

**Comité du label de la statistique publique
Secrétariat du comité du label
Commission « Ménage »**

**Enquête de cheminement tous niveaux de sortie du
système éducatif - Génération 2017**

**Dossier de présentation au Comité du label de la statistique
publique**

**Demande de label d'intérêt général et de qualité statistique
sans caractère obligatoire**

Séance du 18 décembre 2019

Sommaire

1. Demande de label d'intérêt général et de qualité statistique sans caractère obligatoire.....	5
2. Contexte de l'enquête.....	5
2.1 - Avis d'opportunité du Cnis.....	5
2.2 - Service producteur et service réalisant la collecte : maîtrise d'ouvrage et maîtrise d'œuvre ...	5
2.3 - Objectifs détaillés et principaux thèmes abordés.....	6
2.4 - Instances de concertation avec les partenaires sociaux.....	14
2.5 - Autres éléments de la comitologie (maîtrise d'ouvrage et maîtrise d'œuvre).....	17
2.6 - Calendrier de l'enquête.....	18
2.7 - Exploitation et diffusion des résultats.....	18
2.8 - Bilan d'exécution de l'enquête précédente et des résultats produits.....	21
2.9 - Bilan de la mise en œuvre des recommandations du Comité du label.....	22
3. Méthodologie statistique.....	24
3.1 - Champ, unités enquêtées.....	24
3.2 - Paramètres d'intérêt de l'enquête.....	25
3.3 - Description du sondage.....	27
3.4 - Traitements statistiques.....	40
4. Description du mode de collecte.....	47
4.1 - Protocole de collecte.....	47
4.2 - Sécurisation des données et protection du secret statistique.....	50
4.3 - Lettres-avis définitives, notices ou plaquettes d'informations.....	57
5. Questionnaire.....	64
5.1 - Compte-rendu détaillé des tests.....	64
5.2 – Questionnaire définitif (incluant les modifications résultant des tests).....	74
6. Coût des enquêtes.....	75
Liste des annexes.....	77
A.1- Avis d'opportunité du Cnis.....	77
A.2- Fiche de présentation de l'enquête présentée au Cnis pour examen d'opportunité.....	77
A.3 – Compte-rendu du Comité de concertation du 14 Novembre 2018.....	77
A.4 – Publications récentes.....	77
A.5 - Cahier des clauses techniques particulières (CCTP).....	77
A.6 – Inscription au registre des traitements de données à caractère personnel du Céreq.....	77

A.7 – Analyse d’impact (PIA) de l’enquête Génération 2017	77
A.8 – Questionnaire de l’enquête 2020 auprès de la Génération 2017 (dans un document séparé)	77

1. Demande de label d'intérêt général et de qualité statistique sans caractère obligatoire

Le Céreq demande le **label d'intérêt général et de qualité statistique – sans caractère obligatoire** - pour l'enquête de cheminement tous niveaux de sortie du système éducatif 2017, dite Génération 2017. Cette demande porte sur la première interrogation de la Génération 2017, qui aura lieu en 2020.

2. Contexte de l'enquête

2.1 - Avis d'opportunité du Cnis

Cette enquête a reçu l'avis d'opportunité au cours de la réunion du 12 avril 2019 (courrier en date du 14 mai 2019 - N° 88/H030).

L'avis d'opportunité délivré par le CNIS figure en annexe 1.

Il n'y a pas de changement par rapport à la fiche présentée pour l'avis d'opportunité (annexe 2) aux exceptions suivantes :

- Choix de l'organisme collecteur (voir 2.2)
- Détail des extensions géographiques et des extensions thématiques (voir 2.3)
- Le coût de l'enquête (voir partie 6. Coût de l'enquête)

2.2 - Service producteur et service réalisant la collecte : maîtrise d'ouvrage et maîtrise d'œuvre

Le Céreq assure la maîtrise d'ouvrage de l'enquête Génération 2017. Cette maîtrise d'ouvrage lui a été confiée par ses ministères de tutelle (ministères en charge de l'éducation et du travail). Elle s'inscrit dans l'un de ses grands axes d'étude sur les conditions de passage de la formation à l'emploi, selon les types et les niveaux de formation. Cet axe correspond à la mission inscrite au 1° de l'article R313-38 du Code de l'éducation. Elle est rappelée dans le Contrat d'objectifs et de performance 2019-2022, signé par le Céreq et ses tutelles (téléchargeable ici : <https://www.cereq.fr/le-cereq-organisation/direction-du-cereq>)

Le Céreq réalise également une partie de la maîtrise d'œuvre de l'enquête. Il organise la collecte de la base de sondage et la constitue par divers traitements, il définit et met en œuvre le protocole d'échantillonnage, il élabore le questionnaire, suit la collecte effectuée par le prestataire choisi et réalise les traitements post-enquête.

Le Céreq fait appel à des sous-traitants (choisis selon la procédure officielle des marchés publics) concernant certaines phases de la réalisation de l'enquête :

- La société H2A a été choisie pour réaliser la collecte de fichiers constituant une partie de la base de sondage (voir détails en 3.3)
- La société Médiamétrie est en charge de développer le questionnaire CATI-CAWI et réaliser la collecte multimode. Cette entreprise travaille elle-même avec un sous-traitant, Askia, pour le développement du questionnaire. Le CCTP lié à ce marché figure en annexe 5.
- La société Cap Adresse a été sélectionnée pour la mise à disposition de webservices permettant d'effectuer des recherches d'adresses postales normalisées et d'établissements employeurs dans la base SIRENE (Cap Adresse et Askia sont ainsi amenés à collaborer, le Céreq servant d'intermédiaire pour définir précisément les besoins)

Nous sommes sur le point d'opérer la notification des marchés pour les tâches suivantes :

- Restructuration, normalisation et validation des adresses postales des individus figurant dans l'échantillon
- Recherche de coordonnées téléphoniques pour ces mêmes individus
- Édition et envoi des lettres-avis

2.3 - Objectifs détaillés et principaux thèmes abordés

Les enquêtes « Génération » s'intéressent à l'insertion et au cheminement des sortants du système éducatif lors de leurs premières années de vie active. Elles ont pour objectifs principaux de produire des indicateurs d'insertion (taux d'emploi, taux de chômage, taux d'emploi à durée indéterminée, etc.), selon les niveaux de formation, les filières, les spécialités, à destination des acteurs publics et sociaux. Elles proposent ainsi des informations qui contribuent à la compréhension des processus d'insertion et des différenciations des parcours en début de carrière.

L'enquête 2020 auprès de la Génération 2017 permettra d'analyser l'insertion des jeunes sortis de formation initiale durant l'année scolaire 2016-2017 sur leurs trois premières années de vie active et de comparer les cheminements professionnels de ces jeunes à ceux des enquêtes Génération 1998, 2001, 2004, 2007, 2010 et 2013¹.

Le cœur de l'enquête est constitué par le calendrier d'activité qui permet de suivre mois par mois la situation des jeunes à l'issue de leur formation initiale, et de décrire les situations successives d'emploi et de non-emploi.

L'enquête décrira le parcours professionnel des jeunes depuis leur sortie de formation en 2017 jusqu'à la date de l'enquête au printemps 2020. Seront connues les situations des jeunes mois par mois : emploi, chômage, inactivité, formation/reprise d'études, avec une description de tous leurs emplois et de toutes leurs situations hors de l'emploi jusqu'à la date de l'enquête. Des éléments détaillés du parcours scolaire seront également recueillis (diplômes antérieurs, expériences de séjours à l'étranger, expériences de travail en cours d'études).

¹ La rénovation du dispositif et les choix faits en matière de modification du champ de la population enquêtée n'empêcheront pas les comparaisons mais réduiront cependant le périmètre des populations comparables.

Les enquêtes « Génération » permettent aussi de répondre à des demandes d'extensions nationales ou régionales portant sur les jeunes issus de certains niveaux ou spécialités de formations, de certains territoires ou sur les jeunes ayant bénéficié de certaines mesures pour la formation. Les échantillons peuvent donc être élargis en conséquence.

▪ Objectifs et contenus du questionnaire pour la partie nationale

Le tableau 1 présente les principales évolutions des modules thématiques entre l'enquête de 2013 auprès de la Génération 2010 et la prochaine enquête de 2020 auprès de la Génération 2017.

Au cours du processus de rénovation, des choix ont été faits visant à définir un « tronc commun » de questionnement, constitutif du cœur de l'enquête. Les objectifs de ce tronc commun sont d'identifier les informations qui à la fois apparaissent nécessaires pour reconstruire les parcours des jeunes et qui représentent des « forces » de l'enquête pour la réalisation d'études. Ainsi, l'architecture de ce tronc commun s'appuie sur deux piliers : d'une part, le parcours scolaire, d'autre part, le calendrier professionnel. La richesse des analyses produites dépend à la fois de la richesse des informations collectées et de leur pertinence dans l'éclairage des processus d'insertion.

En amont : le parcours scolaire

Celui-ci vise à identifier le plus haut diplôme atteint, clé de voute des analyses produites par l'enquête.

Les derniers travaux sur la Génération 2010 pointent que, au-delà du plus haut niveau de diplôme atteint, le parcours scolaire antérieur peut influencer sur la réussite du parcours professionnel ultérieur. Il s'agit donc de saisir l'hétérogénéité des différents cursus (en même temps que leur richesse) en récapitulant l'ensemble des diplômes obtenus.

Cependant, il s'agit aussi de comprendre les processus d'orientation qui contribuent à dessiner ces parcours scolaires. Les processus de sélection internes au système éducatif représentent des indications de performance des individus en son sein en même temps qu'ils renseignent sur le caractère plus ou moins contraint de ces orientations en cours de scolarité. Outre l'éclairage qu'elles apportent sur les processus d'orientation proprement dits, ces informations constituent des facteurs d'analyse précieux dans des analyses économétriques pour capter une part de l'hétérogénéité non observée entre individus.

Enfin, les conditions matérielles et économiques qui permettent - ou pas - aux jeunes de suivre des études, constituent un enjeu sociétal et de politique publique de plus en plus aigu (cf. infra le module sur le logement en cours d'études, sollicité par la DGESCO). L'enquête de 2020 auprès de la Génération 2017 développera donc un module sur ce thème, qui complètera le module sur les expériences de travail en cours d'études.

En aval : le parcours professionnel

Le calendrier d'activité mensuel représente le socle des enquêtes Génération depuis leur création. Il est à la fois la source de l'information longitudinale produite et la cheville ouvrière du recueil

d'information sur les différentes situations vécues par la personne enquêtée. Pour cette nouvelle enquête, il va connaître quelques aménagements. D'une part, il va être enrichi par l'ouverture de pop-up en cours de saisie, qui simplifieront la suite du déroulé du questionnaire. D'autre part, sa nouvelle mouture entérine le fait que la description de toutes les situations de non-emploi vécues au cours des trois dernières années peut constituer un exercice difficile pour un certain nombre de jeunes. Aussi, il a été acté de supprimer un descriptif détaillé de ces situations révolues pour se restreindre à un petit nombre de questions ayant principalement pour fonction de s'assurer un classement correct des individus dans les différentes situations.

Concernant le détail des situations recueillies dans le calendrier, quelques évolutions sont à noter.

Dans le prolongement des recommandations du rapport Gazier sur la diversité des formes d'emploi, un effort sera fait pour identifier les personnes travaillant avec un statut d'auto-entrepreneur. De même, des questions sur la multi-activité (au moment de l'enquête) sont introduites. Enfin, l'enquête Génération mettra en œuvre la PCS 2020 dans l'identification de la profession exercée.

Concernant les questions liées à la formation et l'obtention de nouveaux titres ou diplômes, dans un contexte de porosité croissante entre les deux univers, il a été décidé de se concentrer sur l'acquisition de nouveaux titres professionnels ou diplômes en fin de questionnaire, afin d'assurer la mesure d'une éventuelle élévation du niveau de diplôme. En conséquence, les questions autour des formations suivies ont été allégées.

Un module sur les intermédiaires du marché du travail est introduit. Il a pour vocation à synthétiser l'information auparavant recueillie de façon éparse dans les différentes situations de non-emploi et présente l'avantage d'être aussi posé désormais aux jeunes en emploi.

Informations sociodémographiques sur la personne

L'évolution majeure concerne l'introduction d'un tableau des mobilités résidentielles qui devrait permettre de mieux suivre les mobilités géographiques de la personne enquêtée.

Tableau 1 – Évolutions du questionnaire par rapport aux enquêtes précédentes

	THÉMATIQUES MAINTENUES	NOUVELLES THÉMATIQUES	THÉMATIQUES SUPPRIMÉES*
Parcours de formation initiale	Arrêt des études	Financement des études	Professionnalisation dans l'enseignement supérieur (extension)
	Diplôme de sortie, plus haut diplôme obtenu, diplôme(s) intermédiaire(s)	Double diplomation	Séjours à l'étranger (extension)
	Orientation (après la troisième, après le bac selon le niveau d'études atteint)		Stage(s) en cours d'études
	Expérience(s) de travail en cours d'études		Modules post initiaux (extension)
			Module Thèse (extension)
	Développement durable (extension)		
Parcours professionnel	Calendrier d'activité mensuel (Description de tous les emplois)	Auto-entrepreneuriat Multi-activité (indépendant, auto-entrepreneuriat, intermittent du spectacle, etc.) <i>PCS 2020</i>	Calendrier d'activité mensuel (description des situations de non emploi) <i>PCS 2003</i>
	Opinion sur l'emploi	Intermédiaires du marché du travail	Formation dans l'entreprise (allégée)
		Acquisition de nouveaux diplômes, certifications (PHD à 3 et 6 ans)	
Caractéristiques de l'individu	Origine sociale	Mobilités résidentielles	Secteur d'activité des parents
	Ressenti de discrimination à l'embauche	Ressenti de discrimination dans l'emploi	Caractéristiques sur la fratrie
	Perspectives professionnelles		Module sur le handicap (extension)
	Calendrier du mode de cohabitation		

* Une priorisation des thématiques a été réalisée pour définir le tronc commun de l'enquête. Les modules non maintenus dans ce tronc commun peuvent toutefois faire l'objet d'une commande par des partenaires d'extensions (c'est le cas du module sur les séjours à l'étranger par exemple).

▪ **Objectifs et contenus du questionnaire des extensions**

Les extensions d'échantillon (pour obtenir davantage de questionnaires) répondent aux demandes émanant d'acteurs publics (régions, ministères, acteurs du champ de la relation formation-emploi), qui souhaitent disposer d'indicateurs comparables à ceux de l'exploitation nationale pour des champs particuliers. Les extensions de questionnement visent, elles, à développer ou approfondir une thématique spécifique éclairant le processus d'insertion professionnelle pour tout ou partie des jeunes enquêtés.

Le tableau 2 présente les différentes extensions signées ou encore en discussion (pour deux d'entre elles).

