et sémantiques, caractérise la norme d'échange.***



L'ensemble de ces règles de contrôle, syntaxiques et sémantiques, caractérise la norme d'échange. Chaque règle décrit « ce qui est interdit ». Il faut souligner que ces contrôles sont intrinsèques au message : Il s'agit de vérifier si dans le message, la structure d'ensemble est correcte et si les données sont bien celles attendues (le bon type, le bon ordre...) et sont conformes à des règles de cohérence.

À ce stade, on n'a fait que décrire des règles. Il faut maintenant les mettre en œuvre, et donc vérifier le message dans toutes ses composantes. C'est nécessaire sur le plan métier car cela permet à toutes les parties prenantes, utilisatrices du message,

<sup>11</sup> Il existe quelques rares cas particuliers où les contrôles syntaxiques sont décrits de façon explicite.

d'être assurées de la qualité du message, et en corollaire de la qualité d'alimentation des référentiels. Cette vérification est implémentée dans des interfaces applicatives (API<sup>12</sup>) qui exécutent des contrôles automatiques pour accepter ou rejeter les éléments contenus dans le flux.

Ces contrôles sont des contrôles bloquants. Le flux, dans le cas d'un traitement unitaire, fait l'objet d'un retour KO (le traitement s'arrête) et dans le cas d'un traitement de masse, il fait l'objet en retour d'un fichier de rejet (le traitement continue quant à lui pour les autres données du flux). Ces contrôles peuvent être complétés par des contrôles non bloquants qui permettent de continuer le traitement mais positionnent une alerte à l'utilisateur.

La succession des étapes, entre émetteur et récepteur du message, lorsque le message passe les contrôles, ou lorsqu'il ne les passe pas, est décrit avec précision dans la documentation de la norme.

### ► Encadré : un langage pour l'écriture des contrôles sémantiques

Le principe général est de décrire formellement une norme et de ne pas écrire de traitements : ceux-ci sont automatiquement générés. Ce principe s'applique aux contrôles sémantiques : on écrit simplement une expression logique (et non un traitement), qu'on associe à une rubrique.

Le langage utilisé n'est donc pas un pseudo-code. Il s'appuie sur une grammaire générale\*, et requiert le vocabulaire nécessaire à l'écriture d'une formule logique dans le contexte d'une norme :

Noms de constantes : nombres, codes, noms de rubriques, nom de la rubrique courante (\$rub),  
Connecteurs logiques (or, and, not, =>, , ...)  
Opérateurs de comparaison (=, !=, <, >=, ...), d'appartenance ensembliste (in), de présence (is\_present)  
Fonctions simples : opérateurs arithmétiques, sur les chaînes de caractères, sur les dates, ...  
Quantificateurs universels et existentiels (every, some, ...), et noms de variables (x, y, z)

Exemple : on veut s'assurer que si la rubrique **aa** vaut 08 et que **bb** est présente, alors **cc** appartient à {1,2}. Ceci s'écrirait : **((aa='08') and isPresent(bb)) => cc in {'01','02'}**

Autre exemple : dans la mesure où un même bloc **aa** peut apparaître plusieurs fois (par exemple un bloc contrat), on peut retrouver plusieurs fois la même rubrique. Si on veut qu'elles soient toutes différentes de la rubrique courante, on écrira :

**isPresent(\$rub) => (every x:aa satisfies (\$x != \$rub))**

La génération automatique des traitements est une opération complexe, car il faut « aller chercher » les différentes rubriques dans le message, qui a une structure arborescente, avec de multiples blocs. Cette complexité est cachée à l'utilisateur, à qui il suffit d'écrire les formules logiques.

L'existence de ce langage s'est révélée très utile au-delà des contrôles, pour le développement des transformations : elles sont là aussi exprimées comme fonction mathématique, à partir de laquelle on génère les traitements de transformation de norme proprement dite.

\* Techniquement, une grammaire compatible avec le framework ANTLR.

<sup>12</sup> Une API (*application programming interface* ou « interface de programmation d'application ») est une interface logicielle qui permet de « connecter » un logiciel ou un service à un autre logiciel ou service afin d'échanger des données et des fonctionnalités.

