

LE CODAGE DES CALENDRIERS PROFESSIONNELS

1 CALENDRIERS PROFESSIONNELS : POURQUOI LES CODER ?

Le recueil de données diversifiées concernant les expositions professionnelles est une spécificité de Constances, qui n'a pas d'équivalent à cette échelle. Parmi les données recueillies, figure l'historique professionnel des volontaires.

Lors de leur inclusion dans la cohorte Constances, il est proposé à chaque participant de compléter un questionnaire intitulé « Calendrier professionnel » reconstituant toute leur carrière professionnelle depuis leur premier emploi. Pour chaque épisode de carrière, il leur est demandé d'indiquer en clair la profession exercée et le secteur d'activité économique correspondant, ainsi que quelques informations complémentaires (dates de début et fin d'épisode, statut d'emploi, type de contrat, temps de travail) ; le cas échéant, ils doivent renseigner des épisodes d'interruption d'emploi. Un extrait du questionnaire figure en annexe 1.

La collecte et le codage des calendriers professionnels correspondent à deux types d'usage pour les chercheurs.

1- **Évaluation d'expositions professionnelles à diverses nuisances par appariement avec des matrices emplois-expositions.** Une matrice emplois-expositions est composée de lignes (les emplois) et de colonnes (les nuisances) ; les cellules à l'intersection d'une ligne et d'une colonne contiennent des indicateurs d'exposition à la nuisance concernée : indicateurs de niveau d'exposition (qui peuvent varier selon le degré de précision : uniquement oui-non, proportion du temps de travail avec exposition, ou valeur numérique) et de probabilité d'exposition. Les matrices sont également historisées (pour un même emploi, il peut y avoir plusieurs périodes datées qui correspondent à des indicateurs d'exposition différents).

Les emplois dans une matrice sont une combinaison d'un code Profession et d'un code Secteur d'activité économique. Il est donc nécessaire, pour apparier un historique professionnel et une matrice, de coder les emplois selon la classification utilisée par la matrice.

Les chercheurs peuvent avoir des besoins variés de données d'exposition pour une nuisance donnée :

Un flag exposition vie entière (oui/non)

Nombre d'années d'exposition

Valeur cumulée vie entière ou pour des périodes d'exposition (en années)

Les calculs d'exposition cumulées vie entière ou sur des périodes définies sont à la charge des chercheurs, mais il faut pouvoir leur fournir les données qui permettent les calculs.

2- **Reconstitution de trajectoires professionnelles (avec chronologie).** Certains projets de recherche s'intéressent aux trajectoires d'emploi plutôt qu'aux expositions professionnelles.

Le calendrier professionnel peut également constituer une aide pour l'interrogatoire des participants pour des enquêtes concernant les trajectoires professionnelles, en permettant d'avoir à disposition la trajectoire professionnelle pendant l'interrogation (par entretien ou sur questionnaire papier).

2 COMMENT SONT CODÉS LES CALENDRIERS PROFESSIONNELS ?

Le codage des calendriers professionnels fait l'objet d'une collaboration entre Constances et l'IRSET - Inserm UMR 1085 - Équipe ESTER à Angers (Pr. Yves Roquelaure) qui est en charge du codage.

Le codage consiste à affecter, pour chaque participant de Constances, un code profession et un code secteur d'activité pour chacun de ces emplois d'une durée d'au moins 6 mois ; les périodes d'interruption d'activité sont également codées. Les nomenclatures utilisées sont celles de l'Insee : PCS 2003 pour les professions (www.insee.fr/fr/information/2400059) et NAF 2008 (www.insee.fr/fr/information/2406147) pour les secteurs d'activité économique. Le codage est fait à partir du questionnaire « Calendrier professionnel » rempli à l'inclusion, et doit être mis à jour en cas de changement d'emploi pendant le suivi. Les procédures mises en œuvre sont les suivantes :