Tableau 2 – Extensions de questionnement, de champ et de questionnaires de l'enquête Génération 2017

EXTENSION	THÉMATIQUE(S) DE L'EXTENSION	COMMANDITAIRE(S)
Logement	Conditions de logement au cours de la dernière année d'études : Distances logement / établissement de formation, logement / entreprise (de stage ou d'alternance), difficultés dans la poursuite d'études liées au logement	DGESCO DHUP (Direction de l'habitat, de l'urbanisme et des paysages) SDES-CGDD
Séjour à l'étranger en cours d'études	Séjours à l'étranger effectués à l'étranger en cours d'études et description plus détaillée des conditions du séjour le plus long	Erasmus + DGESIP Injep OFQJ (L'Office franco-québécois pour la Jeunesse)
Service civique	Identification des jeunes ayant effectué un service civique et datation dans le parcours	Agence du service civique
Formations de l'enseignement supérieur	Insertion des jeunes sortant de l'enseignement supérieur (tous diplômés)	DGESIP
Risques physiques et chimiques	Formation : sensibilisation ? Emploi : prévention par l'employeur ? Équipement spécifique ?	DARES
Formations environnementales	Formation et emploi en lien avec l'environnement ? (perception de l'enquêté)	CGDD

Attractivité de la Fonction publique	Perception de la fonction publique : - pour ceux qui n'y ont jamais travaillé - pour ceux qui y travaillent ou y ont travaillé transitoirement Lien (parents) avec la Fonction publique	DGAFP
Formations automobiles	Insertion des jeunes de la filière (CAP, Bac pro, BTS)	ANFA
Jeunes de QPV	Insertion des jeunes qui résidaient dans un QPV à la fin de leurs études	CGET
Sport	Insertion des jeunes sortant des formations des secteurs du sport et de l'animation. L'extension comprend les post-initiaux, habituellement hors-champ Céreq. De plus, le module multi-activité est davantage développé pour ce public (activité en lien avec le secteur ?).	Ministère des Sports Injep
Bretagne	Insertion des jeunes ayant terminé leurs études en Bretagne	Conseil régional de Bretagne
Pays de la Loire*	Insertion des jeunes ayant terminé leurs études en Pays de la Loire	Conseil régional des Pays de la Loire

* Extension à confirmer

Trois extensions (Bretagne, Pays de la Loire, jeunes des quartiers prioritaires de la politique de la ville (QPV)) correspondent à des extensions définies territorialement (basées sur l'adresse de l'établissement de formation pour les deux régions et sur l'adresse de la personne enquêtée pour la troisième). Elles constituent des extensions d'échantillon qui ont pour but de permettre la diffusion d'information sur ces zonages. L'extension financée par le CGET se singularise pour deux raisons. D'une part, le CGET est intéressé à recueillir l'adresse des personnes enquêtées au moment de leur bac (pour celles arrivées au moins en terminale). Ce type d'information permettra d'analyser, en plus des questions d'insertion (extrêmement difficiles pour ces jeunes), les parcours dans l'enseignement supérieur. D'autre part, le CGET a financé une extension portant sur les deux interrogations de la Génération 2017 (2020 et 2023).

Quatre autres extensions visent des filières de formation spécifiques (CGDD, ANFA, DGESIP, INJEP/Ministère des Sports). Ces extensions d'échantillon doivent permettre de produire de l'information statistiquement robuste sur les segments visés du système de formation. Deux de ces extensions s'accompagnent d'une demande de questionnement limité. Le CGDD souhaite savoir si la personne enquêtée estime, de son point de vue, que sa formation, le métier qu'il exerce et l'activité de l'entreprise qui l'emploie a bien une dimension environnementale. Le Ministère des Sports et l'Injep souhaitent, de leur côté, identifier, en cas de multi-activité déclarée, si le deuxième emploi décrit est dans le domaine du sport ou de l'animation.

Dans le même esprit, la DARES a sollicité une extension de questionnement auprès des jeunes sortis de formations professionnelles dans le but d'identifier quel est le degré d'information dont ils disposent (qu'ils ont retenus) sur les risques physiques et chimiques dans le travail. Le questionnement sur la prévention de ces risques concerne donc à la fois la formation suivie (ont-ils bénéficié d'un module de sensibilisation ?) et l'entreprise qui les emploie (ont-ils bénéficié de consignes, de rappels, de mise en garde, concernant les risques physiques et chimiques ? A-t-on mis à leur disposition un équipement spécifique ? A quels risques sont-ils exposés ?)

La DGESCO, la DHUP et le SDES-CGDD sont demandeurs, de leur côté, d'un module de questionnement sur le logement, au cours de la dernière année d'études. La préoccupation qui motive cette extension concerne la distance qui peut exister entre le domicile familial et l'établissement de formation professionnelle que fréquente (ou aurait souhaité fréquenter) l'individu. Cette distance peut constituer, pour les ménages aux revenus limités, une barrière à l'accès à la formation souhaitée. Dans cette perspective, le module s'attache à saisir les conditions de logement de la personne au cours de sa dernière année d'études. Il distingue la possibilité que la personne doive se partager entre deux logements, l'un lui permettant de suivre sa formation dans l'établissement scolaire, l'autre de réaliser la période en entreprise (stage, alternance) dictée par sa formation le cas échéant. Le module cherche à identifier les potentielles difficultés induites par cet éloignement, voire les impossibilités en termes d'orientation qu'il génère.

Concernant encore le parcours scolaire, le module sur les séjours à l'étranger en cours d'études est de nouveau reconduit. Le module a été adapté pour que l'on puisse connaître la durée totale passée à l'étranger pour chaque jeune et répondre aux deux objectifs européens concernant les mobilités en cours d'études (Benchmark). Le module vise à pouvoir mesurer l'incidence des séjours à l'étranger tant dans l'enseignement supérieur que dans l'enseignement secondaire. Ce module détaille les conditions du séjour le plus long effectué dans la période la plus récente (primaire, secondaire, supérieur). Si ce séjour est effectué dans le cadre scolaire, celui-ci est décrit plus précisément : type de validation qu'il permet, diplômé préparé, cadre dans lequel il a été financé, etc.

L'agence du service civique souhaite, pour sa part, pouvoir identifier les jeunes passés dans le dispositif qu'elle organise. D'une durée de six à douze mois, ce dispositif propose aux jeunes le souhaitant un engagement dans une activité. Si le dispositif est en général bien identifié (par les bénéficiaires au moins), le statut qu'il octroie n'est pas toujours très clair. Certains jeunes tendent à l'assimiler à une période d'emploi. Aussi, pour faciliter les traitements post-enquête, la question sur le service civique sera proposée à trois endroits distincts ; dans le pop-up des « autres situations » du calendrier d'activité, dans les statuts d'emploi, et en fin de questionnaire, explicitement, pour ceux qui n'en auront pas déclaré dans les étapes précédentes.

Le module sur l'attractivité de la Fonction publique (module de questionnement uniquement) a pour objectif d'éclairer la DGAFP sur la perception qu'ont les jeunes de ce secteur, leurs intentions et/ou leurs motivations. Le module distinguera donc ceux qui n'y ont jamais travaillé de ceux qui y travaillent ou y ont travaillé transitoirement. En complément, deux questions concerneront l'existence d'un lien dans la famille (parents) avec la Fonction publique.

- **Les points forts du dispositif Génération**

Un cadre d'analyse homogène et cohérent

Contrairement à d'autres enquêtes d'insertion qui visent des publics segmentés (apprentis, lycéens, sortants de grandes écoles ou d'université...), seul le dispositif « Génération » propose un questionnement, une méthodologie et un cadre d'analyse homogène pour tous, quels que soient le parcours scolaire, les diplômes obtenus, les domaines et voies de formation. Il est donc possible de comparer et d'évaluer l'impact de ces différentes caractéristiques sur les variations observées au cours des premières années de vie active : qui accède rapidement à un emploi ? Qui reste durablement au chômage ? À quel type d'emploi accède-t-on ? À quel niveau de rémunération ? Telles sont les questions auxquelles le dispositif permet de répondre. Plus généralement, il met en évidence les phénomènes de concurrence ou de complémentarité entre niveaux, domaines et voies de formation.

Des informations riches et diversifiées

Grâce à un questionnaire détaillé et un échantillon important, les enquêtes permettent, au-delà des caractéristiques du parcours scolaire et des diplômes obtenus, de prendre en compte d'autres critères. Le genre, l'origine sociale, l'origine nationale, le lieu de résidence, les mobilités géographiques, le statut familial, les réseaux sociaux, mais aussi la place et le rôle des dispositifs publics sont autant de dimensions que le dispositif Génération permet d'intégrer pour analyser les différences observées au cours des premières années de vie active.

Un recul temporel nécessaire

Certaines enquêtes d'insertion sont réalisées quelques mois seulement après la sortie du système scolaire. L'option retenue est alors de disposer d'indicateurs qui peuvent être mis rapidement à disposition des décideurs, des familles et des étudiants. Avec le dispositif « Génération », la première interrogation est réalisée trois ans après la sortie du système scolaire. Les résultats des premières enquêtes ont mis en évidence l'importance de ce recul temporel. En effet, il faut attendre plusieurs années pour que la stabilisation professionnelle soit établie pour le plus grand nombre. Enquêter tôt après la sortie de formation donne une photographie faussée des situations par rapport à l'emploi, qui accentue fortement les différences, alors que les enquêtes « Génération » montrent que celles-ci tendent à se réduire avec le temps.

Un suivi longitudinal

Le questionnaire permet aux jeunes débutants de décrire systématiquement, mois par mois, les différentes situations qu'ils ont connues depuis leur sortie du système éducatif. Ce mode d'interrogation permet de construire différents indicateurs comme le taux de chômage ou le taux

d'emploi, et d'aborder la qualité de l'emploi (niveau de rémunération, type de contrat). Il permet aussi de construire des typologies de parcours à partir de la description des situations mois par mois. Ces typologies offrent une vision synthétique des premières années sur le marché du travail : trajectoire d'accès rapide à l'emploi, trajectoire d'accès différé à l'emploi, trajectoire de décrochage, etc. L'insertion est une réalité multidimensionnelle qui ne peut se réduire à un ou deux indicateurs.

La même conjoncture pour tous

Les « générations » sont construites en fonction de la date de sortie de formation et non de l'année de naissance. Quel que soit leur niveau de formation, les jeunes arrivent donc dans un contexte de marché du travail plus ou moins favorable, mais identique pour tous. Il est donc plus facile a priori de comparer les trajectoires d'accès à l'emploi. Mais cette conjoncture a-t-elle les mêmes effets pour tous : à qui profitent les embellies ? Qui souffre le plus des retournements ? Quels effets sur les taux de chômage, l'importance des CDD ou de l'intérim, et pour qui ? Telles sont les questions auxquelles le caractère récurrent des enquêtes « Génération » permet de répondre.

Une enquête panéalisée

L'enquête « Génération », dans sa forme renouvelée, comprend une ré-interrogation en 2023. La première interrogation, trois ans après la sortie du système éducatif, est principalement tournée vers la caractérisation du processus d'insertion ; la ré-interrogation, à six ans, sera plutôt centrée sur les usages analytiques (notamment sur la question des parcours et des mobilités sur moyen terme) ; cette ré-interrogation permettra aussi d'approfondir les premières analyses issues de l'exploitation de la première interrogation.

2.4 - Instances de concertation avec les partenaires sociaux

Le dispositif des enquêtes Génération dispose d'un comité de concertation, qui se réunit deux fois par an pour examiner la préparation des enquêtes (questionnaire en particulier), le bilan des enquêtes administrées, et les premiers résultats d'exploitation des données.

Ce comité est composé des représentants des principaux partenaires institutionnels du Céreq : Direction de l'évaluation, de la prospective et de la performance (Depp), Direction de l'animation, de la recherche, des études et des statistiques (Dares), Direction générale de l'enseignement scolaire (Dgesco), Direction générale de l'enseignement supérieur et de l'insertion professionnelle (Dgesip), Délégation générale à l'emploi et de la formation professionnelle (Dgefp), Direction Générale de l'administration et de la fonction publique (Dgafp), France Stratégie, Pôle emploi, Institut national de la jeunesse et de l'éducation populaire (Injep), Institut national de la statistique et des études économiques (Insee). Y figurent également des partenaires sociaux (Medef, Cfdt, CGT, FO, Unsa, etc.), des représentants d'observatoires, d'organismes d'études et de recherche, des associations d'étudiants et de familles, de l'Union Nationale des Missions Locales et des représentants de la Politique de la Ville. Le tableau 3 liste l'ensemble des organismes membres du comité de concertation, en précisant s'ils sont financeurs de l'enquête. Ce financement peut se faire via les subventions versées par les deux ministères de tutelle du Céreq, ou via conventionnement avec des partenaires d'extension du questionnaire.

Tableau 3 – Composition du comité de concertation

Organisme membre du comité de concertation	Financier
CÉREQ	
Céreq, Deeva (Département Entrées et Évolutions dans la vie active)	
3 Centres associés régionaux au Céreq	
Les deux ministères de tutelle	
Ministère en charge de l'éducation nationale /DEPP	Subvention
Ministère en charge de l'éducation nationale /DGESCO	Extension
Ministère en charge du travail/DARES	Subvention+Extension
Ministère en charge du travail/DGEFP	non
Autres ministères ou opérateurs	
Ministère en charge de l'enseignement supérieur, de la recherche et de l'innovation/DGESIP	Extension
Ministère en charge de l'enseignement supérieur, de la recherche et de l'innovation/SIES	non
Ministère en charge de l'action et des comptes publics/DGAFP	Extension
Ministère de la Cohésion des territoires et des relations avec les collectivités territoriales/CGET	Extension
Premier ministre/CNEFOP	non
Pôle emploi	non
Institut national de la jeunesse et de l'éducation populaire (INJEP)	Extension
Ministère en charge de l'éducation nationale /ONISEP	non
<i>Premier ministre/France stratégie</i>	non
<i>Ministère en charge de la transition écologique et solidaire/CGDD</i>	Extension
<i>Ministère en charge de la santé/DREES</i>	non
Observatoires, organismes d'études et de recherche	
Institut national de la statistique et des études économiques (INSEE)	non
Réseau des observatoires de l'enseignement supérieur (RESOSUP)	non
<i>Observatoire régional des métiers (ORM PACA)</i>	non
Partenaires sociaux ou associations	
Confédération française démocratique du travail (CFDT)	non
Confédération générale du travail (CGT)	non
Force ouvrière (FO)	non
Mouvement des entreprises de France (Medef)	non
Union nationale des syndicats autonomes (UNSA)	non
Fédération des associations générales étudiantes (FAGE)	non
<i>Union nationale des associations familiales (UNAF)</i>	non

Note : en italique, les nouveaux membres par rapport à la précédente Génération enquêtée, Génération 2013

Lors de la rénovation du dispositif, les ministères de tutelles ainsi que d'autres ministères partenaires d'extensions à l'enquête, le conseil scientifique du Céreq, le comité de concertation ainsi que des

chercheurs extérieurs au Céreq ont été consultés. Les principaux éléments de cette rénovation ont été présentés et discutés lors de la séance du 9 novembre 2017 de la commission « Emploi, qualification et revenus du travail » du Cnis.

Les réunions du comité de concertation consacrées à l'enquête Génération 2017 se sont déroulées les 4 octobre 2016 et 14 novembre 2018. La réunion du 4 octobre 2016 faisait état des deux premières étapes de réflexions sur la rénovation de l'enquête Génération notamment sur les aspects méthodologiques et techniques, sur le contenu, l'architecture et le financement de l'enquête ainsi que la description des premiers scénarios de dispositifs rénovés.

La réunion du 14 novembre 2018, spécifiquement dédiée à l'enquête Génération 2017, a permis de présenter les conclusions de la troisième et dernière étape de la rénovation c'est-à-dire les principales évolutions de l'architecture du dispositif, du champ, du mode de collecte ainsi que celles du questionnaire. Elle a également permis de lister les possibilités de partenariats d'extension à venir.

Voici les points essentiels évoqués lors de cette réunion :

- l'architecture du dispositif : passage de 3 à 4 années d'intervalle entre deux Générations ; toutes les Générations sont interrogées à 3 et 6 ans (au lieu d'alterner des Générations « pleines » interrogées à 3-5-7 ans, et des Générations « légères » interrogées une seule fois à 3 ans)
- le champ de l'enquête : le champ demeure celui des primo-sortants de formation initiale de tous niveaux de formation. Trois changements sont introduits : la redéfinition des périodes de césure (de 12 à 16 mois), l'inclusion des sortants de contrats de professionnalisation dans le champ, et l'inclusion des résidents à l'étranger au moment de l'enquête (voir partie 3.1 pour plus de détails)
- le mode de collecte : passage à une collecte multimode (internet et téléphone) avec intégration d'innovations notamment pour la collecte de l'activité économique des établissements employeurs et le codage de la PCS à partir de la profession.
- la structure du questionnaire : une partie est définie comme « tronc commun », supposé fixe d'une Génération à l'autre. Des modules thématiques viennent s'y ajouter, en nombre limité et afin d'approfondir certaines thématiques. Un état des lieux des thématiques envisagées dans le tronc commun (maintien, suppression, ajout) a été présenté (mis à jour dans la partie 2.3 de ce dossier)
Le passage au multimode a par ailleurs impliqué de repenser le questionnaire (même pour les thématiques maintenues). Il a nécessité, d'une part, un travail de reformulation et d'adaptation des questions et modalités pour limiter au maximum les effets de mode et d'autre part, une réécriture complète des filtres. L'objectif étant de faciliter l'appropriation du questionnaire tant par les chargés d'études que par le prestataire de développement de l'outil de collecte.
- les possibilités d'extensions : l'appel à projets d'extension a été présenté. Il y a possibilité de conventionner pour étendre le champ (cas des sortants post-initiaux du sport), d'augmenter

la taille de l'échantillon pour un sous-champ (sortants d'une région, de QPV...) ou d'ajouter un module de questionnement sur une thématique non présente dans le tronc commun.