## ► Les livrables associés à une norme

La norme représente le format d'échanges pour l'envoi et la réception de flux entre les différents acteurs et en assurer la standardisation et la qualité. Les normes d'échange concernent de nombreux acteurs et sont susceptibles d'évoluer fréquemment (changements réglementaires, nouvelles demandes...). Un ensemble de livrables cohérents entre eux et systématiquement associés à un numéro de version doit être impérativement transmis à tous les acteurs.

Ces livrables sont les suivants :

- les documentations décrivant la norme ;
- le fichier informatique de description de la structure utilisable pour les développements ;
- les interfaces applicatives (APIs) des contrôles ;
- la brique de contrôle destinée à la validation de gros volumes de flux (dans la DSN, une performance de 1 million de lignes en 20 minutes en moyenne) ; elle contient la « base de connaissance » avec le modèle et les logiques de validation spécifique à une norme et elle est configurable pour les besoins d'industrialisation de la production ;
- un outil d'auto-contrôle pour que l'émetteur du flux puisse faire les contrôles avant envoi pour tests au destinataire.

## ► La documentation



***Le premier livrable, documentaire, est le cahier technique de la norme.***



Le premier livrable, documentaire, est le cahier technique de la norme. Il contient les objectifs, le périmètre, la description de la structure, ainsi que la définition de chaque bloc, chaque rubrique avec son format (*datatype*) et la gamme de valeurs associée. Le cahier

technique est proposé au format pdf et au format html pour faciliter la mise à disposition ainsi que la recherche d'information à partir du poste de travail.

Ce document est riche et très détaillé : dans sa version 2022, le cahier technique de la norme NEODeS fait 354 pages<sup>13</sup>. Il commence par une centaine de pages d'explications générales sur la DSN, ses usages, et sur toutes les notions liées à une norme d'échange. Le document décrit ensuite les différents modèles de message, en l'occurrence les modèles de déclaration (DSN mensuelle, DSN signalement fin de contrat, DSN signalement arrêt de travail), et pour chaque modèle, l'arborescence de blocs correspondante. Il explicite ensuite tous les blocs, toutes les rubriques dans les blocs avec pour chaque rubrique son type et les contrôles associés (*figure 4*).

<sup>13</sup> Voir <https://www.net-entreprises.fr/media/documentation/dsn-cahier-technique-2022.1.pdf>.

La documentation produite, au-delà du besoin de standardisation, comprend l'ensemble des informations et leur précision associée (blocs, rubriques, types de données, gammes de valeurs, contrôles appliqués...), tout en étant accessible grâce à une composition, un choix de maquette, des polices et des graphiques qui en facilitent la lecture.

La norme fait l'objet d'un affichage précis des versions successives ainsi que de la traçabilité des évolutions d'une version à l'autre, et elle est « à jour » et en cohérence avec les autres livrables techniques.

D'une version à l'autre, la liste de toutes les évolutions apportées est tracée dans un document spécifique livré avec chaque nouvelle version de norme.

Le cahier technique est complété par les **fichiers techniques de description du flux** (format xsd<sup>14</sup>) exposant la norme sous forme de schémas XML utilisables par les équipes de développement pour en dériver des objets.

► **Interface applicative pour les contrôles automatiques**

L'API ou brique de contrôle est utilisée par les applications qui garantissent la qualité du flux (et, dans le cas de la norme R, la bonne alimentation du référentiel de carrières). Cette API ou brique de contrôle applique l'ensemble des vérifications et contrôles de la norme pour autoriser ou rejeter les enregistrements en fonction du résultat du contrôle. Il existe une API ou une brique par norme.

► **Figure 4 - Extrait du cahier technique de la norme Neodes pour la rubrique « Nature juridique de l'employeur »**

**Nature juridique de l'employeur**  
**Etablissement.NatureJuridiqueEmployeur**

S21.G00.11.017



La nature juridique de l'employeur constitue ce qui définit en droit un employeur. Elle précise s'il est de nature privée ou publique. L'employeur est une personne physique ou morale qui a conclu un contrat de travail avec un salarié. Il exerce des pouvoirs de direction, de contrôle et de sanction. Il assume envers le salarié et à l'égard des administrations fiscale et sociale les obligations liées au contrat de travail.  
Nature du droit applicable à l'employeur.



CCH-11 : Cette rubrique est obligatoire si le Type de gestion de l'Assurance chômage (S21.G00.40.029) d'au moins un contrat de travail est renseigné de la valeur "03 - employeur ayant adhéré au régime d'Assurance chômage (adhésion révocable)" ou "04 - employeur ayant adhéré au régime d'Assurance chômage (adhésion non révocable)".