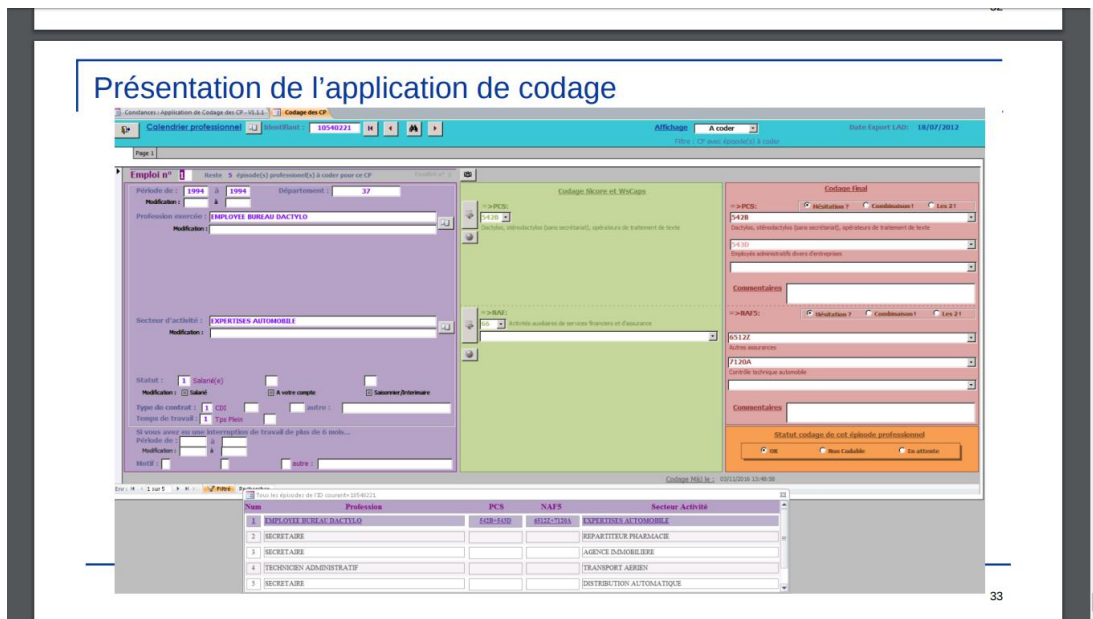
- Les questionnaires « Calendrier professionnel » sont saisis à l'Imprimerie Nationale, puis enregistrés par Constances, de même que le scan des questionnaires papier originaux ; le fichier est envoyé à ESTER-IRSET à Angers, à l'équipe en charge du codage proprement dit, coordonné par Fabien Gilbert.
- Des règles de vérification, de nettoyage et d'imputation de données incomplètes ou erronées ont été définies en amont et sont appliquées à Angers par ESTER-IRSET. Les données saisies dans l'application de codage sont directement intégrées dans un fichier de l'infrastructure Constances.
- Divers traitements sont effectués afin d'obtenir un livrable à fournir aux utilisateurs en fonction de leurs besoins :
 - Correction automatique de certaines dates erronées
 - Imputation de certaines dates absentes en tenant compte de l'enchaînement des épisodes professionnels pour un même historique professionnel
 - Mise en ordre chronologique des épisodes professionnels des participants Constances
 - Recodage systématique des codes NAF et PCS en fonction de certaines associations de libellés secteur d'activité et profession
 - Nettoyage de certains codes PCS ou NAF pour lesquels il n'y a pas d'ambiguïté pour le choix entre un code court ou long : on code systématiquement avec le code long (par exemple code PCS étudiant 8400 à la place de 84###)
 - Construction de variables complémentaires

Les règles de codage

Le codage de chaque épisode d'un calendrier professionnel est complexe, et il peut être difficile de coder correctement un épisode quand les informations fournies par les participants dans leur questionnaire sont imprécises ou incomplètes. Ce travail est fait manuellement par des codeurs, qui sont souvent amenés à prendre des décisions demandant une importante expertise.

Il faut tout d'abord que les codeurs analysent l'ensemble du calendrier professionnel, car toutes les informations fournies sont essentielles au codage. Celui-ci s'effectue par le biais d'une application créée par Constances et qui donne accès aux questionnaires numérisés. Elle permet de voir entièrement chaque calendrier professionnel avec un dérouleur qui apparaît et qui montre tous les épisodes du calendrier. Elle se compose de 3 cadres de couleurs différentes :

- Le cadre violet rassemble toutes les données écrites sur le questionnaire qui a été numérisé.
- Le cadre vert, qui est relié à SICORE (logiciel de codification automatique développé par l'Insee), est une aide pour le codage, mais doit être absolument vérifié avant de saisir le code.
- Le cadre rose est le cadre de saisie. Quand on commence à saisir, il y a un dérouleur automatique qui permet de voir le code plus précisément.