Le compte-rendu de cette réunion est joint en annexe 3.

2.5 - Autres éléments de la comitologie (maîtrise d'ouvrage et maîtrise d'œuvre)

- Conception de l'enquête

La conception de l'enquête, en termes d'élaboration du questionnaire, est prise en charge par le DEEVA (Département Entrées et Évolutions dans la Vie Active) au Céreq. Ce département est composé d'1 chef, d'1 cheffe adjointe, d'1 secrétaire et de 10 chargés d'études - chercheurs économistes, sociologues, statisticiens Insee - et de 5 chargés d'enquête qui forment une équipe dédiée à la production de l'enquête (l'EIGE, Equipe Ingénierie et Gestion d'Enquête).

Les réflexions sur les évolutions à apporter d'une enquête à l'autre (ajout ou suppression de modules, ajout ou suppression de questions ou de modalités) sont menées de façon collégiale au sein du département, en croisant l'approche technique de l'équipe dédiée à la production, et l'expérience des chargés d'études dans l'exploitation des précédentes enquêtes. Comme mentionné plus haut (2.4), l'ensemble des choix effectués ont également été soumis à la discussion de façon élargie, notamment lors des séances de comité de concertation.

Pour le cas particulier des extensions thématiques, une co-construction s'opère avec les commanditaires. Les partenaires sont en effet porteurs d'une problématique qu'ils souhaitent pouvoir étudier, le Deeva les accompagne dans la formalisation de l'écriture du module dédié.

- Pilotage de l'enquête

Au sein du Deeva, l'EIGE assure le pilotage de l'enquête. Cette équipe de 6 personnes suit le travail réalisé par le prestataire (Médiamétrie) chargé du développement du questionnaire. Des réunions hebdomadaires permettent de suivre l'évolution du questionnaire, et l'équipe teste au fil de l'eau les versions disponibles. Les retours sont alors envoyés à Médiamétrie pour correction.

Par ailleurs, divers « pilotes » (tests en réel du questionnaire) sont organisés. Au cours des pilotes, l'EIGE analyse et répond aux mails reçus par les individus sollicités en CAWI (sur une boîte mail dédiée en cas de problème rencontré) et est présente sur le plateau du CATI.

Suite au pilote du mois de septembre, une réunion est prévue prochainement entre EIGE et Médiamétrie afin d'établir la feuille de route en vue du prochain pilote, notamment en ce qui concerne les modifications du questionnaire à intégrer et le format des exports de données.

2.6 - Calendrier de l'enquête

Étape	Date
Collecte des fichiers pour la constitution de la base de sondage	Octobre 2018 – Juin 2019
Lancement de l'AO pour le développement et l'administration du questionnaire	Janvier 2019
Réunion de lancement avec le prestataire chargé du développement du questionnaire	22 mars 2019
Présentation au Cnis pour avis d'opportunité	12 avril 2019
Pilote 1 (1er test en réel du questionnaire)	16 -28 septembre 2019
Remise du dossier au Comité du label	12 Novembre 2019
Passage au comité du secret	Décembre 2019
Passage au Comité du label	18 Décembre 2019
Pilote 2	Fin janvier 2020
Pilote 3	Fin février – début mars 2020
Démarrage de la collecte	2 Avril 2020
Fin de la collecte	31 juillet 2020
Premières publications	1er semestre 2021

2.7 - Exploitation et diffusion des résultats

▪ Cadre général

L'exploitation des résultats nationaux sera principalement réalisée par le Céreq et son réseau de centres associés régionaux. Les partenaires d'extension pourront également réaliser des publications sur le champ de leur extension à partir des fichiers de diffusion qui leur seront remis par le Céreq en juillet 2021.

Les différentes exploitations porteront sur les données nationales, incluant les départements d'Outre-Mer. Des exploitations régionales seront réalisées par les régions ayant sollicité une extension.

Les résultats feront l'objet d'une première publication rapide par le Céreq à destination du grand public : un Bref de premiers résultats (premier semestre 2021) et un ouvrage « Quand l'école est finie » développant ces premiers résultats (automne 2021). À partir de l'automne 2021, des travaux d'approfondissement, davantage thématiques, seront publiés soit dans les collections habituelles du

Céreq (Bref, Céreq Etudes, Céreq Echanges, Céreq Essentiels, Working papers) soit dans d'autres publications : collections d'institutions partenaires (participation à l'État de l'école, l'État de l'enseignement supérieur, de la recherche et de l'innovation, Insee Références selon les dossiers proposés, collections de l'Injep, etc.), revues scientifiques, ouvrages en coédition ou participation à des chapitres d'ouvrages édités hors du Céreq.

Un groupe d'exploitation sera mis en place afin de faciliter et coordonner les premières exploitations thématiques autour de l'enquête. Ce groupe d'exploitation sera ouvert à la fois aux chargés d'études du Céreq et de ses centres associés, aux chercheurs extérieurs et chargés d'études des partenaires institutionnels du Céreq relevant du service de la statistique publique. L'appel à participation sera lancé au début de l'année 2021. La durée de vie de ce groupe d'exploitation est de dix-huit mois pour aboutir à des publications dans les collections du Céreq et dans des revues scientifiques.

Les fichiers d'exploitation - rendus anonymes- seront également transmis au centre Quetelet et mis à disposition de la communauté scientifique via le portail Quetelet-PROGEDO Diffusion, courant 2022.

Dans le cadre de sa politique d'information et de diffusion, un retour des résultats de l'enquête est prévu auprès des jeunes ayant participé à sa réalisation (envoi du Bref de premiers résultats). Ce retour s'inscrit également dans la démarche de fidélisation des jeunes enquêtés dans la perspective de leur ré-interrogation en 2023.

Dans le cadre de sa démarche de mise en valeur des données d'enquête qu'il recueille et exploite, le Céreq entend développer les possibilités de data-visualisation à partir de son site internet (recrutement prévu d'une personne en 2020 pour développer ce projet). Dans ce cadre, les données recueillies par l'enquête s'inscriront de façon prioritaire dans cette démarche.

▪ **Programme d'études**

Le programme d'études couvrira les principaux enjeux liés à l'enquête. Il ciblera des sous-populations pour la réalisation d'études sur leur insertion professionnelle :

- les sortants de l'enseignement supérieur, selon la nature du plus haut diplôme atteint
- les sortants de l'enseignement secondaire, avec des focus sur les sortants de la voie professionnelle et les sortants non diplômés
- les jeunes préparés à un même diplôme par des voies de formations alternatives : alternance versus voie scolaire. Au sein de l'alternance, seront distingués les sortants d'apprentissage des sortants de contrats de professionnalisation
- les jeunes sortis de formations ayant un lien avec des préoccupations environnementales (partenariat CGDD)
- les jeunes issus de différents sous-groupes sociaux. Ces études mettront en évidence les spécificités des conditions de leur insertion professionnelle et éclaireront l'impact des inégalités sociales sur la qualité de l'insertion et l'hétérogénéité qu'elles induisent dans la

valorisation du diplôme sur le marché du travail. Ses approches concernent le genre, l'origine sociale, l'origine migratoire, les jeunes des quartiers défavorisés

- les disparités territoriales d'insertion (selon les régions, départements, zones d'emplois...)

Certaines thématiques se focaliseront sur les parcours des jeunes, en formation comme sur le marché du travail, avec des problématiques comme :

- Au-delà du diplôme atteint, quelle est l'influence des processus d'orientation et du parcours de formation qui y a conduit ?

- Caractérisation des séjours à l'étranger effectués en cours d'études et leur lien avec le parcours professionnel

- Les conditions de logement des jeunes dans la dernière année d'études, leurs liens avec l'arrêt des études et leurs premiers pas sur le marché du travail

- Les modes de financement des études, leurs liens avec l'arrêt des études et leurs premiers pas sur le marché du travail

- Les risques physiques et chimiques : les jeunes y ont-ils été sensibilisés pendant leur formation ? Par leur employeur ?

- Comment caractériser les trajectoires professionnelles des jeunes de cette Génération, ont-elles évolué par rapport à celles de leurs aînés ?

- Analyse des mobilités professionnelles des jeunes de la Génération sur le marché du travail

- Analyse des mobilités géographiques des jeunes de la Génération sur le marché du travail, et leurs relations avec les mobilités en cours d'études

- Caractérisation des jeunes passés dans le dispositif du service civique pendant leur parcours professionnel

- Attractivité de la Fonction publique pour les jeunes de la Génération 2017, qu'ils y aient travaillé (même transitoirement) ou non au cours de leurs trois premières années de vie active

- Les « nouvelles » formes d'emploi sur lesquelles la rénovation du dispositif a mis l'accent : l'auto-entrepreneuriat, la multi-activité

Le programme d'études ciblera également certains apports de la rénovation du questionnaire, à la fois sous un angle thématique et méthodologique :

- travaux sur les interruptions d'études (inférieures à 16 mois), dont une éventuelle année de césure pendant la formation

- jeunes résidant à l'étranger au moment de l'interrogation

- implémentation de la nouvelle PCS « 2020 » dans l'enquête Génération

La rénovation du dispositif devrait également permettre de relancer des études ciblant de façon plus précise le type d'entreprises qui embauchent des jeunes, grâce à la rénovation de la collecte des informations sur les établissements employeurs.

Plus largement, dans le cadre d'une démarche de réflexion globale de la part du Céreq, seront envisagées les possibilités d'appariement des données de l'enquête avec des sources administratives complémentaires, concernant à la fois le parcours en formation et le parcours sur le marché du travail.

Enfin, le programme d'études abordera également différents aspects méthodologiques, liés à l'enquête. En particulier, la question de l'agrégation des données du fait du protocole multimode retenu dans l'enquête donnera lieu à des études méthodologiques, en lien avec les réflexions sur la question menées au sein du service statistique public.

2.8 - Bilan d'exécution de l'enquête précédente et des résultats produits

L'enquête 2020 auprès de la Génération 2017 est la première interrogation des jeunes trois ans après leur sortie du système éducatif. Les jeunes enquêtés à cette occasion seront ré-interrogés en 2023, soit six ans après leur sortie. Des deux Générations précédentes enquêtées trois ans après leur sortie de formation initiale, à savoir la Génération 2013 (une interrogation unique) et la Génération 2010 (à deux ré-interrogations), la seconde apparaît dans l'esprit la plus proche de la nouvelle enquête, compte tenu de sa structure de panel. Le bilan de l'enquête 2013 auprès de la Génération 2010 a été présenté dans un précédent dossier du label (session du 28 janvier 2015, dossier remis le 15 décembre 2014, partie 4, pages 5 à 15). Il a par ailleurs donné lieu à une publication dans les collections du Céreq, téléchargeable à l'adresse : <https://www.cereq.fr/bilan-methodologique-enquete-generation-2010-interrogation-2013>. Il n'est pas repris dans ce dossier.

L'enquête de 2013 auprès de la Génération 2010 a produit de nombreux résultats et alimenté les publications du Céreq au cours des cinq dernières années.

Ces publications ont débuté par des études de premiers résultats pointant la très nette dégradation des conditions d'insertion des jeunes de cette Génération vis-à-vis de ses aînées, en premier lieu du fait du contexte économique présidant à leur arrivée sur le marché du travail. Les moins diplômés ont particulièrement souffert de cette dégradation, voyant leur accès à l'emploi se compliquer, tant en CDD qu'en CDI. Les différentes études ont donné lieu à des exploitations centrées sur le devenir des jeunes selon leur niveau de sortie ou passés par une voie de formation par apprentissage. La question de la professionnalisation de l'enseignement supérieur, dans la foulée de la promulgation de la loi LRU de 2008, a été abordée via un questionnaire spécifique sur les différentes actions mises en place par les universités.

D'autres études se sont focalisées sur les parcours des jeunes. Ces études ont établi que les parcours en formation ont une importance dans le devenir ultérieur des jeunes sur le marché du travail, au-delà du seul plus haut diplôme atteint. Cette influence passe par la construction des cursus dans laquelle les initiatives, individuelles ou collectivement organisées, contribuent à davantage individualiser les

parcours que par le passé. Ces initiatives passent entre autres par des mobilités en cours d'études, qu'elles soient temporaires (séjours à l'étranger) ou plus durables (mobilités d'études).

Au-delà de la formation, les transformations des parcours professionnels dans les premières années de vie active ont aussi été étudiées. Elles ont mis en évidence la prégnance plus grande des retours en formation après quelques années passées sur le marché du travail, surtout le fait de jeunes sortis du système éducatif sur un échec scolaire. Cette montée des reprises d'études souligne la porosité croissante en France de la frontière entre système éducatif et système productif. Dans le même esprit, l'apprentissage, voie de formation particulière qui permet de poursuivre sa formation tout en se positionnant déjà dans le marché du travail, a continué à être étudié ; les travaux se sont particulièrement intéressés à son développement dans l'enseignement supérieur et ont souligné son impact différencié sur l'insertion selon le niveau du diplôme de sortie.

Finalement, différentes études ont pointé le rôle des inégalités sociales comme facteur d'hétérogénéité pesant sur la valorisation du diplôme au moment de l'insertion. Par exemple, plusieurs études soulignent la remarquable progression des conditions d'insertion des jeunes femmes ; cependant, une fois remis en contexte leur progression dans la scolarité plus remarquable encore, elles concluent à la persistance d'inégalités sur le marché du travail, se traduisant notamment par des différences d'accès à la catégorie cadres entre jeunes hommes et femmes diplômés de l'enseignement supérieur.

Différentes études méthodologiques ont également été menées, à partir de l'enquête Génération 2010. En effet, cette enquête a servi de support à des expérimentations préparant au passage programmé à une collecte multimode dans le cadre du nouveau dispositif Génération. Ce passage au multimode constitue une adaptation nécessaire du dispositif, à la fois aux évolutions de comportement et de pratiques des jeunes enquêtés et au cadre plus structurel visant à un contrôle plus strict des coûts de l'enquête. Ces études ont permis de poser les jalons du protocole multimode qui sera déployé dans la collecte de 2020 auprès de la Génération 2017.

L'annexe 4 présente la liste des publications issues des différentes exploitations réalisées à partir des données collectées dans l'enquête de 2013, classées par grandes thématiques.

2.9 - Bilan de la mise en œuvre des recommandations du Comité du label

Lors de la précédente édition de l'enquête Génération, le Comité du label a émis plusieurs recommandations à propos du champ de l'enquête, de la méthodologie, du protocole, du questionnaire et de la diffusion des résultats.

Concernant le champ, le Comité invitait à veiller à sa complétude pour les enquêtes futures. Celui-ci a donc été redéfini en prenant en compte les individus résidant à l'étranger et en redéfinissant les concepts d'interruption d'études et de césure. Un individu est défini comme sortant du système éducatif s'il n'a pas repris ses études au cours des 16 mois suivant son entrée sur le marché du travail. La césure est définie comme une période de suspension des études durant moins de 16 mois.

Le Comité recommandait également de finaliser les travaux méthodologiques en cours pour la rénovation du dispositif d'enquêtes, notamment ceux sur le multimode et la gestion du tirage de l'échantillon. Trois articles présentant les conclusions de ces travaux ont été publiés dans les actes des XIIIèmes Journées de Méthodologie Statistique en 2018² ; deux d'entre eux ont été présentés lors du colloque. Par ailleurs, une Note sur l'attrition et la déformation de l'échantillon des enquêtes après de la Génération 2010³ a été publiée conformément à la demande du Comité.

Autres aspects méthodologiques, les questions de calage et de correction de la non réponse. Le calage de l'enquête 2017 portant sur la Génération 2010 a été réalisé sur des marges de calage fournies à la fois par les enquêtes 2013 et 2015 ainsi que des sources externes ; de la même façon, il est prévu de caler les résultats de la deuxième vague de l'enquête auprès de la Génération 2017 (interrogation de 2023) sur ceux de la première vague (interrogation de 2020).

Des parodonnées ont été incluses aux modèles de correction de non-réponse, comme le recommande le Comité (avoir reçu une relance mail, avoir appelé le numéro gratuit...).