1. =

2. =

 x  [2,2]



01 - Privée

02 - Publique

03 - Etablissement privé à capitaux majoritaires publics

Pour chaque rubrique, sont indiqués le nom, le libellé, l'explication détaillée, le ou les contrôles sémantiques (ici le « CCH-11 »), les caractéristiques de domaine (type « liste de valeurs »), la longueur minimum et maximum et la liste de valeurs effective (01,02,03).

14 Xsd : xml schema definition : fichier de description du flux XML.



***Cette brique de contrôle applique l'ensemble des vérifications et contrôles de la norme pour autoriser ou rejeter les enregistrements en fonction du résultat du contrôle.***



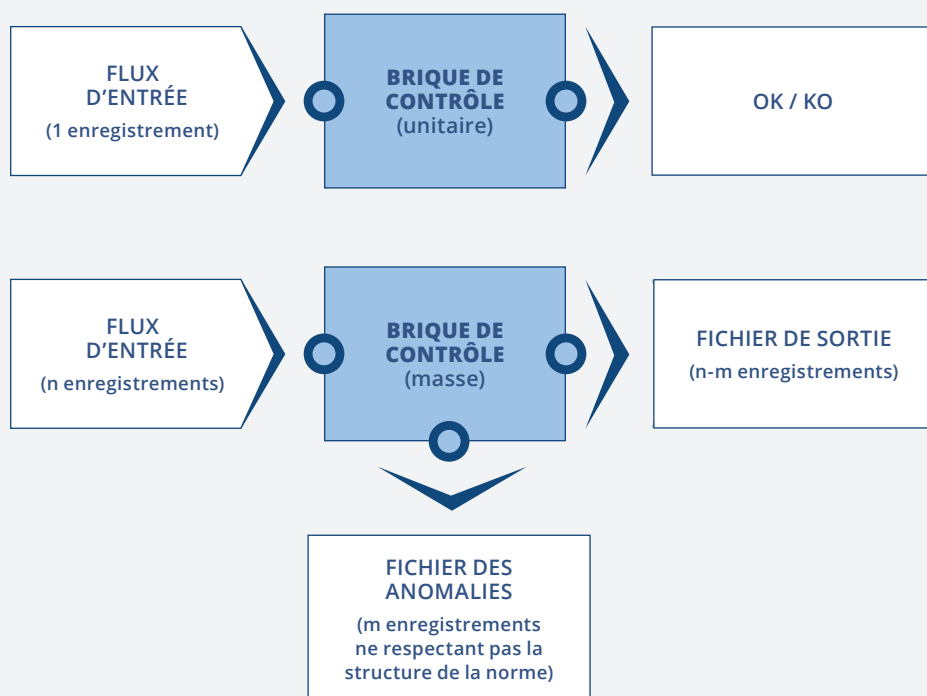
**L'API ou brique de contrôle** traite deux types de flux, les flux unitaires et les flux de masse (**figure 5**).

Dans le cadre d'une saisie par un technicien sur une application ou par un assuré sur le web (cas du flux unitaire), si un

seul contrôle s'avère négatif, le flux est rejeté : dans le cas du référentiel de carrières, cela implique que le flux n'est pas intégré dans le référentiel.

Dans le cadre d'échange en masse entre systèmes d'information, les enregistrements contrôlés « incorrects » sont rejetés. Les enregistrements qui satisfont les contrôles constituent un flux valide à intégrer. Un bilan de traitement du flux est envoyé à l'émetteur pour information et redressement des enregistrements incorrects.

► **Figure 5 - La brique de contrôle**



## ► Application d'autocontrôle

L'outil d'autocontrôle (ou application d'autocontrôle) est destiné aux émetteurs de données afin qu'ils vérifient la validité de leur flux avant de le transmettre à la plateforme de collecte, pendant les phases de test principalement.

Il s'agit d'une application qui comporte une interface graphique pour présenter le message à valider. Il génère le même bilan de traitement que la brique de contrôle implantée sur le serveur de collecte en production.

L'outil d'autocontrôle décrit la norme et identifie, dans le flux, toutes les données en anomalie. Il comporte aussi une aide en ligne et un mécanisme qui appelle le site de référence. En comparant sa version de la norme avec celle de référence, il propose sa mise à jour le cas échéant sans avoir à télécharger un nouvel outil complet.