Avant de saisir et coder manuellement, plusieurs étapes sont indispensables pour avoir un codage de qualité.

1^{ère} étape : Comprendre le calendrier professionnel

La première question que l'on doit se poser quand commence à coder, est : « comment la personne a-t-elle répondu au questionnaire ? » En effet, en analysant le calendrier professionnel, on doit analyser les réponses données surtout au niveau du secteur d'activité. Il ressort en général trois façons de répondre au questionnaire :

- Des réponses très générales (ex : industrie, bâtiment, transport, ...). Il est donc impossible de coder précisément les secteurs d'activité et on se concentrera plus sur le code de profession qui donnera souvent plus d'information.
- Des réponses plus précises qui donnent des informations sur le secteur d'activité, mais pas assez pour aider à coder précisément (ex : comptabilité, informatique, commerce, ...). On devra bien prendre en compte le secteur d'activité et le relier à la profession pour coder au mieux.
- Des réponses très précises qui informent soit du nom de l'entreprise, soit des produits que l'entreprise fabrique ou vend, ou bien encore dans quel secteur d'activité la personne travaille précisément. Le codage sera alors au plus juste.

2^{ème} étape : Analyser les informations

Les informations fournies sont essentielles pour un meilleur codage. On doit toutes les prendre en compte.

- L'individu est-il salarié ou à son compte ? selon le statut, les codes ne seront pas les mêmes (ex : cadre informaticien salarié, le code sera 388#, alors que pour un indépendant le code sera 312e).
- Dans quel département travaille-t-il ? certains départements ont des incidences sur le codage (ex : un pompier qui exerce dans les départements 75 ou 13 sera un pompier militaire donc avec un code NAF 8422Z, alors que les pompiers des autres départements seront des pompiers avec un code NAF 8425Z).
- En quelle année la profession a-t-elle été exercée ? on peut déterminer si le secteur est public ou privé selon l'année et l'appliquer automatiquement sur le code de la profession (ex : France Telecom en 1960 était du secteur public, alors qu'à partir de 2010 ce sera du secteur privé).

3^{ème} étape : Coder les calendriers professionnels

Pour coder les épisodes des calendriers professionnels, on peut utiliser, soit des nomenclatures papier éditées par l'Insee, soit le moteur de recherche CAPS géré par Sante Publique France qui permet d'interroger l'ensemble des informations disponibles dans les nomenclatures par mot-clé. En premier, on commence toujours par coder le secteur d'activité en prenant en compte de la profession.

- Les nomenclatures papier sont très utiles quand on commence à coder pour voir l'ensemble des codes de la même famille et elles aident à les mémoriser plus vite. On a une vue panoramique de tous les codes de chaque catégorie et de leurs spécificités.
- L'outil informatique CAPS est un moteur de recherche par mot-clé. Il donne un résultat certes plus rapide, mais moins précis. En effet, il ne fournit que les résultats par mot-clé, il faut donc souvent vérifier plusieurs codes avant de trouver les plus pertinents.

Quand on débute le codage, pour coder au mieux, il est recommandé d'utiliser les 2 sources. On peut commencer par faire une recherche sur CAPS, en faisant attention de bien choisir le code en le vérifiant dans les nomenclatures papier qui donnent plus de précisions sur chaque code (cf. exemples de procédure de codage des calendriers professionnels en annexe 2).

4^{ème} étape : Vérifier les codes et obtenir une cohérence

Lorsque l'on code un calendrier professionnel, il est important d'avoir une cohérence entre chaque épisode.

- On doit prendre en compte l'évolution de carrière. Exemple : si une personne a été cadre à un moment de sa carrière, elle doit rester à ce rang pour les épisodes suivants.
- On doit aussi vérifier son parcours et voir si chaque épisode est logique dans celui-ci. Souvent une personne qui reste dans le même domaine aura une évolution de carrière qui devra être représentée dans les codes. Exemple : les codes profession se suivront généralement : 5###, 4###, 3###.
- Pour coder au plus juste, on prend en compte la hiérarchie et/ou la qualification. Exemple : ouvrier qualifié ou non, employé ou agent de maîtrise, cadre ou pas cadre, ...