Au niveau du questionnaire, nous avons pris en compte l'une des préconisations du Comité, à savoir de ne pas parler de « nouvelle enquête » lors de l'invitation à une éventuelle interrogation en post-enquête (Question AC050). En revanche, dans la question sur les motifs de discrimination (DI030), nous n'avons pas ajouté l'item « pour raison de santé ou de handicap », car l'analyse du champ ouvert n'a pas fait ressortir souvent ce motif. Il reste possible pour les enquêtés de déclarer ce motif via la modalité « autre » (qui donne lieu à précision).

Enfin, les délais de mise à disposition des données aux chercheurs via le centre Quetelet ont été raccourcis et sont désormais de deux ans.

² Barret, C., Cissé, M., Dzikowski, C. "Plan de sondage des enquêtes Génération : utilisation d'un calage pour suréchantillonner les extensions". *XIIIèmes Journées de Méthodologie Statistique*, 2018.

Barret C., Cissé M., Gaubert E., Mazari Z., Olaria M., "Efficacité d'un protocole multimode (téléphone et internet)". *XIIIèmes Journées de Méthodologie Statistique*, 2018.

Barret C., Cissé M., "Agrégation de données multimode : Impact sur la modélisation des variables présentant un effet de mesure", *XIIIèmes Journées de Méthodologie Statistique*, 2018.

³ Cissé M., "Note sur l'attrition et la déformation de l'échantillon des enquêtes - Enquête Génération 2010", *Céreq Etudes*, n°28, 2019, 14p. <https://www.cereq.fr/note-sur-lattrition-et-la-deformation-de-lechantillon-des-enquetes-enquete-generation-2010>

3. Méthodologie statistique

3.1 - Champ, unités enquêtées

▪ Définition du champ de la Génération 2017

La Génération 2017 concerne les individus « primo sortants » de formation initiale en 2016-2017 (année scolaire), qu'ils soient diplômés ou non. Dans « primo-sortants » sont inclus les individus ayant interrompu leurs études par le passé pendant une durée comprise entre 1 et 16 mois. Tous les niveaux et domaines de formations sont concernés.

De façon plus précise, les critères d'éligibilité pour être dans le champ retenu, nommé ensuite « champ Céreq », sont les suivants :

- Avoir été inscrit dans un établissement de formation en France métropolitaine ou dans un DROM durant l'année scolaire 2016-2017 ;
- Avoir quitté le système éducatif entre octobre 2016 et décembre 2017 ;
- Ne pas avoir interrompu ses études durant plus de 16 mois avant l'année scolaire 2016-2017 (sauf pour raison de santé) ;
- Ne pas avoir repris ses études au cours des 16 mois qui ont suivi l'entrée sur le marché du travail ;
- Avoir 35 ans ou moins en 2017.

Ces conditions sont cumulatives.

Quelques points particuliers concernent l'application de ces critères :

- Bien qu'il s'agisse de contrats de travail, les contrats d'apprentissage et les contrats de professionnalisation à visée diplômante et qui interviennent dans la continuité du parcours scolaire initial sont considérés comme relevant de la formation initiale. Une personne repérée comme sortant de formation en 2016-2017 qui poursuit par un contrat d'apprentissage ou un contrat de professionnalisation en 2017-2018 est donc considérée en poursuite d'études, et donc hors champ.
- Une personne sortie d'un établissement de formation en 2016-2017 qui poursuit des cours par correspondance ou des cours du soir en 2017-2018 est considérée comme en poursuite d'études, donc hors champ, si elle n'a pas d'emploi en parallèle. Si elle a un emploi en parallèle, sa situation d'emploi prime.

Cas particulier, celui de l'extension d'échantillon sur les formations du sport et de l'animation : pour cette extension, les individus dits « post-initiaux » seront également interrogés. Il s'agit de personnes sorties de ces formations en 2016-2017, mais qui avaient déjà interrompu leurs études plus de 16 mois

par le passé (ils ne sont pas « primo-sortants » de la Génération 2017). Cette demande habituelle émanant du commanditaire (Ministères des sports et INJEP) se justifie par la spécificité de ces formations qui comptent un grand nombre de post-initiaux. Les questionnaires concernés ne font pas partie du fichier de l'enquête principale mais sont livrés au partenaire dans des fichiers dédiés.

▪ Évolutions par rapport aux enquêtes précédentes

Cette définition du champ pour la Génération 2017 présente des évolutions par rapport aux précédentes enquêtes :

- Auparavant, les jeunes en contrat de professionnalisation étaient considérés comme entrants sur le marché du travail dès l'obtention d'une rémunération au cours de leur formation. Autrement dit, ils étaient considérés sortants du système de formation initiale avant le début de leur contrat. Désormais, les jeunes dont les contrats de professionnalisation interviennent dans la continuité du parcours scolaire initial et ont une visée diplômante au même titre que l'apprentissage (*i.e.* titres professionnels et Certifications de Qualification Professionnelle exclus) seront considérés comme sortants du système éducatif au moment de leur sortie de la formation pour laquelle ils ont eu un contrat de professionnalisation.
- Suite à la parution de la circulaire de 2015 visant à faciliter et encadrer la pratique de la « césure » en cours d'études supérieures, il nous est apparu opportun d'étudier l'effet de cette nouvelle pratique sur les parcours d'insertion professionnelle. Pour cela, nous avons modifié la durée d'interruption des études à partir de laquelle l'individu est considéré hors du champ de Génération. Auparavant, ce seuil était de 11 mois. Or, la césure telle qu'encadrée par la circulaire de 2015 consiste souvent en une année universitaire complète. Ainsi, afin de percevoir les pauses d'une année scolaire, le délai à partir duquel un individu est réputé avoir interrompu ses études est dorénavant de 16 mois. Par conséquent seront désormais considérés comme primo-sortants les individus ayant interrompu leurs études par le passé pendant une durée comprise entre 1 et 16 mois ; ne seront plus considérés comme primo-sortants les individus ayant repris des études entre 0 et 16 mois après leur sortie. Ce « seuil de suspension » concernera tous les individus de l'enquête, quelle que soit la raison de l'interruption (inférieure à 16 mois) de leurs études.
- Les jeunes résidant à l'étranger (*i.e.* jeunes primo-sortants du système éducatif français qui résident à l'étranger – *i.e.* hors France métropolitaine ou DOM) à la date d'interrogation sont à présent intégrés dans le champ de l'enquête.

3.2 - Paramètres d'intérêt de l'enquête

L'enquête 2020 sur la Génération 2017 a pour objet principal de caractériser l'insertion professionnelle des jeunes, dans une perspective multicritère. Dans cette optique, elle doit permettre :

- De distinguer, une fois la personne sortie de formation initiale, **mois par mois, ses différentes positions sur le marché du travail**. Le premier paramètre d'intérêt distingue différents types de séquences : emploi, chômage, formation ou reprise d'étude, autre situation. Parmi les séquences collectées, celle correspondant au moment de l'interrogation (situation actuelle) revêt une importance particulière ;
- Parmi les séquences d'emploi, d'identifier pour les salariés **le contrat de travail** prévalant à l'embauche, à la fin de la séquence, et le moment où s'est opéré ce changement. Ces informations permettent ainsi de distinguer les personnes employées en CDD (sous ses différentes formes) de celles employés en CDI ou bénéficiant du statut de fonctionnaire, et du moment du passage de l'un à l'autre.

Ces deux premiers points aboutissent à reconstituer des calendriers professionnels couvrant les trois premières années selon les positions d'activité successives des personnes. Ces calendriers permettent, à leur tour, de :

- Calculer les **durées passées** dans les différents états
- Calculer le **nombre de séquences** par état.

Ces deux types de paramètres permettent une première caractérisation des cheminements professionnels des jeunes (par exemple, ils peuvent permettre de distinguer les situations de chômage récurrent des situations de chômage persistant). Ils s'avèrent importants pour compléter l'information sur la seule situation observée en coupe, au moment de l'enquête, trois ans après la sortie et permettent de distinguer les rythmes d'insertion, au-delà du mouvement de convergence vers l'emploi qui caractérisent une majorité des trajectoires, particulièrement parmi les diplômés de l'enseignement supérieur. Le calendrier permet également de :

- Calculer la **durée mise pour atteindre un état donné**. Typiquement, la durée d'accès au premier emploi est un paramètre utilisé (particulièrement parmi les peu diplômés) ; la durée d'accès au premier CDI est également un paramètre largement mobilisé.
- Réaliser une **analyse typologique** des calendriers professionnels, aboutissant à identifier des trajectoires-type intégrant les différentes dimensions. Classiquement, cette analyse a identifié sur les Générations précédentes une trajectoire de référence de stabilisation rapide et durable à l'emploi / en EDI.

L'enchaînement des séquences, la nature des transitions constituent également des indicateurs clés. Parmi les séquences autres que la situation à la date d'enquête, le **premier emploi** représente une séquence-clé des analyses puisqu'il permet de caractériser les conditions initiales d'accès à l'emploi. Différents travaux ont pointé l'ambivalence des CDD, entre opportunité (tremplin vers le CDI) et risque (trappe à précarité).

Cependant, pour les jeunes entrant sur le marché du travail, l'accès à l'emploi, ou non, et la nature du contrat de travail ne suffisent pas à caractériser les conditions d'insertion des différentes sous-populations étudiées. En effet, ces paramètres permettent d'opposer les grands niveaux de diplômes (non diplômés / diplômés du secondaire / diplômés du supérieur) mais ils sont beaucoup moins informatifs concernant des distinctions plus fines, particulièrement à la sortie de l'enseignement supérieur. Ils nécessitent donc d'être complétés par d'autres paramètres d'intérêt :

- La **rémunération perçue**,
- Le **niveau de qualification des emplois**

Ces paramètres permettent de développer des analyses en termes de rendement des diplômes mais également des approches en termes de déclassement (normatif, statistique ou salarial). Cet aspect permet du coup de souligner la place dans l'enquête de l'information sur **l'acquisition de diplôme après** la sortie de formation initiale.

Au-delà de la qualité de l'emploi occupé, l'enquête s'intéresse également à la nature de l'activité exercée et son éventuel lien avec la formation suivie. La **profession occupée** (PCS 2020 pour l'enquête de 2020) constitue donc un autre paramètre d'importance, au-delà du seul niveau de qualification du poste.

Outre ces informations factuelles caractérisant les situations professionnelles des personnes, un apport important de l'enquête est de pouvoir rendre compte et **caractériser le niveau de satisfaction / insatisfaction des personnes par rapport à leur activité de travail** ; sentiment d'être utilisé en dessous ou à son niveau de compétence, sentiment d'être (plutôt) bien ou mal rémunéré, sentiment de se réaliser professionnellement, etc., de pouvoir caractériser son rapport au travail exercé. Dans le même esprit, l'enquête permet de caractériser la dynamique temporelle des parcours, des aspirations et des anticipations des personnes, par des questions sur leurs priorités professionnelles, leurs aspirations à la mobilité (personnes en emploi déclarant rechercher un autre emploi), le sentiment d'insécurité qu'ils sont susceptibles d'exprimer⁴ (« être plutôt inquiet / optimiste pour son avenir professionnel »), le ressenti de discrimination à leur encontre, etc. Ces approches prennent davantage de relief encore dans la mesure où la Génération 2017 est un dispositif panélisté permettant donc de suivre les évolutions professionnelles y compris dans leur dimension subjective à travers ce rapport au travail exprimé.

Enfin, au-delà de la seule caractérisation de la « vie professionnelle » des personnes enquêtées, l'enquête constitue une ressource pour étudier les conditions de vie des jeunes. Ainsi, le **calendrier habitat** pointe les difficultés qu'éprouvent une partie de ces jeunes à accéder à l'autonomie résidentielle, quitter sa famille d'origine pour éventuellement en fonder une nouvelle. Dans le dispositif Génération, ces problématiques sont habituellement amorcées au cours de la première interrogation et développées dans l'interrogation suivante. Plus largement, les variables sociodémographiques permettant de caractériser sociologiquement les sous-populations étudiées constituent des variables importantes de l'enquête.

3.3 - Description du sondage

a. La base de sondage

Il n'existe pas de base nationale d'élèves nominative couvrant l'ensemble des sortants de formations initiales dispensées en France (métropole et DOM) pour une année donnée. Le Céreq constitue donc

⁴ cf. Séance du 5 novembre 2019 de la Commission Emploi, qualification et revenus du travail du CNIS sur la mesure du sentiment d'insécurité sur le marché du travail)

cette base de sondage pour chaque nouvelle enquête, à partir de différentes sources et en opérant divers traitements.

▪ **Une collecte de données organisée en deux temps**

Deux principales opérations de collecte ont eu lieu pour récupérer des listes nominatives d'élèves ou d'étudiants inscrits dans un établissement de formation initiale français en 2016-2017.

Dans un premier temps, une collecte de données nominatives a été réalisée en récupérant des bases nationales centralisées auprès de divers ministères :

- le fichier des élèves inscrits dans le second degré⁵, provenant du Système d'Information Statistique Consolidé Académique (SYSCA), fourni par le ministère de l'Éducation nationale (par la Depp, Direction de l'évaluation, de la prospective et de la performance)
- le fichier des élèves inscrits dans des formations en apprentissage, issu du Système d'Information de la Formation des Apprentis (SIFA), fourni par ce même ministère
- les données sur les étudiants inscrits en écoles d'architecture et d'arts provenant de la plateforme TAIGA, fournies par le ministère en charge de la Culture
- les données des élèves inscrits dans les lycées agricoles ou équivalents fournies par le ministère de l'Agriculture (DGER, Direction générale de l'enseignement et de la recherche)
- les données sur les inscrits dans des formations jeunesse et sport, issues de l'outil de gestion FORMations et diplÔMES de l'animation et du sport (FORÔMES), fournies par de la Direction de la Jeunesse de l'Education Populaire et de la Vie Associative du Ministère de l'éducation nationale
- les données sur les contrats de professionnalisation (base « Extrapro ») fournies par la Direction de l'Animation de la recherche, des Études et des Statistiques (Dares) ;
- les données des inscrits dans un dispositif CIFRE, fournies par l'ANRT (Association Nationale Recherche Technologie)

Dans un second temps, une collecte auprès de tous les autres établissements de formation (universités, écoles d'ingénieurs, de commerces, écoles de la santé et du social, etc.) est réalisée spécifiquement pour le Céreq avec l'aide d'un prestataire externe pour le contact des différents établissements. La liste des établissements à enquêter a été établie à partir d'une sélection dans la BCE (Base centrale des établissements – ministère en charge de l'Education), en sélectionnant les établissements ne figurant pas dans les bases centralisées. La collecte s'est déroulée entre janvier et juillet 2019.

Dans les deux cas, pour la collecte auprès des ministères comme celle auprès des établissements, les informations demandées concernant les élèves inscrits sont les suivantes : nom et prénom(s), INE

⁵ SYSCA inclut, en plus du secondaire, les classes de BTS.

(lorsque celui-ci est disponible), adresse, numéro(s) de téléphone, mail, date et lieu de naissance, sexe, diplôme et spécialité de formation, obtention ou non du diplôme préparé.

▪ Couverture de la base de sondage

La qualité de couverture des fichiers d'élèves est très variable selon le type d'établissement. Ainsi, les listes d'élèves recueillies à partir de fichiers centralisés par les ministères sont considérées comme exhaustives. En revanche, malgré tous les efforts de relance consentis auprès des autres établissements contactés directement, certains d'entre eux ne transmettent pas leurs fichiers d'élèves inscrits. Au total, 40% des établissements sollicités par notre prestataire ont transmis les fichiers demandés (Tableau 4). À noter que ce taux de réponse moyen gomme les fortes disparités entre établissements en termes d'effectifs. Par exemple, les « universités » et établissements rattachés, qui concernent les effectifs les plus importants, ont fait l'objet d'un effort particulier ce qui a conduit à un taux de réponse de 100%. De plus, par rapport aux enquêtes précédentes, davantage de fichiers ont été collectés de façon centralisée, ce qui permet de limiter le défaut de couverture de la base de sondage prise dans son ensemble. Ainsi, le taux de couverture global de la base de sondage est estimé à 87%.