Dans le cadre de la DSN, cet outil est très utile pour les entreprises, qui peuvent tester les déclarations sociales « chez elles » avant l'envoi dans les référentiels de la sphère sociale.

## ► Un besoin d'outillage et d'automatisation

Les structures d'échanges peuvent comporter un grand nombre de blocs, de rubriques et de contrôles (pour la norme Neodes ou la norme R, plusieurs centaines de contrôles) et avec des évolutions fréquentes. Le risque d'incohérence entre les livrables est important et peut devenir très coûteux pour les opérations de vérification voire de remise en cohérence de la norme et des développements qui l'utilisent.

Il y a une dizaine d'années, à la Cnav (par exemple sur la N4DS et la DSN), le principe était d'écrire le cahier technique et de le faire évoluer régulièrement en fonction des nouveautés

réglementaires ou des nouvelles demandes, puis de développer les contrôles automatiques et l'outil d'autocontrôle. Cela générerait un coût élevé et le risque d'incohérence était permanent entre les différents livrables.

***Pour gagner en efficacité et cohérence, il fallait disposer d'une description formelle en un seul endroit pour y décliner tous les livrables.***

Pour gagner en efficacité et cohérence, il fallait disposer d'une description formelle en un seul endroit pour y décliner tous les livrables... sans besoin de développer les contrôles ou réécrire la

documentation à chaque fois. C'était un gage de réactivité et de cohérence compte-tenu du nombre d'évolutions à gérer :

- plus de mille par an dans le cas de la N4DS à l'origine du projet ;
- plus de cinq cents par an par version de norme pour NEODES, sachant que trois versions de NEODES sont administrées de front : la norme en production de l'année N, celle de l'année suivante N+1 et celle de l'année d'après N+2.

Le principe appliqué est celui de la séparation des aspects métier et technique en construisant les outils les plus génériques possibles applicables à une norme d'échange. Le code relatif aux données et le code relatif aux traitements doivent être séparés (figure 6).

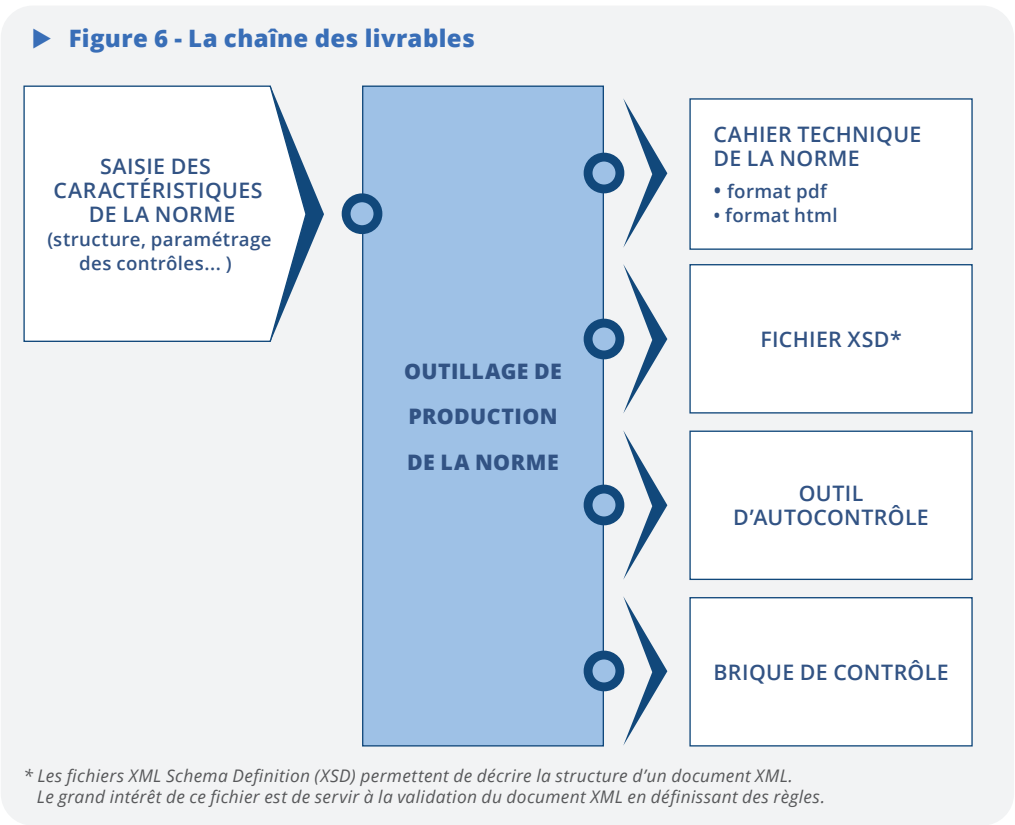
L'outil doit être basé sur la description d'un **modèle** et la déclinaison des livrables à partir de celui-ci.