Pour coder le plus correctement possible, il faut bien réaliser toutes ces étapes. Lorsque l'on observe l'ensemble des informations et qu'un doute subsiste, il ne faut pas hésiter à regarder le questionnaire scanné. L'important est de bien prendre en compte toutes les données fournies et de les rassembler pour obtenir le meilleur codage. Si on travaille en équipe, il est essentiel d'échanger et de discuter des points de vue de chacun pour faire le bon choix. Avant de valider le calendrier professionnel, on doit avoir, au mieux, le sentiment que le choix est cohérent et le plus juste possible. C'est la meilleure façon d'effectuer un codage précis.

3 LIVRABLES

Les résultats du codage se présentent sous forme de différentes bases de données.

- La base de données initiale des calendriers professionnels corrigés contenant les variables source et les variables nettoyées, des variables de signalement de modification de dates de périodes ou d'indication de cohérence ou d'incohérence d'historique professionnel.
- Une base de données avec la construction par ordre chronologique d'un historique professionnel par participant. Cette base peut être organisée de deux façons, selon la demande de l'utilisateur :
 - Une ligne correspond à un épisode professionnel complet avec les dates de début et de fin. Un historique peut comporter un ou plusieurs épisodes professionnels.
 - Les périodes d'emploi et d'interruption apparaissent par année calendaire, année par année. Une ligne correspond à une année. Il n'y a plus de dates de début et de fin de période. Il y a autant de lignes que dure en années la période d'emploi ou d'interruption.

- La base de données fournie à l'utilisateur contient l'ensemble des variables construites utiles pour juger de la qualité de l'enchaînement des dates des calendriers ; elle est accompagnée d'une documentation détaillée de tous les contrôles et traitements réalisés, des variables créées et de l'utilisation de la base de données des calendriers professionnels.

En décembre 2020, les calendriers professionnels complétés à l'inclusion de **198 255 participants**, comprenant **637 796 épisodes de carrière** avaient été codés. Les inclusions n'étant pas complètement terminées, le codage doit se poursuivre, ainsi que pour les changements d'emploi pendant le suivi.

ANNEXE 1 : QUESTIONNAIRE « CALENDRIER PROFESSIONNEL »



Comment remplir ce questionnaire

- Décrivez tous les emplois de **plus de 6 mois** que vous avez eus, **en commençant par votre premier emploi**.
- Si vous avez occupé plusieurs emplois très différents dans la même entreprise, décrivez-les séparément.
- Si un emploi a été suivi d'une période sans emploi (chômage, raisons de santé...), décrivez-la dans le cadre réservé à cet effet en précisant la raison.
- Notez la période de début et de fin d'emploi. Si vous ne vous souvenez plus de l'année exacte, indiquez une année approximative. Si votre dernier emploi est encore en cours, ne remplissez pas l'année de fin.
- Pour la profession exercée, évitez de noter des informations trop vagues, comme « fonctionnaire » ou « mécanicien » mais précisez exactement votre emploi (professeur ou mécanicien poids lourd).
- Ce calendrier est prévu pour 7 emplois, si vous en avez exercé un plus grand nombre, continuez sur une feuille libre sur le même modèle.

Exemple :

Période de : 1 9 9 8 à 2 0 0 1 Département : 3 3

Profession exercée : P E I N T R E E N C A R R O S S E R I E

Production ou secteur d'activité : R É P A R A T I O N A U T O M O B I L E

Statut : Salarié(e) A votre compte Saisonnier / intérimaire

Type de contrat : CDI CDD Autres

Temps de travail : Temps plein Temps partiel

Si vous avez eu une interruption de travail de **plus de 6 mois** entre cet épisode professionnel et le suivant, précisez la durée et le motif :

Période de : 2 0 0 1 à 2 0 0 2

Motif : Santé Chômage Autres

CALENDRIER PROFESSIONNEL

Premier emploi

Période de : A A A A à A A A A Département :

Profession exercée :

Production ou secteur d'activité :

Statut : Salarié(e) A votre compte Saisonnier / intérimaire

Type de contrat : CDI CDD Autres

Temps de travail : Temps plein Temps partiel

Si vous avez eu une interruption de travail de **plus de 6 mois** entre cet épisode professionnel et le suivant, précisez la durée et le motif :