Du fait de leur mode de construction, les bases de sondage de l'enquête Génération ont toujours présenté un défaut de sur-couverture, lié à des jeunes présents dans la base de sondage alors qu'ils sont en réalité hors champ. Pour tenter de réduire le taux de hors champ de la base de sondage de l'enquête Génération 2017, la méthode a été en partie revue cette année. Historiquement, dans le cadre de la collecte via le prestataire, le Céreq collectait des fichiers de « présumés sortants » des établissements. Il s'agissait de personnes inscrites l'année scolaire de référence dans l'établissement et non réinscrites l'année scolaire suivante dans *le même* établissement. Elles pouvaient néanmoins s'être réinscrites dans un autre établissement et donc ne pas faire partie du champ de l'enquête. Dans certains cas, nous arrivions à repérer ces poursuites d'études : nous disposions en effet des fichiers d'inscrits durant l'année scolaire suivant l'année scolaire de référence pour le secondaire (Sysca, anciennement BEA), pour les apprentis (SIFA), et pour les universités. Cette fois-ci, nous avons *systématiquement* collecté les fichiers des inscrits durant l'année scolaire de référence (c'est-à-dire 2016-2017) ainsi que ceux des inscrits durant l'année scolaire suivante (2017-2018), et nous repérons nous-mêmes les présumés sortants de formation initiale. Cela nous permet de détecter davantage de poursuites d'études ; nous espérons ainsi réduire le taux de hors champ de notre base de sondage (et le taux de hors champ parmi les répondants à l'enquête in fine).

Tableau 4 – Bilan de la collecte auprès des établissements

	Nombre d'établissements mis en production	Nombre d'établissements collectés	Taux de réponse
ÉCOLES D'INGÉNIEURS	149	81	54
ÉCOLES DE COMMERCE ET DE GESTION	289	102	35
UNIVERSITÉS ET ÉTABLISSEMENTS RATTACHES	69	69	100
UNIVERSITÉ DE TECHNOLOGIE	3	3	100
AUTRE ETAB.PUBLIC ENSEIGNEMENT UNIVERSITAIRE	10	6	60
PÔLE RECHERCHE ET ENSEIGNEMENT SUPÉRIEUR	3	0	0
AUTRE ETAB.PUBLIC ENSEIGNEMENT UNIVERSITAIRE	214	31	14
COMMUNAUTÉ D'ÉTABLISSEMENT	18	3	17
ÉCOLES DE FORMATIONS SOCIALES	350	132	38
ÉCOLES DE FORMATIONS DE LA SANTÉ	623	426	68
ÉCOLES DE FORMATIONS SANTÉ ET SOCIAL	56	31	55
COLLEGE SPECIALISE	1	0	0
ÉCOLE PROFESSIONNELLE SPECIALISEE	3	0	0
ECOLE SECONDAIRE SPECIALISEE (2 D CYCLE)	2	0	0
ÉCOLES COMPOSÉES UNIQT DE STS ET OU CPGE	240	65	27
ÉTABLISSEMENT RÉGIONAL D'ENSEIGNT ADAPTE	1	0	0
LYCÉE POLYVALENT	15	5	33
ESPE	25	22	88
ÉCOLES DANS LE SECTEUR DU SERVICE	312	77	25
ÉCOLES DANS LE SECTEUR INDUSTRIEL	54	20	37
ÉCOLES DE FORMATION AGRICOLE	10	5	50
CENTRES OU FACULTÉS PRIVÉES	4	1	25
LYCÉE D'ENSEIGNEMENT GÉNÉRAL ET/OU TECHNOLOGIQUE	42	6	14
CENTRES DE FORMATION DES APPRENTIS	33	2	6
ÉTABLISSEMENTS D'ENSEIGNEMENT SECONDAIRE AGRICOLE	376	109	29
CENTRES DE FORMATION PÉDAGOGIQUE PRIVÉS	31	7	23
ÉCOLES ADMINISTRATION PUBLIQUE	38	4	11
ÉCOLES NORMALES SUPÉRIEURES	4	3	75
ÉCOLES DE LA DGA	23	11	48
INSTITUTS ETUDES POLITIQUES	7	7	100
ÉCOLES ARCHITECTURES ET ARTISTIQUES	362	123	34
ÉCOLES JURIDIQUES	35	5	14
ÉCOLES DE LA DGAFP	45	19	42
Total	3447	1375	40

- **Les étapes de la construction de la base de sondage**

Après réception des divers fichiers, par les ministères ou par les établissements, un premier travail consiste à les agréger pour obtenir deux bases : une base d'élèves inscrits⁶ en 2016-2017, et une base d'élèves inscrits en 2017-2018. Dans un second temps, sont supprimés tous les individus de la base d'élèves 2016-2017 qui sont retrouvés dans la base 2017-2018, le but étant de ne conserver que les sortants de formation initiale dans notre base 2016-2017. Autant que possible, ce repérage des jeunes réinscrits durant l'année 2017-2018 se fait par un appariement des bases 2016-2017 et 2017-2018 sur la variable INE (Identifiant National Etudiant). D'année en année, cette variable est de mieux en mieux remplie et de façon plus homogène, ce qui rend efficace cet appariement. Cette année, les bases constituant la majeure partie de la base de sondage, à savoir celles concernant les élèves du secondaire, des apprentis, des établissements agricoles et des universités disposaient d'un INE pour l'ensemble des élèves. Pour les autres, lorsque l'INE est manquant, un appariement est effectué en créant une variable concaténant nom-prénom-date de naissance pour repérer les poursuivants.

Dans un troisième temps, sont supprimés autant que possible les autres individus hors champ de la base, par exemple les individus âgés de plus de 35 ans.

Enfin, un repérage des individus apparaissant plusieurs fois dans la base est effectué. Certains individus sont en effet inscrits en 2016-2017 dans plusieurs formations et/ou plusieurs établissements : dans ce cas, nous ne conserverons qu'une seule inscription en utilisant un ordre de priorité. Ce travail est actuellement en cours.

Tableau 5 – Effectifs provisoires de la base de sondage Génération 2017

Type d'établissement de sortie en 2016-2017	Effectifs provisoires
Présumés sortants Secondaire (SYSCA)	489 352
Présumés sortants Apprentis (SIFA)	182 327
Présumés sortants étab. Agricoles	46 796
Présumés sortants Universités (dont ESPE)	395 648
Présumés sortants formations du Sport	29 117
Présumés sortants Contrats de Professionnalisation (Extrapro)	147 572
Présumés sortants Ecoles architecture et arts (min. Culture)	15 426
Présumés sortants thèses CIFRE	1 829
Présumés sortants Ecoles santé/ social	51 279
Présumés sortants Ecoles d'ingénieurs	36 302
Présumés sortants Ecoles de commerce	26 063
Présumés sortants Ecoles de la Fonction publique	3 555
Présumés sortants Autres établissements (IEP,...)	30 436
Total	1 455 702

fichiers récupérés par la collecte prestataire
fichiers récupérés auprès des ministères

⁶ Inscrits dans un établissement ou sous contrat de professionnalisation

Les effectifs figurant dans le tableau ci-dessus sont provisoires, à double titre : nous avons encore certains fichiers de la collecte auprès des établissements à intégrer (fichiers arrivés tardivement, que nous devons mettre en forme avant de les compiler dans la base). Par ailleurs, certaines sélections n'ont pas encore été faites pour supprimer des individus hors champ (utilisation possible de la situation antérieure dans certains fichiers) et pour supprimer les doublons, comme évoqué plus haut.

- **Coordonnées dans la base de sondage : progression du taux de mails**

Le type de coordonnées et leur nombre varie fortement d'une base à l'autre. En moyenne, sur l'ensemble de la base provisoire constituée, 88,4% des individus ont une adresse complète. 80,1 % d'entre eux ont au moins un numéro de téléphone (tout comme les deux dernières Générations).

Au niveau des adresses électroniques, le taux de remplissage a nettement augmenté par rapport à la base de sondage de Génération 2013 : le taux de remplissage était seulement de 32%, il est actuellement de 49% en ne tenant compte que des mails d'élèves.

À noter que SYSCA (base du secondaire) est un cas particulier : seuls 17% des individus ont un mail élève renseigné mais 44% d'entre eux ont au moins un mail de parent. Ainsi, dans Sysca, 50% des individus ont au moins un mail parent ou élève. Étant donné le nombre d'individus concernés, la prise en compte des mails des parents dans Sysca conduit à un taux de mail global de la base de 60%, soit 11 points de gagnés par rapport au cas où seuls les mails d'élèves seraient pris en compte. La question de l'utilisation des mails des parents se pose donc. Avec pour question corollaire : doit-on adapter les mails-avis dans ce cas ? Cette question est à instruire et à trancher dans les prochaines semaines.

Tableau 6 - Récapitulatif des coordonnées présentes dans la base de sondage

	Présence d'une adresse complète (%)	Nombre moyen de numéros par élève	Nombre moyen de numéros de parents	Présence d'au moins un numéro élève ou parent (%)	Présence d'un mail élève (%)	Présence d'un mail parent (%)	Présence d'un mail parent ou élève (%)	Total individus dans la base de sondage
Base SIFA	97	0,6	0	63	0	0	0	182 327
Base SYSCA	79	0,6	2,7	98	17	44	50	489 352
Base étab. Agricoles	50	0,8	0	45	22	0	22	46 796
Base Min. Culture	100	1,2	0,8	92	99	0	99	5 958
Formations du sport	99	1,0	0	99	98	0	98	29 117
Base Extrapro	97	0,0	0	0	44	0	44	147 572
Formations Fonction publique	98	1,4	0	97	95	0	95	3 349
Formations santé/social	46	1,1	0	81	90	0	90	51 279
Thèses CIFRE	0	0	0	0	100	0	100	1 829
Universités gérées par Apogée	96	2,4	0	95	93	0	93	352 443
Autres établissements (autres universités, écoles de commerce, écoles d'ingénieurs...)	91	1,4	0	90	94	0	94	145 680
TOTAL	88,4	1,1	0,9	80,1	48,9	14,9	59,9	1 455 702

b. Échantillonnage

Pour la définition de la méthode d'échantillonnage de l'enquête Génération 2017, le Céreq a bénéficié de l'accompagnement de la DMCSI (Direction de la méthodologie et de la coordination statistique et internationale) de l'Insee. Des réunions régulières ont eu lieu pour valider les choix effectués.

L'échantillon de l'enquête auprès de la Génération 2017 est construit pour répondre aux objectifs suivants, sous la contrainte de 20.000 questionnaires financés par le Céreq pour l'enquête 2020 :

- permettre de réaliser 20.000 questionnaires financés par le Céreq
- répondre aux problématiques sur l'enquête 2020 et la ré-interrogation de 2023 en anticipant la non-réponse et l'attrition
- tenir compte de certains types de formations sur lesquelles le Céreq souhaite réaliser une analyse fine
- permettre de satisfaire les demandes des partenaires d'extension, en termes de nombre de questionnaires sur un sous-champ pour l'enquête 2020

- permettre de disposer d'une réserve, utilisable totalement ou partiellement, en fonction des taux de réponse observés en cours d'enquête.

Pour satisfaire au mieux à ces objectifs, l'échantillon est composé d'un échantillon principal « tronc commun », afin de répondre aux problématiques Céreq, ainsi que d'échantillons tirés indépendamment pour satisfaire les besoins individuels de chaque extension. La majorité des extensions d'échantillon concerne uniquement l'enquête 2020. L'échantillon de la ré-interrogation sera composé des individus répondants à l'enquête 2020 sélectionnés au sein des échantillons du tronc commun et des extensions reconduites.

▪ Échantillonnage du tronc commun

Le tronc commun est sélectionné par un tirage à probabilités inégales équilibré sur des variables de niveaux de formation plus fins et sociodémographiques au sein de strates de niveaux de formation de sorties.

Calcul des probabilités de tirage

Les probabilités d'inclusion sont définies pour limiter la dispersion des poids après redressement post collecte, en essayant de corriger des défauts de couverture et d'anticiper la non réponse afin que la structure des répondants à l'issue de la phase de collecte soit proche de la structure théorique de la population des sortants de formation initiale. Pour cela, les probabilités de tirage intègrent des coefficients pour corriger d'une part les défauts de couverture de la base de sondage et d'autre part anticiper les forts taux de non réponse de certaines sous populations.

Pour corriger de la sous-couverture de la base de sondage, et mieux représenter les individus issus de types de formations moins couvertes par la base de sondage, la probabilité de tirage d'un individu est inversement proportionnelle à sa probabilité d'être dans la base de sondage. Pour un type d'établissement donné, un taux de couverture au sein de la base de sondage est estimé en utilisant la meilleure information externe disponible. Lorsqu'un effectif d'inscrits de référence est disponible, il est comparé à l'effectif présent dans la base de sondage pour calculer le taux de couverture du type d'établissement concerné. En l'absence d'informations externes, le taux de couverture est estimé par le taux de réponse à la collecte des établissements pour un type donné. Un individu i se voit attribuer un coefficient t_i correspondant au taux de couverture du type du dernier établissement scolaire qu'il a fréquenté afin de représenter sa probabilité d'être dans la base de sondage.

Les taux de réponse des précédentes enquêtes Génération sont très hétérogènes selon les différentes niveaux et types de formations. Afin d'obtenir un nombre suffisamment élevé de répondants de sortants de chaque niveau de formation, les populations ayant la plus faible propension à répondre sont sur-échantillonnées. Pour ce faire, la probabilité d'inclusion à l'échantillon du tronc commun de chaque individu est inversement proportionnelle à sa probabilité de répondre à l'enquête estimée a priori. Ainsi, une probabilité de réponse individuelle, théorique, à la ré-interrogation à 6 ans $p_{rep6,i}$ est calculée à l'aide des observations des comportements de réponses sur les enquêtes précédentes, afin

d'anticiper la non-réponse et l'attrition et d'assurer une représentation de l'ensemble des individus lors de l'enquête 2023 auprès de la Génération 2017.

Une probabilité de réponse à la première interrogation à 3 ans $p_{rep3,i}$ est estimée à l'aide des enquêtes précédentes afin de pouvoir calibrer la taille de l'échantillon, de manière à assurer l'objectif de 20000 répondants de l'échantillon de tronc commun.

Malgré un effort croissant Génération après Génération de détection des individus poursuivant leurs études, une grande partie des individus présents dans la base de sondage se révèlent hors champ lors de la passation de l'enquête. Afin de corriger ce défaut de sur-couverture et de représenter de manière plus homogène tous les individus sortants de formation initiale, les probabilités d'inclusion à l'échantillon du tronc commun de chaque individu sont inversement proportionnelles à la probabilité d'être dans le champ de l'enquête $p_{champ,i}$.

La probabilité d'inclusion d'un individu i à l'échantillon du tronc commun S_0 est alors :

$$\pi_{i,0} = P[i \in S_0] = \frac{c}{t_i \cdot p_{rep6,i}}$$

Où c est un coefficient calibré de manière à obtenir le nombre de réponses souhaitées à la première interrogation en 2020 :

$$\sum_{i \in U} \pi_{i,0} \cdot p_{rep3,i} \cdot p_{champ,i} = 20000$$

$$d'où : c = \frac{20000}{\sum_{i \in U} \frac{p_{champ,i} \cdot p_{rep3,i}}{t_i \cdot p_{rep6,i}}}$$

Plan de sondage et tirage de l'échantillon du tronc commun et de sa réserve

L'échantillon tronc commun, ayant pour but de satisfaire les besoins du Céreq, est stratifié selon les différents niveaux de sortie d'études des individus de la Génération 2017. Chaque strate peut ensuite être subdivisée en niveaux plus fins, qui interviennent dans la phase d'équilibrage.