La déclinaison automatique d'un modèle unique permet de maintenir tout au long des années **un modèle intègre, valide et tracé**. Toute opération sur le modèle est attachée à une série de spécifications s'inscrivant dans un cycle de publication versionné. À son tour, le pilotage projet reprend ce versionnage pour construire les plans de déploiement.

Cet automatisme permet d'associer une équipe « maîtrise d'ouvrage » qui spécifie la norme et une équipe « maîtrise d'œuvre » qui réalise la chaîne de validation des flux. Ainsi, cette équipe mixte avec ces compétences couvre la modélisation de la norme et la génération de la base de connaissance associée qui sera lue par un moteur de validation unique.

► **Saturne : un outil basé sur la modélisation**

Cet outil Saturne pour « Suite Applicative pour des Traitements Unifiés et Rationalisés fondés sur une Norme d'Échange » (Rivière P. et Rosec O., 2013), possède une interface, **l'Outil de gestion de la norme (OGN)**, pour décrire la norme. Il constitue alors un « référentiel de norme » fondé sur la modélisation. Cet outil permet de générer automatiquement la documentation, les formats XML et les outils de contrôle automatique ou d'autocontrôle.



Techniquement, l'implémentation fondée sur un modèle a été réalisée avec un *Domain Specific Language* (DSL). Il s'agit d'un langage concis et adapté à un domaine métier (Spinellis D., 2001 ; Fowler M., 2010). La mise en œuvre de ces principes a permis de faciliter l'implémentation, la généricité et la réactivité.

L'outil Saturne s'appuie sur un *framework*<sup>15</sup> de modélisation des normes. En première approche, la personne qui crée la norme n'a pas besoin de faire de développement. Elle modélise son projet par la saisie de la structure de la norme, des paramètres de contrôles dans Saturne, qui génère l'ensemble des livrables indiqués précédemment.

## ► Extension et industrialisation

---

La richesse et la complexité d'une norme nécessitent une approche industrielle pour la développer et pour en garantir la reproductibilité. Un processus et un outillage sont créés en ce sens et intégrés à une chaîne continue comprenant la validation de son niveau de qualité.

Le modèle d'une norme comporte des milliers de propriétés. La qualification de la norme passe par la rédaction d'un cahier de tests très complet et la réalisation de milliers de tests. La charge de réalisation d'une norme dense peut être de l'ordre d'une quinzaine de jours mais la couverture de test peut, elle, prendre 2 à 4 fois plus de temps.

L'approche industrielle de Saturne est fondée sur une démarche et un outillage d'intégration continue, complétés, compte-tenu du nombre de cas de tests liés au nombre de blocs et rubriques de la norme, par un outillage spécifique pour tester la norme avec un **outil de qualification de norme (OQN)**. Il permet de dupliquer à partir de stéréotypes d'instances de norme valides les jeux de tests pour gagner en délai et en qualité de ceux-ci. Deux modèles auxiliaires représentent d'une part les jeux de tests et d'autre part le bilan d'une campagne de test.

## ► Un outil complémentaire : l'outil de transformation de norme

---



*Un nouveau type d'outil  
a été créé : la transformation  
d'un flux de données  
d'une norme à une autre.*



Compte tenu des caractéristiques génériques de Saturne basées sur des modèles, un nouveau type d'outil a été créé : la transformation d'un flux de données d'une norme à une autre. La grammaire utilisée pour représenter les règles de contrôle dans la norme (**encadré**) est utilisée pour définir les règles de transformation d'une norme à une autre.

Selon le même principe, des programmes sont générés automatiquement à partir de la description mathématique des règles de transformation : c'est l'**outil de transformation de norme (OTN)**.