Période de : A A A A à A A A A

Motif : Santé Chômage Autres

ANNEXE 2 : EXEMPLES DE CALENDRIER PROFESSIONNEL

Exemple de calendrier professionnel à 1 épisode professionnel

DATE DEBUT DE PERIODE	DATE DE FIN DE PERIODE	STATUT	LIBELLE SECTEUR D'ACTIVITE	LIBELLE PROFESSION	CODE NAF 2008	CODE PCS 2003	COMMENTAIRES NAF	COMMENTAIRES PCS
1995	2008	Salarié	Tertiaire	Employé de bureau	#####	543d	Le terme « tertiaire » est beaucoup trop vague pour coder la NAF. Ce libellé ne peut pas être codé.	Résultat de la recherche par mot-clé sur CAPS avec précision QND (qualification non définie), sachant que nous n'avons pas connaissance de son rang.

Exemple de calendrier professionnel à 3 épisodes professionnels

DATE DEBUT DE PERIODE	DATE DE FIN DE PERIODE	STATUT	LIBELLE SECTEUR D'ACTIVITE	LIBELLE PROFESSION	CODE NAF 2008	CODE PCS 2003	COMMENTAIRES NAF	COMMENTAIRES PCS
1975	1978	Salarié	Métallurgie	Tourneur fraiseur	24####	623f ou 623g	N'ayant pas plus de précision du secteur d'activité, on ne peut coder que les 2 premiers chiffres de la NAF.	En observant son calendrier professionnel en entier, on voit qu'il est dans l'épisode suivant responsable de production. C'est certainement un ouvrier qualifié. Par contre, on ne sait pas s'il travaille sur des petites ou grandes séries, on ne peut donc pas choisir entre les 2 codes.
1978	1991	Salarié	Sidérurgie	Responsable production	2410Z	474c	Résultat de la recherche par mot-clé sur CAPS	Comme on ne sait pas si c'est un cadre ou non, nous le laissons en agent de maîtrise.
1991	2003	A son compte	Chaudronnerie	Chef d'entreprise	3320A	212b	Résultat de la recherche par mot-clé sur CAPS	Comme il est à son compte, il a un code de statut indépendant, et sachant son activité et au vu de son calendrier professionnel, on le met en chaudronnier artisan.

Exemple de calendrier professionnel à 5 épisodes professionnels

DATE DEBUT DE PERIODE	DATE DE FIN DE PERIODE	STATUT	LIBELLE SECTEUR D'ACTIVITE	LIBELLE PROFESSION	CODE NAF 2008	CODE PCS 2003	COMMENTAIRES NAF	COMMENTAIRES PCS
2008	2011	Salarié	Education à l'image	Chargé de mission	85###	543d	Nous savons que le secteur d'activité est l'enseignement mais nous ne savons pas à quel niveau	En observant l'épisode suivant, nous pouvons en déduire qu'il est chargé de communication en apprentissage. On peut en conclure qu'il est certainement employé administratif à un poste de classe 5 et qu'il se forme à l'épisode suivant
2011	2013	Salarié	Mécénat culturel	Chargé de communication	8412z	464a	Cela fait partie du Ministère de la Culture	Sachant qu'il est encore en « apprentissage » et qu'à l'épisode précédent il était déjà chargé de mission, nous pouvons le mettre en classe 4 car il continue sa formation
2014	2014	Salarié	Cinéma	Production assistant	591##	465b	Le cinéma est un vaste secteur d'activité mais comme il est assistant dans la production, nous avons un peu plus d'informations pour coder, nous pouvons donc coder les 3 premiers chiffres de la NAF	Résultat de la recherche par mot-clé sur CAPS
2015	2018	Salarié	Syndicat Professionnel (cinéma)	Responsable des affaires économiques	9420z	461f	Résultat de la recherche par mot-clé sur CAPS	Il est responsable administratif et financier. Ce code est celui qui est le plus proche de sa fonction
2019	----	A son compte	Syndicat Professionnel (cinéma)	Délégué général	9420z	335a	Syndicat Professionnel (cinéma)	Même s'il est à son compte, il exerce un mandat syndical, on le code donc dans cette catégorie