Tableau 7 - Stratification et niveaux élémentaires envisagés des études

Strate	Niveaux fins
01-Niveaux non diplômants	01A-Sortie Collège
	01B-Sortants SEGPA
	01C-Année non terminale CAP
	01D-Année non terminale BAC PRO-BT-BP
	01E-Année non terminale bac general ou technologique
02-CAP	02A-CAP Tertiaire
	02B-Cap industriel
03-Bac Pro	03A-Bac Pro Tertiaire
	03B-Bac Pro industriel
	03C-BT BP
04-Bac Général	04A-Bac général
	04B-Bac techno
05-Autres niveaux secondaire	05-Autres niveaux secondaire
06- Niveaux non dipomants du supérieur	06A-Niveau non-diplômant STS
	06B-Niveau non-diplômant d'université
07-BTS/DUT	07A-BTS Production
	07B-BTS Services
	07C-DUT Production
	07D-DUT Services
	07E-Autre Bac +2
08-Formations Santé/social	08-Formations Santé/social
09-Licence Pro	09A-Licence Pro LSH
	09B-Licence Pro Sciences techniques
10-Licence Générale	10A-Licence générale LSH
	10B-Licence générale Sciences techniques
	10C-Master 1
11-M2	11A-M2 Droit sciences po
	11B-M2 eco AES
	11C-M2 Art lettres langues SHS
	11D-M2 sciences
	11E-M2 staps
	11F-M2 santé
	11G-M2 Autres
12-Ecoles supérieur	12A-Ecole ingénieur
	12B-Ecole commerce
	12C-Autres écoles bac+5
13-Doctorat	13A-Doctorat santé
	13B-Doctorat hors santé

La stratification a été créée avec l'objectif d'avoir des profils selon les principales variables d'analyse de l'enquête et des comportements de réponse les plus homogènes possible au sein de chaque strate. La stratification se devait cependant de correspondre aux principaux domaines de diffusion de

l'enquête que sont les plus hauts niveaux de diplôme obtenus. Le niveau de sortie est l'information disponible dans la base de sondage qui permet d'anticiper au mieux ces plus hauts niveaux de diplôme. Enfin, les strates devaient contenir suffisamment d'individus pour permettre un équilibrage sur des variables plus fines.

Au sein de chaque strate, l'échantillon sera équilibré sur des variables de sexe de l'individu, région et type du dernier établissement de formation, niveau de sortie plus fin, résidence ou non en QPV (Quartier de la Politique de la Ville) à la sortie des études ainsi que le fait d'avoir ou non réalisé sa dernière formation en apprentissage. Le tirage de l'échantillon principal du tronc commun et de sa réserve sera effectué sous R à l'aide du package sampling, développé par Yves Tillé et Alina Matei. Pour des raisons de durée de traitement, la phase d'atterrissage s'effectuera en relâchant les contraintes d'équilibrage une à une ; le tableau suivant résume les différentes variables d'équilibrage au sein de chaque strate, classées par ordre d'importance.

Tableau 8 - Variables d'équilibrage par strate

Strate	Variables d'équilibrage					
	1	2	3	4	5	6
01-Niveaux non diplômants	Niveau fin	Résidence en QPV	Type d'établissement de sortie	Sexe	Région de l'établissement	
02-CAP	Niveau fin	Formation en apprentissage	Résidence en QPV	Sexe	Type d'établissement de sortie	Région de l'établissement
03-Bac Pro	Niveau fin	Formation en apprentissage	Résidence en QPV	Sexe	Type d'établissement de sortie	Région de l'établissement
04-Bac Général	Niveau fin	Résidence en QPV	Sexe	Type d'établissement de sortie	Région de l'établissement	
05-Autres niveaux secondaire	Type d'établissement de sortie	Formation en apprentissage	Résidence en QPV	Sexe	Région de l'établissement	
06- Niveaux non diplômants du supérieur	Niveau fin	Résidence en QPV	Sexe	Type d'établissement de sortie	Région de l'établissement	
07-BTS/DUT	Niveau fin	Formation en apprentissage	Résidence en QPV	Sexe	Type d'établissement de sortie	Région de l'établissement
08-Formations Santé/social	Type d'établissement de sortie	Sexe	Résidence en QPV	Région de l'établissement		
09-Licence Pro	Niveau fin	Formation en apprentissage	Résidence en QPV	Sexe	Région de l'établissement	
10-Licence Générale	Niveau fin	Formation en apprentissage	Résidence en QPV	Sexe	Type d'établissement de sortie	Région de l'établissement
11-M2	Niveau fin	Formation en apprentissage	Résidence en QPV	Sexe	Région de l'établissement	
12-Ecoles supérieur	Type d'établissement de sortie	Formation en apprentissage	Résidence en QPV	Sexe	Région de l'établissement	
13-Doctorat	Niveau fin	Résidence en QPV	Sexe	Région de l'établissement		

Le tirage s'effectuera en deux temps : dans un premier temps l'ensemble (échantillon principal + réserve) sera tiré au sein de la base de sondage avec les probabilités $\pi_{i,0} \times \frac{11}{10}$ équilibré sur les variables sus-citées au sein de chaque strate afin de disposer d'une réserve de 10% de la taille de l'échantillon.

Dans un second temps, les individus faisant partie de l'échantillon principal seront distingués de la réserve par un tirage équilibré sur les mêmes variables, à probabilités égales à $\frac{10}{11}$ parmi les individus de l'ensemble (échantillon + réserve).

De cette manière, il est assuré que l'échantillon soit équilibré sur les variables souhaitées que la réserve soit utilisée ou non, et les probabilités d'inclusion à l'échantillon principal hors réserve sont $\pi_{i,0}$.

- **Tirage des échantillons d'extension**

Pour chaque extension d'échantillon, un échantillon est tiré indépendamment du tronc commun et des autres extensions afin de satisfaire les objectifs en nombre de questionnaires fixés par les partenaires. Les objectifs des échantillons d'extensions sont les suivants :

Tableau 9 - Objectifs des extensions

Extension	Partenaire	Champ	Niveau	Objec tif
E1	CGET	Jeunes résidant en QPV à la fin de leurs études	Tous	2300
E2	DGESIP	Sortants du supérieur	Supérieur	3500
E3	INJEP	Diplômés de formations du secteur du sport et de l'animation	BPJEPS DEJEPS DESJEPS	1500
E4	ANFA	Sortants de formations automobiles	CAP-Bac Pro	2000
E5	CGDD	Sortants de formation en environnement	CAP	500
			Bac	950
			Bac+1, Bac+2	360
			Bac+3, Bac+4	390
			Bac+5 et plus	700
E6	Région Bretagne	Sortants de formation en région Bretagne	Non diplomant	180
			CAP tertiaire	220
			CAP industriel	220
			Bac Pro tertiaire	180
			Bac Pro techno ou industriel	180
			Bac Général ou techno tertiaire	200
			Bac+2 industriel	220
			Bac+2 tertiaire	220
			Bac+3/4	200
			Bac+5 et plus	180
E7	Région Pays de la Loire	Sortants de formation en région Pays de la Loire	Non diplomant	180
			CAP	310
			Bac	310
			Bac+2	300
			Bac +3/4	200
			Bac+5 et plus	200

Note : l'extension E7 reste à confirmer

Le plan de sondage des extensions E_1 et E_2 seront stratifiés et équilibrés sur les mêmes strates que l'échantillon principal. Afin d'atteindre les objectifs souhaités, les effectifs de la base de sondage des extensions E_3 et E_4 seront échantillonnés exhaustivement. Les échantillons des extensions E_5 , E_6 et E_7 seront stratifiés selon les niveaux d'objectifs ciblés et équilibrés sur les mêmes variables que la strate correspondant au même niveau dans l'échantillon principal

Pour un individu de la base de sondage appartenant au champ d'une extension d'échantillon, le calcul de sa probabilité d'inclusion à l'échantillon de cette extension est fixé proportionnellement à celle d'appartenir à l'échantillon du tronc commun, de manière à atteindre l'effectif de répondants souhaité

pour chaque objectif, et indépendamment du fait qu'il ait tiré ou non dans l'échantillon du tronc commun ou d'une autre extension d'échantillon.

La probabilité d'inclusion d'un individu i appartenant au champ de l'extension d'échantillon E_j pour l'objectif O_k est :

$$\pi_{i,j} = P[i \in S_j] = c_k * \pi_{i,0}$$

$$\text{avec } \sum_{i \in O_k} \pi_{i,jk} \cdot p_{rep,i} \cdot p_{champ,i} = obj_k$$

en notant obj_k le nombre d'individus répondants souhaités pour l'objectif et S_j l'échantillon tiré pour l'extension E_j .

3.4 - Traitements statistiques

- **Estimation composite et partage des poids**

Chacun des échantillons tirés pour l'enquête Génération, tronc commun ou extension, répond à un objectif propre et permet d'obtenir un jeu de poids $\left\{ w_{i,j}^{ech} = \frac{1}{\pi_{i,j}} I[i \in S_j] \right\}_{i \in U}$ permettant d'estimer sans biais leur domaine d'intérêt (tronc commun si $j = 0$, extension E_j pour $j \in \llbracket 1, 7 \rrbracket$) à partir des individus de l'échantillon. Ces différents domaines se recoupent fortement ; en particulier, le tronc commun inclut les domaines d'intérêt de toutes les extensions. Il semblerait dommageable de ne pas inclure tous les individus ayant répondu à l'enquête lors d'estimations sur l'ensemble du champ de l'enquête, et de ne garder que les individus tirés pour le tronc commun. Afin de tirer parti du maximum d'informations disponibles, une estimation composite est réalisée à partir des différents jeux de poids issus des échantillons tirés pour le tronc commun et les extensions. Ainsi, pour estimer le total d'une variable y un sous-domaine D du champ de l'enquête Génération, il s'agit d'agréger les estimations provenant des estimations de ce total de chacun des échantillons, en tenant compte du fait que ces échantillons ne représentent pas tous intégralement D . Par exemple, l'échantillon extension Bretagne ne permet pas d'estimer correctement les agrégats sur la population entière, mais peut améliorer l'estimation par ce qu'elle apporte sur le champ des sortants de la région Bretagne ; elle ne permet cependant pas d'améliorer les estimations sur les individus sortant d'autres régions.

Pour agréger les différentes estimations, il faut préalablement segmenter le champ de l'enquête selon les sous-domaines que les échantillons permettent ou non de représenter, en considérant la partition décrite par l'ensemble des intersections des différents champs des extensions.

Tableau 10 - Partition de la base de sondage de l'enquête auprès de la Génération 2010 selon les intersections d'extensions prévues pour l'interrogation 2020 de la Génération 2017

Partie	E1	E2	E3	E4	E5	E6	E7	Effectif
1	0	0	0	0	0	0	0	478730
2	0	0	0	0	0	0	1	35032
3	0	0	0	0	0	1	0	29017
4	0	0	0	0	1	0	0	10512
5	0	0	0	0	1	0	1	847
6	0	0	0	0	1	1	0	867
7	0	0	0	1	0	0	0	4631
8	0	0	0	1	0	0	1	189
9	0	0	0	1	0	1	0	322
10	0	0	1	0	0	0	0	10031
11	0	0	1	0	0	0	1	895
12	0	1	0	0	0	0	0	429351
13	0	1	0	0	0	0	1	29099
14	0	1	0	0	0	1	0	20576
15	0	1	0	0	1	0	0	13344
16	0	1	0	0	1	0	1	773
17	0	1	0	0	1	1	0	1092
18	0	1	0	1	0	0	0	515
19	0	1	0	1	0	0	1	39
20	0	1	0	1	0	1	0	20
21	1	0	0	0	0	0	0	41316
22	1	0	0	0	0	0	1	1813
23	1	0	0	0	0	1	0	1031
24	1	0	0	0	1	0	0	510
25	1	0	0	0	1	0	1	25
26	1	0	0	0	1	1	0	18
27	1	0	0	1	0	0	0	690
28	1	0	0	1	0	0	1	17
29	1	0	0	1	0	1	0	20
30	1	0	1	0	0	0	0	487
31	1	1	0	0	0	0	0	29187
32	1	1	0	0	0	0	1	1084
33	1	1	0	0	0	1	0	702
34	1	1	0	0	1	0	0	490
35	1	1	0	0	1	0	1	24
36	1	1	0	0	1	1	0	19
37	1	1	0	1	0	0	0	37
38	1	1	0	1	0	0	1	3
39	1	1	0	1	0	1	0	1

Note de lecture La partie n°6 est composée des individus appartenant aux sous-populations E5 et E6 sans appartenir à aucune autre sous-population. Dans la base de sondage de l'enquête auprès de la Génération 2010, ils étaient 867 individus au sein de la partie 6.

Pour chaque élément p de cette partition, il s'agit alors d'estimer sans biais le total de y sur p à partir des différents échantillons et de pondérer ces estimations à l'aide d'un jeu de coefficients $a_{p,j}$. L'estimateur composite du total de Y sur p est alors :

$$\hat{Y}_{comp,p} = \sum_{j \in \llbracket 0,7 \rrbracket} \left(a_{p,j} \cdot \sum_{i \in S_j} w_{i,j}^{ech} \cdot y_i \right).$$

$$\text{Sous condition : } \sum_{j \in \llbracket 0,7 \rrbracket} a_{p,j} = 1$$

L'estimateur du total de Y sur l'ensemble de la population est alors :

$$\hat{Y}_{tot} = \sum_{p \in P} \hat{Y}_{comp,p}$$

Au niveau individuel, cela se traduit par un partage des poids. Pour chaque individu i d'un sous-domaine p , son poids de l'estimateur composite est alors

$$w_i^{comp} = \sum_{j \in \llbracket 0,7 \rrbracket} a_{p,j} \cdot w_{i,j}^{ech}$$

Pour chaque sous-domaine p de la partition P , le choix des coefficients $a_{p,j}$ dépend des extensions constituant p . Si p contient l'extension E_3 ou E_4 (qui sont exhaustives), alors $a_{p,3} = 1$ ou $a_{p,4} = 1$ et les autres coefficients sont nuls ; E_3 ou E_4 garantit que tous les individus de p sont au sein de l'échantillon.

Sinon, $a_{p,j}$ est égal à la part d'individus de p répondants provenant de l'extension E_j , de telle sorte que l'estimation du total à partir d'un échantillon (tronc commun ou extension) fournissant un grand nombre de répondants de p ait plus de poids que celle d'un échantillon en fournissant moins. Le poids de l'estimation composite d'un individu sera alors la moyenne pondérée par le nombre de répondants des poids d'échantillonnage de chaque échantillon (avec un poids nul si l'individu n'est pas échantillonné). De cette manière, les moyennes des $a_{p,j} \cdot w_{i,j}^{ech}$ par échantillon E_j seront a priori proches et les poids moyens des individus échantillonnés une unique fois seront en moyenne proches, indépendamment de l'échantillon initial dans lequel ils étaient tirés.

▪ Traitement de la non-réponse totale

Une correction de la non-réponse totale est envisagée par repondération par l'inverse de la probabilité de réponse estimée $p_{red,i}$. L'estimation sera réalisée par la méthode des groupes homogènes de réponse suite à une estimation par régression logistique, de manière analogue aux enquêtes précédentes.

Lors de la précédente enquête, la modélisation était décomposée en deux étapes, afin d'affiner l'estimation en intégrant des données relatives au processus de collecte. Une première étape était d'estimer la probabilité de contacter les individus échantillonnés ou un de leur proches à l'aide des variables présentes dans la base de sondage ainsi que du processus de contact.

La deuxième étape modélisait la probabilité d'accepter de répondre à l'enquête sachant que l'individu a été contacté et est identifié dans le champ de l'enquête. Dissocier cette modélisation de la première

permet notamment d'intégrer des variables relatives à la phase de contact du questionnaire : nature du contact, prise de rendez-vous...

Pour exemple, sur la première enquête auprès de la Génération 2013, les variables utilisées dans les modèles étaient :

- Niveau et type de formation
- Âge de l'individu (en 2013)
- Variables indicatrices sur le type de coordonnées téléphoniques disponibles dans l'échantillon (numéro de téléphone fixe, portable, issu de la base de sondage ou de recherches de coordonnées téléphoniques,...)
- Mode d'envoi de la lettre avis
- Formation effectuée par apprentissage en 2013
- Caractéristiques de la commune de résidence en 2013
- Département de l'établissement de formation en 2013
- Genre

Pour l'enquête 2020 auprès de la Génération 2017, pour inclure la dimension multimode de cette enquête, des variables renseignant le type de contact utilisé, la phase de l'enquête où le premier contact a été effectué, ou les différents types de relances utilisées pour contacter l'individu pourront être utilisées dans la modélisation en addition aux variables sus-citées.