<sup>15</sup> Un *framework* est un cadre de travail, sorte de boîte à outils qui facilite le travail de développement informatique.

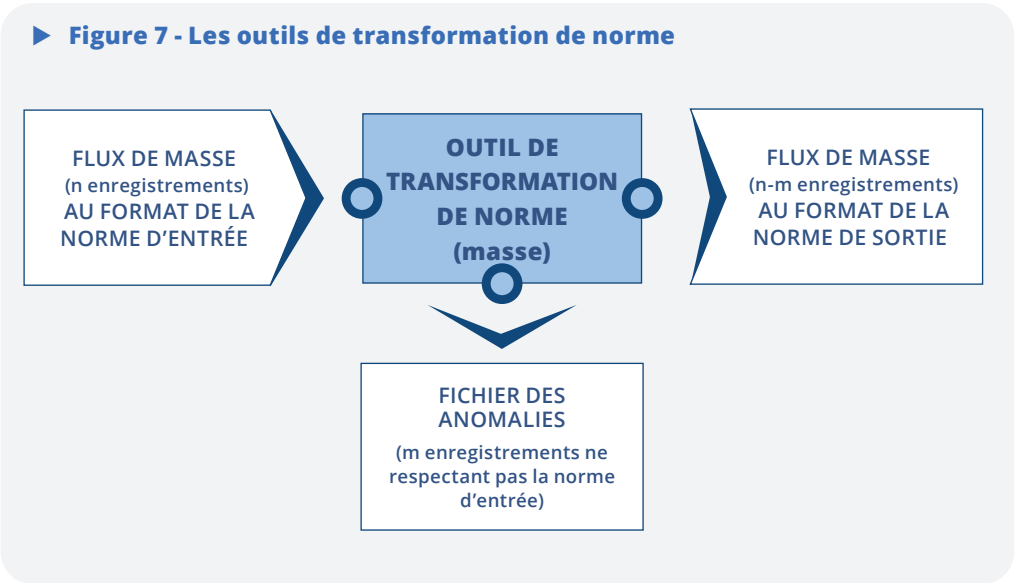
Les OTNs<sup>16</sup>, par l'utilisation de Saturne, bénéficient donc des mêmes avantages que la norme pour la cohérence des livrables, la standardisation de la documentation, la reproductibilité et la réactivité pour des demandes d'évolutions.

L'outil de transformation de norme (OTN) peut être conservé alors que le nouveau référentiel est généralisé, si des applications métier conservent les interfaces de l'ancien référentiel (figure 7).

La difficulté principale est de réunir les partenaires pour normaliser la transformation, écrire les règles de passage d'une norme à une autre, mettre en place la couverture de tests et le processus de qualification qui étend au processus de transformation la garantie de qualité que l'on veut faire peser sur le flux d'alimentation nativement à la norme d'entrée (Miotto E., 2011).

Dans le cadre du remplacement du SNGC (Système National de Gestion des Carrières) par le RGPU, les OTNs sont des éléments déterminants. En effet le SNGC était alimenté par de nombreux organismes de la sphère sociale (Cnam<sup>17</sup>, Pôle emploi, etc.) utilisant pour la plupart leur propre norme d'échange pour l'alimentation et la restitution des éléments de carrière. La bascule vers le RGPU de ces flux d'alimentation et de restitution a nécessité de les adapter à la norme R. Depuis, une vingtaine de normes ont été ainsi transformées vers la norme R.

De plus, dans le cas de l'alimentation du RGPU par la DSN, sachant que les deux normes sont partenariales, il faut écrire une spécification inter-régimes qui s'étend au fur et à mesure que ceux-ci basculent dans le RGPU. La spécification est publiée et éventuellement amendée au regard des remarques des représentants des régimes avant de devenir opposable pour la modélisation et le développement.



<sup>16</sup> Par abus de langage, «les» OTNs sont les différentes transformations de norme réalisées à partir de l'OTN, outil générique.

<sup>17</sup> Caisse nationale de l'Assurance Maladie.

## ► Une mise en œuvre de plusieurs années sur plusieurs normes

À l'origine, la suite Saturne devait supporter la N4DS (Norme pour la dématérialisation des déclarations de données sociales) qui, en 2012, remplaçait la norme DADS-U. Mais à la suite d'un prototype et à la relance du projet DSN (déclaration sociale nominative, norme NEODeS), fin 2011, un rapport Igas – IGF a posé comme prérequis la conception d'un format d'échange plus simple et concis que la N4DS.

La déclaration sociale nominative (DSN, norme NEODeS) a été déployée par phases portant des scénarios métier de plus en plus ambitieux. En phase 1, elle transmettait mensuellement à un bloc applicatif de stockage les données qui servaient à générer à la place de l'employeur les attestations maladie ou chômage dont a besoin le salarié. En phase 2, elle collectait les cotisations précomptées par les employeurs et l'instrument de leur paiement. En phase 3, elle collecte les données carrière qui serviront à calculer les retraites des salariés.

Fin 2016, le prototype de la norme NEORAu a été utilisé par les employeurs hors DSN (les fonctions publiques jusqu'en 2022) puis par tous les collecteurs d'impôt sur le revenu dans le cadre du prélèvement à la source entré en production en 2019.

En septembre 2014, la conception de la norme R (comme retraite) a porté toutes les interactions entre le répertoire de gestion des carrières unique (RGCU) et les systèmes d'information (SI) métier des plus de trente régimes de retraite français qui peu à peu vont y basculer. L'ancien système national de gestion des carrières (SNGC) était alimenté par

une multitude de flux hétérogènes. Dans le cas du RGCU toutes les interactions avec le répertoire sont unifiées dans la même norme, la norme R.

Pour l'ensemble des projets pour laquelle la norme a été implémentée, le gain est indéniable. La compréhension et le partage des concepts liés à la qualité de la documentation sont facilités et la cohérence des livrables est toujours présente. La réactivité grâce au nommage des

versions est vérifiée et est indispensable pendant la montée en charge des projets où le nombre d'évolutions reste très important. La mise en œuvre des contrôles automatiques permet d'assurer la qualité d'alimentation du référentiel dans le cadre d'arrivée régulière de nouveaux fournisseurs de flux. La mise à disposition d'outils d'autocontrôle autorise les émetteurs de flux à tester par anticipation l'adéquation de ceux-ci avec la structure de la norme et ainsi faciliter leur intégration lors des phases de tests, limitant ainsi les coûts et les délais.

D'un *Proof of Concept*<sup>18</sup> en 2012, Saturne est devenu un outil irremplaçable au service de deux programmes nationaux majeurs, la DSN et le RGCU.

<sup>18</sup> Le PoC, *Proof of Concept*, ou preuve de concept en français, est une méthode qui permet d'évaluer la faisabilité d'un projet.

## ► Perspectives : étendre cette offre

La mise en œuvre de grands référentiels avec l'accroissement du volume, de la variété des données et des échanges associés génère un besoin de gouvernance de la donnée. La qualité des données est une condition indispensable à l'adoption de ces référentiels. La gouvernance de la donnée s'inscrit dans la stratégie *data* destinée à fournir des services à valeur ajoutée.

Cette stratégie est aujourd'hui portée par une direction dédiée à la Cnav, la direction de la gestion de la donnée (DGD), chargée de délivrer des services liés à la donnée grâce à une plateforme technologique, et en amont d'une démarche de mise en qualité de la donnée indispensable à la fourniture de services à haute valeur ajoutée.



***Cette mise en qualité est une orientation forte de la stratégie data.***



Cette mise en qualité est une orientation forte de la stratégie *data*, que ce soit pour les SI internes de la Cnav ou pour les référentiels majeurs utilisés par toute la sphère sociale.

Les processus et les outils mis en œuvre autour de Saturne s'inscrivent pleinement dans cette orientation. Ils constituent aujourd'hui une suite industrielle qui tire parti des leviers de productivité offerts en matière de production logicielle par la génération de code par les modèles, la réutilisation de composants et la notion de ligne de produits.

C'est un gage de qualité pour l'intégration des flux dans les référentiels mais aussi un moyen de partage et de compréhension des données avec les différents partenaires. C'est aussi un gage de réactivité pour la prise en compte de modifications dans les phases de spécifications ou dans les phases de tests.

La documentation détaillée des données et l'apport de qualité sur les référentiels facilitent le travail de l'ensemble des acteurs utilisant les données, en particulier, les statisticiens pour qui la phase de prise de connaissance des données et de préparation est une composante lourde de leur processus.