▪ Calage sur marges et pondération finale

Les Repères et références statistiques (RERS) du ministère de l'Éducation nationale publient un tableau donnant une estimation de la répartition par sexe et niveau de diplôme des sortants du système éducatif français. Les poids des individus répondant à l'enquête 2020 auprès de la Génération 2017 seront calés sur la répartition publiée dans le RERS. Les poids de l'enquête 2023 seront calés sur les principaux agrégats de l'enquête 2020, pour assurer une cohérence des indicateurs entre les deux enquêtes. Pour exemple, les indicateurs ayant servi au calage de l'enquête 2017 auprès de la Génération 2010 sont :

- Région d'habitant en 2013 (regroupées en nouvelles régions) : *CA0_REG13*
- Région d'habitat en 2015 (regroupées en nouvelles régions) : *CA0_REG15*
- Situation professionnelle en Avril 2013 : *MOIS42*
- Situation professionnelle en Mars 2015 : *MOIS65*
- Mode de cohabitation à la date d'enquête en 2013 : *HABDE_13*
- Mode de cohabitation à la date d'enquête en 2015 : *HABDE_15*
- Lieu de naissance (France ou étranger) : *LIEUNAIS*
- Sexe : *Q1*
- Age en 2010 : *AGE10*
- Sexe*Plus Haut Diplôme (*Q1*PHD*)

Les poids finaux des individus répondants de l'enquête Génération sont alors

$$w_i^{tot} = \frac{w_i^{comp}}{p_{red,i}} \cdot cal_i$$

Où cal_i est le coefficient de calage.

▪ **Traitement de la non-réponse partielle : les salaires**

Durant le questionnaire, pour chaque emploi déclaré par l'individu, le salaire au début et à la fin de la séquence d'emploi est demandé. L'information sur le salaire étant sensible, il est laissé l'opportunité aux individus ne souhaitant pas livrer le montant exact de leur salaire de ne pas répondre ou de donner un intervalle dans lequel il est compris (déclaration de tranches de salaires). Cette information est cependant un des principaux paramètres d'intérêt et objets d'études des enquêtes Génération ; les salaires manquants sont imputés afin d'avoir des données sur chacun des emplois renseignés.

Le traitement de la non-réponse des salaires est envisagé de manière similaire aux enquêtes précédentes : il s'effectue par imputation par la régression linéaire. Les salaires déclarés en début et fin de séquence emploi sont modélisés séparément. Les salaires dont les résidus studentisés de ces modèles sont supérieurs à 3 sont jugés aberrants et mis à blanc. Les informations non renseignées sont alors imputées de manière déterministe par la projection de ce modèle.

Une modélisation similaire est effectuée pour l'imputation des salaires déclarés en tranche, à l'aide des salaires déclarés au sein de l'intervalle de salaire correspondant à la tranche.

Pour exemple, les variables permettant la modélisation des salaires pour la Génération 2013 ont été :

- Sexe
- Âge de l'enquêté
- Niveau de sortie : non diplômé, secondaire, bac+2, bac+3/4, bac+5
- Plus haut diplôme obtenu en 15 positions
- Spécialité de formation : général, industriel ou tertiaire
- Type de contrat de travail : indépendant, fonctionnaire, cdi, cdd, contrats aidés
- Catégorie socio-professionnelle : ouvrier, profession intermédiaire, cadre, employé, autre
- Ancienneté
- Région de l'entreprise : Île-de-France, autre région, étranger
- Taille de l'entreprise : moins de 20 salariés
- Activité de l'entreprise : NAF en 8 postes

▪ **Traitement du biais de mesure entre les différents modes**

La collecte multimode implique des effets de mesure dans les réponses des individus. De fait, un même individu ne répond pas forcément de la même manière selon le mode d'interrogation. Les causes de ces différences peuvent être multiples : en présence d'un enquêteur, un individu se sent plus concerné et se concentre plus pour trouver la réponse la plus adéquate à une question difficile et peut être aidé en cas d'incompréhension de la question. Lorsqu'au contraire, lors d'un mode de collecte auto-

administré, l'enquêté se contente de donner une réponse qu'il juge bonne au lieu de chercher la meilleure réponse qu'il aurait pu donner ; cet effet est dénommé *satisficing*.

Entre les collectes Internet et téléphone, les informations ne sont pas reçues identiquement : dans le cadre de la collecte par téléphone, l'enquêté doit se rappeler de la liste des modalités, alors que dans la passation en auto administré, l'enquêté lit les modalités qui apparaissent toutes ensemble sur son écran. À partir de cette différence de transmission de l'information, deux effets de mesure opposés ont été définis. Dans une collecte par téléphone, on parle de *recency effect* : l'enquêté donnera davantage une réponse parmi les modalités dont il se souvient, c'est-à-dire une des dernières modalités entendues. À l'inverse, dans le cas de la collecte par Internet, les répondants auraient tendance à choisir une des premières modalités lues. Il s'agit, alors d'un effet de primauté (*primacy effect*).

Dans le cas d'un mode de collecte où un enquêteur est présent, l'enquêté peut craindre le jugement de l'enquêteur et donner des réponses socialement plus acceptables à des questions d'opinion ou sensibles. Cet effet est appelé biais de *désirabilité sociale*.

Le biais impliqué par les effets de mesure sur les estimations est problématique : les estimations des variables sujettes à un effet de mesure réalisées sur les populations de répondants diffèrent selon le mode de réponse à structure de population égale. La vraie valeur du paramètre n'est sûrement pas égale à l'une des estimations, effectuée sur un seul mode ; l'estimation de ce paramètre sur l'ensemble de la population dépend de la répartition des individus selon le mode de réponse.

Afin de tenir compte de ce biais de mesure dans l'enquête auprès de la Génération 2017, une quantification de l'effet de mesure sur chacune des variables du questionnaire sera réalisée par un matching sur score de propension. Usuellement utilisées dans le cadre de l'évaluation des politiques publiques, le but des méthodes de matching est d'évaluer l'effet d'un traitement T sur une variable d'intérêt Y. Dans le contexte de l'enquête auprès de la Génération 2017, le traitement est le fait de répondre au questionnaire par Internet, par rapport à une réponse par téléphone ; les variables d'intérêt seront des variables du questionnaire susceptibles de présenter un effet de mesure, par exemple des variables portant sur l'opinion sur l'emploi des individus, les raisons de son arrêt des études, ou sur le sentiment de discrimination au sein de l'entreprise.

La population suivant le traitement a une variable de traitement $T=1$ tandis que celle ne le subissant pas a pour valeur $T=0$. Dans ce cadre, deux variables latentes Y_1 et Y_0 correspondent aux valeurs de Y que l'individu aurait renseigné en répondant, respectivement, par Internet ou par téléphone. Une seule de ces variables latentes est observée, mais la relation suivante existe :

$$Y = T.Y_1 + (1 - T).Y_0$$

L'hypothèse de cette approche par matching est de considérer que, conditionnellement aux variables qui expliquent la sélection, les variables latentes sont indépendantes du traitement. C'est-à-dire que conditionnellement aux covariables, il n'y a plus d'effet de sélection et l'assignation du traitement peut être considérée comme aléatoire.

L'objectif est d'alors d'estimer la valeur latente non observée. Pour cela, à chaque individu qui a reçu le traitement est attribué un contrefactuel qui n'a pas reçu le traitement et qui lui ressemble selon les variables qui expliquent la sélection. Dans le contexte de l'enquête auprès de la Génération 2017, la détermination d'un contrefactuel de chaque individu sera par la méthode d'appariement sur score de

propension ; dans le cas d'indépendance des variables latentes Y1 et Y0 par rapport aux covariables expliquant la sélection, alors il y a aussi indépendance par rapport à la probabilité de recevoir le traitement, le score de propension. Ce score de propension sera estimé par modélisation logistique de la probabilité de répondre par Internet plutôt que par téléphone à partir de variables non sujettes à effet de mesure.

La valeur de la variable latente inobservée Y0 est alors estimée par la variable Y du contrefactuel. Réciproquement, la valeur de la variable latente inobservée Y0 des individus qui n'ont pas suivi le traitement peut être estimée par la valeur Y de leur contrefactuel qui a suivi le traitement.

L'effet de mesure est estimé par l'effet moyen du traitement (Average Treatment effect on the Treated ou ATT) :

$$ATT = E(Y1 - Y0|X, T = 1)$$

Les variables sujettes à un fort effet de mesure seront renseignées dans le dictionnaire des variables de l'enquête afin que chaque personne utilisant les données de l'enquête soit sensibilisée au problème de mesure.

Une correction de l'effet de mesure pourra être apportée dans un second temps pour certaines variables importantes présentant un effet de mesure. Une méthode⁷ de correction de l'effet de mesure par imputation a été proposée par S.Legleye dans le cadre du groupe de travail de l'INSEE concernant les enquêtes multimode et pourrait être employée dans le cadre de l'enquête auprès de la Génération 2017. Elle repose sur la mise à blanc des variables présentant un effet de mesure pour une partie de la population des répondants d'un mode alternatif puis de l'imputation de ces variables selon une modélisation effectuée sur le mode de référence afin d'annuler l'effet de mesure. Un jeu de variables imputées sera alors livré aux chargés d'étude en complément des données brutes.

7 Legleye S., Razafindranovona T., De Peretti G., "AGREGER LES ÉCHANTILLONS D'UNE ENQUÊTE MULTIMODE EN LIMITANT L'EFFET DE MESURE: UNE PROPOSITION D'IMPUTATION RAISONNABLE ET PRAGMATIQUE". *XIIIèmes Journées de Méthodologie Statistique*, 2018.

4. Description du mode de collecte

4.1 - Protocole de collecte

- **Une collecte organisée en 3 phases**

Le protocole d'enquête multimode internet-téléphone prévoit la mise en place d'un plateau téléphonique sur 4 mois, d'avril à juillet 2020. Durant l'entièreté de la période d'enquête, les individus échantillonnés pourront contacter le Céreq via un numéro gratuit ou par le biais d'une adresse mail afin de répondre à leurs interrogations à propos de l'enquête, de les assister à la réalisation de l'enquête par internet ou de réaliser l'enquête par téléphone. Seul l'individu échantillonné pourra répondre à l'enquête, un parent ou un proche ne pourra servir de proxy qu'en cas d'incapacité physique de l'individu à répondre. Le protocole permet le changement de mode de collecte pour un même enquêté : un individu ayant commencé à répondre par internet peut terminer le questionnaire par téléphone, et inversement.

Le protocole de collecte est construit autour de trois phases distinctes dans le temps et privilégie une collecte internet.

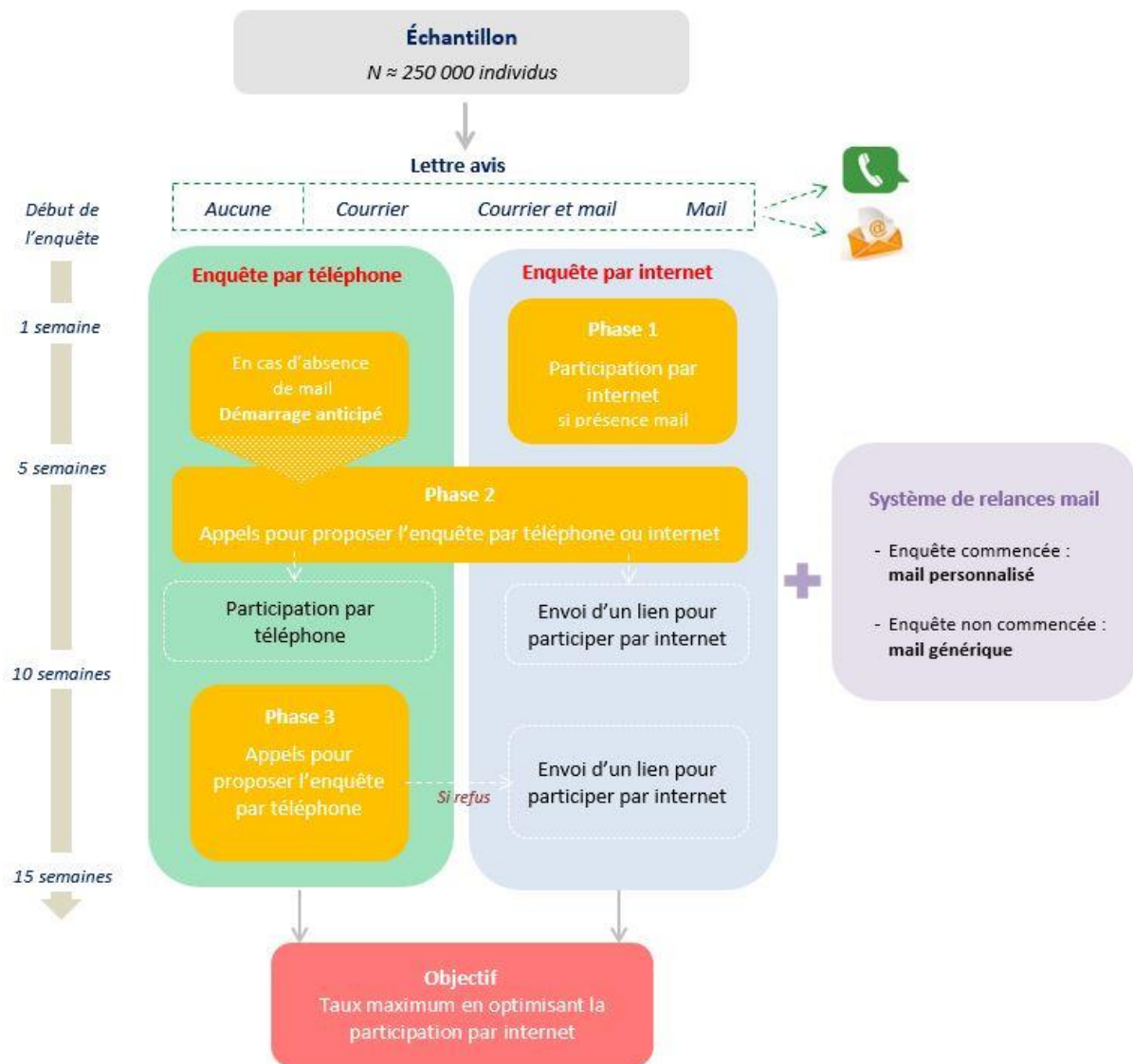


Figure 1 - Logigramme du protocole de l'enquête multimode

Phase 1 - Incitation à répondre par internet

La première phase, d'une durée de quatre semaines, débute par l'envoi d'un mail avis ou d'une lettre-avis papier destinée à informer du démarrage de l'enquête. Le mail comme la lettre incluent un lien personnalisé pour participer à l'enquête par internet. Le protocole diffère selon la disponibilité ou non d'une adresse mail valide et/ou d'une adresse postale pour l'individu.

Présence d'un mail valide : Un premier mail avis est envoyé au démarrage de l'enquête, il permet à la fois d'informer les individus et de vérifier la validité de l'adresse mail. Par la suite, des mails de relance seront envoyés hebdomadairement en variant les jours et heures d'envoi. Pour certaines sous-populations (individus faisant partie d'une extension par exemple), il pourra être envisagé d'envoyer une lettre-avis papier en complément du mail-avis.

Absence de mail valide et présence d'une adresse postale : L'ensemble des individus ne possédant pas d'adresse mail valide mais possédant une adresse postale recevra une lettre-avis.

Absence de mail valide et d'adresse postale : ces individus seront contactés par téléphone durant les trois premières semaines de l'enquête. Il s'agira de qualifier les individus, c'est-à-dire de s'assurer d'avoir joint la bonne personne et qu'elle appartient au champ de l'enquête. Si cette confirmation est réalisée, le choix de continuer l'enquête par téléphone ou par internet sera proposé. Si le choix se porte sur internet, une adresse électronique sera demandée pour l'envoi immédiat d'un mail contenant un lien pour répondre par internet. Si la personne enquêtée n'a pas de mail ou souhaite répondre par téléphone, la passation se fera par téléphone.

Phase 2 - Choix de répondre par internet ou téléphone

La seconde phase, d'une durée de cinq semaines, consiste à contacter par téléphone l'ensemble des individus n'ayant pas encore répondu afin de leur donner le choix de participer soit par internet soit par téléphone selon leur préférence. Cette phase se caractérise par une alternance entre les relances par mail et les tentatives d'appels téléphoniques. Un envoi de mail par semaine est prévu ; l'ensemble de l'échantillon sera contacté par téléphone au moins une fois pendant cette phase.

Phase 3 – Priorisation des réponses par téléphone

La dernière phase de collecte, d'une durée de cinq semaines, est prévue pour proposer aux enquêtés de participer par téléphone. Le plan de dialogue de l'enquêteur sera adapté pour proposer uniquement de répondre par téléphone. En cas de demande spontanée de l'enquêté pour répondre par internet, l'enquêteur devra tout de même inciter à répondre par téléphone. En cas de refus de réponse par téléphone, le lien du questionnaire sur internet sera envoyé après confirmation de l'adresse mail de l'enquêté.