La démarche aujourd'hui est de constituer une offre autour des principes de la norme et de la suite Saturne pour l'étendre à d'autres référentiels et bénéficier de ses atouts.

À la suite de ce retour d'expérience et dans le cadre de la démarche sur la stratégie *data*, la Cnav s'interroge pour étendre cette offre au-delà, à travers un cadre de mise en qualité général des données, dans un contexte d'échanges de plus en plus massifs entre les organismes.

## ► Bibliographie

---

- ALVISET, Christophe, 2020. La troisième refonte du répertoire Sirene : trop ambitieuse ou pas assez. In : *Courrier des statistiques*. [en ligne]. 29 juin 2020. Insee, N° N4, pp. 101-121. [Consulté le 6 juin 2023]. Disponible à l'adresse : <https://www.insee.fr/fr/information/4497083?sommaire=4497095>.
- AMOSSÉ, Thomas, 2020. La nomenclature socioprofessionnelle 2020 : Continuité et innovation, pour des usages renforcés. In : *Courrier des statistiques*. [en ligne]. 29 juin 2020. Insee. N° N4, pp. 62-80. [Consulté le 6 juin 2023]. Disponible à l'adresse : <https://www.insee.fr/fr/information/4497076?sommaire=4497095>.
- FOURMOND Vincent, 2005. *Les expressions régulières par l'exemple*. H & K éditeurs. Collection Technique et pratique. ISBN 978-2-914010-65-8.
- FOWLER, Martin et PARSONS, Rebecca, 2010. *Domain-Specific Languages*. 23 septembre 2010. Addison-Wesley Signature Series. ISBN 978-0321712943.
- GRATIEUX, Laurent et LE GALL Olivier, 2016. *L'optimisation des échanges de données entre organismes de protection sociale*. IGAS, RAPPORT N°2015-090R1 / IGF N°2015-N-055.
- HUMBERT-BOTTIN, Élisabeth, 2018. La déclaration sociale nominative. Nouvelle référence pour les échanges de données sociales des entreprises vers les administrations. In : *Courrier des statistiques*. [en ligne]. 6 décembre 2018. Insee. N° N1, pp. 25-34. [Consulté le 6 juin 2023]. Disponible à l'adresse : <https://www.insee.fr/fr/information/3647025?sommaire=3647035>.
- MIOTTO, Éric et VARDANEGA, Tullio, 2011. *On the integration of domain-specific and scientific bodies of knowledge*. In : *Model Driven Engineering*. [en ligne]. [Consulté le 6 juin 2023]. Disponible à l'adresse : [https://web.archive.org/web/20110724223732/http://adams-project.org/standrts09/proceedings/miotto\\_vardanega\\_standrts09\\_final.pdf](https://web.archive.org/web/20110724223732/http://adams-project.org/standrts09/proceedings/miotto_vardanega_standrts09_final.pdf).
- PRÉVERAUD DE VAUMAS, Joseph, 2022. Un référentiel des identités pour les besoins de la sphère sociale. Le système national de gestion des identifiants (SNGI). In : *Courrier des statistiques*. [en ligne]. 29 novembre 2022. Insee. N° N8, pp. 93-114. [Consulté le 6 juin 2023]. Disponible à l'adresse : <https://www.insee.fr/fr/information/6665190?sommaire=6665196>.
- RIVIÈRE, Pascal et ROSEC, Olivier, 2013. *Model-Based Interchange Formats : a Generic set of Tools for Validating Structured Data against a Knowledge Base*. *Poster Workshop of the Complex Systems Design & Management Conference CSD&M 2013*. [en ligne]. pp. 127-138. [Consulté le 6 juin 2023]. Disponible à l'adresse : [https://ceur-ws.org/Vol-1085/CSDM2013\\_PW.pdf](https://ceur-ws.org/Vol-1085/CSDM2013_PW.pdf).
- SUREAU, Christian et MERLEN, Richard, 2021. Le Répertoire de gestion des carrières unique (RGCU). Un nouveau référentiel ouvrant des perspectives pour l'analyse sociale. In : *Courrier des statistiques*. [en ligne]. 8 juillet 2021. Insee. N° N6, pp. 64-81. [Consulté le 6 juin 2023]. Disponible à l'adresse : <https://www.insee.fr/fr/information/5398687?sommaire=5398695>.