▪ **Les relances multimode**

Des relances seront réalisées par mail, par téléphone, par message laissé sur répondeur et/ou par SMS. En parallèle des relances téléphoniques propres à chaque phase, des relances mails hebdomadaires seront programmées. La relance peut également se faire ponctuellement par SMS.

Les contacts téléphoniques et messages sur le répondeur

Pour chaque individu du fichier de l'échantillon, il est prévu 20 tentatives d'appel par numéro de téléphone disponible. Un message sera laissé sur le répondeur des individus ne répondant pas en milieu d'enquête. Le message présentera l'enquête Génération en quelques mots et indiquera le numéro gratuit afin que l'individu puisse rappeler et prendre rendez-vous ou encore réaliser l'enquête.

Les relances mails

Deux types de relances par mail seront effectuées :

Les relances mail **planifiées** seront effectuées auprès des individus qui n'ont pas commencé le questionnaire (qu'ils aient été contactés par mail ou par téléphone). Ces relances pourront s'étaler sur l'ensemble de la période de collecte à raison d'un mail par semaine et seront effectuées à des horaires et des jours de la semaine différents.

Des relances mails **pour terminer l'enquête** seront envoyées aux individus qui ont commencé l'enquête sans la terminer à J+4, J+8... (le rythme précis reste à définir avec notre prestataire).

Les relances par SMS

Une relance par SMS sera faite aux individus disposant d'un numéro de téléphone portable et qui n'ont pas encore répondu à l'enquête au moment du passage dans la deuxième phase de collecte. Ce message contiendra un lien pour se connecter directement au questionnaire en ligne ainsi que le numéro gratuit pour répondre par téléphone.

4.2 - Sécurisation des données et protection du secret statistique

Il est à noter que l'une des membres de l'équipe EIGE du Céreq est également DPO (Déléguée à la protection des données). Elle veille donc au respect de la sécurisation des données dans toutes les phases de l'enquête.

▪ **Sécurité des données personnelles au Céreq**

Les données personnelles (contenues dans la base de sondage, dans l'échantillon et dans le fichier des répondants) sont stockées dans un répertoire auquel seuls les membres de l'équipe EIGE ont accès.

Comme le stipule l'article 32 du RGPD, des mesures techniques et organisationnelles appropriées doivent être mises en œuvre afin de garantir un niveau de sécurité adapté au risque. Les mesures suivantes sont ainsi mises en place au Céreq :

- **Sécurité physique** (serveurs, postes, réseaux, moyens relatifs au contrôle de l'accès aux locaux)

Les serveurs du Céreq sont situés dans une salle fermée accessible par badge. L'accès à cette salle est restreint au personnel du service technique habilité.

Chaque visiteur dans les locaux du Céreq doit se présenter à l'accueil au troisième étage du bâtiment. Les trois autres étages (premier, deuxième et quatrième) sont accessibles soit, à pied par un passage systématique au troisième étage, soit par ascenseur avec l'utilisation d'un badge personnel. Les ascenseurs sont bloqués dans le sens montant à partir de 18h (ils ne sont jamais bloqués pour descendre). Les locaux ferment et une alarme de sécurité est enclenchée à partir de 21h. En cas de déclenchement, un service de sécurité intervient immédiatement.

De plus, le bâtiment, dans lequel se trouve les locaux du Céreq est surveillé par des agents de sécurité jour et nuit.

- **Sécurité logistique et physique des traitements** (moyens relatifs à la sécurité du système d'information) :

Le Céreq met à disposition de chaque agent un PC portable dont le contenu du disque dur est chiffré. Chaque utilisateur détient des droits qui restreignent son accès à un certain type de données. Ces

droits sont gérés par le service informatique. Chaque ordinateur est protégé par un câble renforcé et sécurisé par code d'accès.

Chaque PC est accessible via un compte personnel, géré par un identifiant et un mot de passe actuellement sur 12 positions alphanumériques, renouvelable obligatoirement tous les 3 mois.

- **Liste des outils destinés à sensibiliser les acteurs à cette problématique**

Les agents sont soumis à la charte informatique du Céreq qui liste les droits et devoirs des agents face à l'usage des outils informatiques. Elle est annexée au règlement intérieur. Une nouvelle version de cette charte est actuellement en cours de rédaction dans le cadre de l'instauration du télétravail au Céreq et dans le cadre de modifications dans la gestion informatique de la traçabilité des accès externes au Céreq. Cette charte imposera également un code de conduite pour le traitement des données (article 40 du RGPD).

La DPO mène des actions de sensibilisation des agents au RGPD. Elle intervient en comité de direction pour former les managers, publie régulièrement des communiqués sur l'intranet et accroche des affiches dans les locaux afin d'informer les agents sur le RGPD et réalisent avec l'aide du service informatique des formations sur la sécurité.

L'intranet constitue aussi un moyen de sensibiliser les agents sur leurs obligations.

- **Mesures prises pour auditer le niveau de sécurité et recommandations effectuées**

L'installation récente de Netwrix, outil de traçabilité des accès permet d'auditer en temps réel le niveau de sécurité (cryptolockers, accès non désirés, comptes administrateurs périmés, etc.)

- **Procédure mise en place pour le transfert des fichiers**

Les transferts de fichiers de données à caractère personnel (entrants ou sortants) sont à minima compressés et protégés par mot de passe. L'envoi de ces fichiers est réalisé via un des moyens de transfert sécurisé suivants : par mail, via la Céreq box (cloud de fichiers), via filesender (outil de transfert sécurisé fourni par Renater) ou tout autre outil permettant de maîtriser la sécurité.

Tous les accès aux outils de transfert de fichiers nécessitent une authentification par login et mot de passe. La totalité des échanges de données sont faits par un canal chiffré (https) permettant de se prémunir des attaques. Chaque mot de passe est transmis par téléphone. Une fois réceptionnés, les fichiers sont enregistrés dans un dossier à accès restreint aux seules personnes habilitées à y accéder. Un code de conduite sera imposé pour le traitement de ces données (article 40).

- **Encadrement des travaux réalisés par les prestataires**

Les prestataires sont choisis sur marché d'appel d'offre. Les clauses administratives listent des obligations citées sur le CCAP :

- Une obligation de transparence et de traçabilité
- La prise en compte des principes de protection des données
- Une obligation de garantir la sécurité des données traitées
- Une obligation d'assistance, d'alerte et de conseil

▪ Sécurité des données personnelles lors de la collecte

Lors de la réponse à l'appel d'offre portant sur la réalisation de la collecte, le prestataire a su démontrer sa capacité à protéger les données à caractère personnel (DCP). En effet, le DPO de Médiamétrie mène des analyses d'impact, tient un registre des traitements, prévoit des actions de sensibilisation des collaborateurs aux enjeux du traitement des DCP, notamment avec la mise en place d'une politique de confidentialité engageant tous les agents. Il mène également une démarche « Privacy by design » et « Security by design » avec l'aide de son Responsable de la sécurité des systèmes d'information (RSSI) et a mis en place une procédure en cas de violations de données.

Le prestataire assure que le Céreq bénéficie de l'ensemble de leurs mesures techniques d'ores et déjà mises en place et prévoit des dispositions spécifiques à notre enquête. Ainsi, les DCP échangées dans le cadre de l'étude le seront via une gateway SFTP sur laquelle le Céreq disposera d'un compte dédié. De plus, l'accès aux questionnaires internet sera toujours réalisé grâce à des liens *https*. Enfin, les clés de session qui permettront d'accéder au questionnaire électronique seront construites de façon aléatoire et composées de plus de 50 caractères. Les serveurs qui contiendront les données relatives à l'enquête « Génération » seront hébergées dans les Datacenters où le prestataire bénéficie d'un espace strictement dédié. Ces Datacenters sont situés en Île de France.

Chez Médiamétrie, la sécurité des informations est encadrée par une Politique de Sécurité de l'Information qui vise à garantir à leurs clients, collaborateurs, panélistes et interviewés (anciens, actuels et futurs), la protection de tous les éléments concourant à leurs études ainsi qu'au bon fonctionnement de leur entreprise. De cette Politique de Sécurité de l'Information découlent une Politique de Sécurité du Système d'Information et une Charte Informatique. Celle-ci est annexée au règlement intérieur de l'entreprise. L'ensemble de ces dispositions donnent lieu à la mise en place des mesures concrètes suivantes :

- Sauvegarde des données : sauvegarde quotidienne sur six jours, une hebdomadaire le septième jour sur quatre semaines et une sauvegarde mensuelle.
- Suivi : La protection et l'audit des données sensibles est assurée par les outils Varonis, permettant un audit sur les mouvements de fichiers (ouverture / déplacement / copie / modification), ainsi que la gestion suivie des autorisations d'accès.
- CERT : Le prestataire est accompagné au quotidien par un CERT (Computer Emergency Response Team) permettant d'assurer une veille sécurité et de gérer les incidents éventuels.
- Politique de mots de passe : les mots de passes doivent être de minimum 10 caractères alphanumériques et doivent être renouvelés tous les six mois.
- Maintien en condition opérationnelle : Gestion et Supervision 24/7 sur l'ensemble du Système d'information et astreinte 24/7
- Protection des postes de travail : les disques durs des ordinateurs portables sont tous systématiquement chiffrés.

- Datacenters : les infrastructures CAWI sont hébergées sur deux Datacenters en clusters aux dernières normes ISO et certifiées Tier 3 minimum avec un temps de disponibilité supérieur à 99,99999 %. Les interconnexions entre les sites sont toutes doublées via deux opérateurs différents avec une gestion de flux Niveau 3 entièrement gérée et monitorée par les équipes du prestataire. La haute disponibilité du réseau est assurée par une redondance des équipements et une configuration en routage dynamique. Les Datacenters bénéficient de contrôles d'accès physiques stricts, réalisés via un dispositif biométrique complété par un badge d'accès nominatif.
- Askia – Chiffrement des données : L'outil Askia qui sera utilisé pour la réalisation de l'enquête « Génération » bénéficie d'une fonctionnalité de chiffrement des informations qui pourra être enclenchée au moment du recueil de l'information.
- Dispositifs techniques de protection : Des firewalls réseau et applicatifs sont en place pour contribuer à la protection du système d'information de l'entreprise. Ces dispositifs sont complétés par l'existence d'une solution d'authentification forte pour les accès à distance (tokens) et d'une solution d'analyse et de maîtrise des flux Internet sortants (Zscaler).
- Infrastructures de protection physique : La protection des données passe également par la mise en place de dispositifs de protection des locaux de l'entreprise contrôle d'accès sécurisé par badge, système de vidéo surveillance, agent de sécurité qui contrôle les entrées des enquêteurs pendant la période de vacation, système d'alarme anti-intrusion, droits d'accès restreints dans les zones sensibles (accès par digicode et badge), système de destruction des documents papiers par bennes, sécurisées et broyeurs, système de sécurité incendie.

▪ Conformité RGPD au Céreq

L'enquête Génération a été enregistrée dans le registre des traitements tenu par la DPO du Céreq et son adjointe. Une copie de cet enregistrement est disponible en annexe 6.

Afin de garantir la sécurité des données à caractère personnel lors de la collecte, la DPO du Céreq et son adjointe ont travaillé en collaboration avec le DPO de la société prestataire et leur juriste. Une analyse d'impact (PIA) a ainsi été réalisée via l'outil PIA mis à disposition par la CNIL.

Nous avons mis en évidence une gravité importante pour les individus concernés en cas d'accès illégitime aux données. Cependant au vu des mesures mises en place, la vraisemblance de cet événement est négligeable. Dans le cas où cet événement aurait lieu, nous lancerions la procédure de violation des données à caractère personnel à la fois au Céreq et auprès de notre prestataire. Pour ce qui est de la modification non désirée des données et de la disparition des données, la gravité est négligeable pour l'individu et la vraisemblance est limitée.

Le résultat de cette PIA est présenté ci-dessous d'une part au travers de la vue d'ensemble des risques et des mesures associées pour y remédier (Figure 1) et d'autre part grâce à la cartographie des risques (Figure 2). L'analyse d'impact complète est disponible en annexe 7.

Figure 1 - Vue d'ensemble des risques et mesures associées

Impacts potentiels

Usurpation d'identité
Sollicitation marketing non désirée
Opération frauduleuse (type phishing)
Divulgaration d'informations confidentielles
Chantage ou Dénonciation (ex : travail non déclaré)
Discrimination à l'embauche
Discrimination à l'accès au logement ou d'un autre service
Atteinte à l'image d'une personnalité publique
Aucun impact pour l'individu

Menaces

Introduction frauduleuse dans le système par un tiers extérieur
Introduction frauduleuse dans le système par un employé non autorisé
Copie et extraction automatisées des DCP par un logiciel
Copie et extraction manuelles des DCP par un employé malveillant
Absence de sécurisation
Erreur de programmation
Bug Logiciel
Malveillance interne
Erreur ou négligence
Erreur de manipulation des données
Malveillance externe
Dysfonctionnement du système d'information

Sources

Tiers malintentionné s'introduisant dans le système d'information
Employé malveillant
Organisation illégale volant des données à grande échelle automatique
Humain
Source logique

Mesures

Politique de sécurité des systèmes d'informations - Charte
Politique de gestion des DCP
Sensibilisation des collaborateurs
Privacy by design
Chiffrement des flux de données
Gestion des droits d'accès aux données
Traçabilité des accès
Hébergement sécurisé des données
Sécurité du code source
Protection internet
Cloisonnement des informations
Protections des locaux
Gestion des identifications
Traçabilité, suivi et audit des accès aux données
Protection globale des infrastructures informatiques
Gestion et supervision de l'ensemble du système d'information
Sauvegarde régulière des données
Veille et alerte des informations de traçabilité

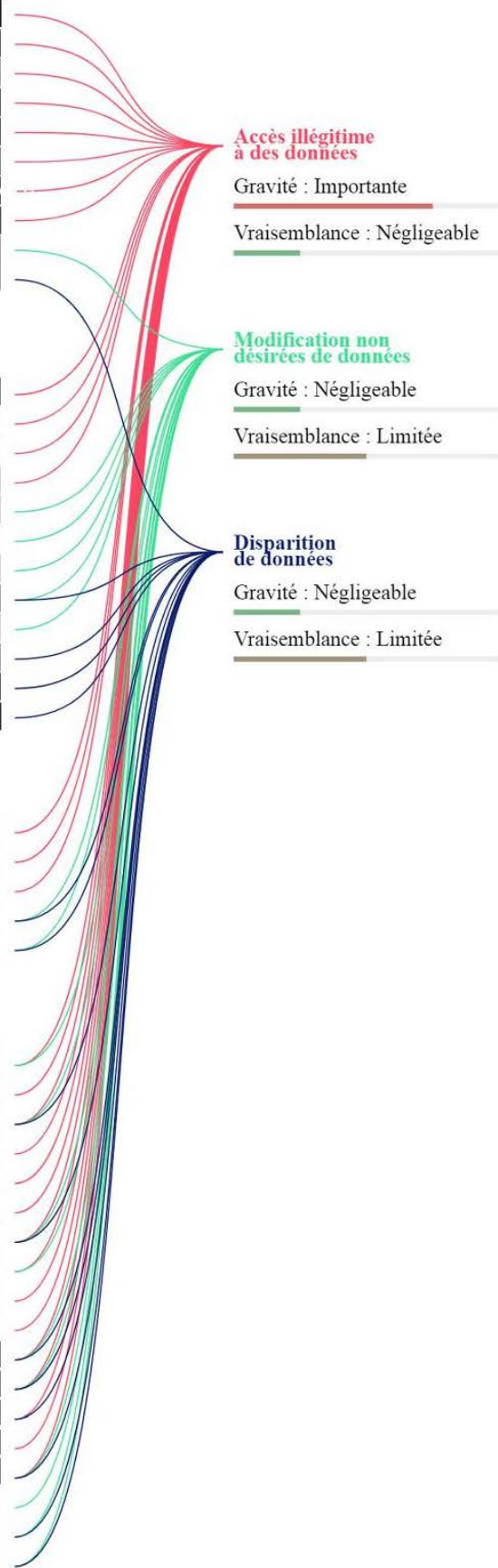
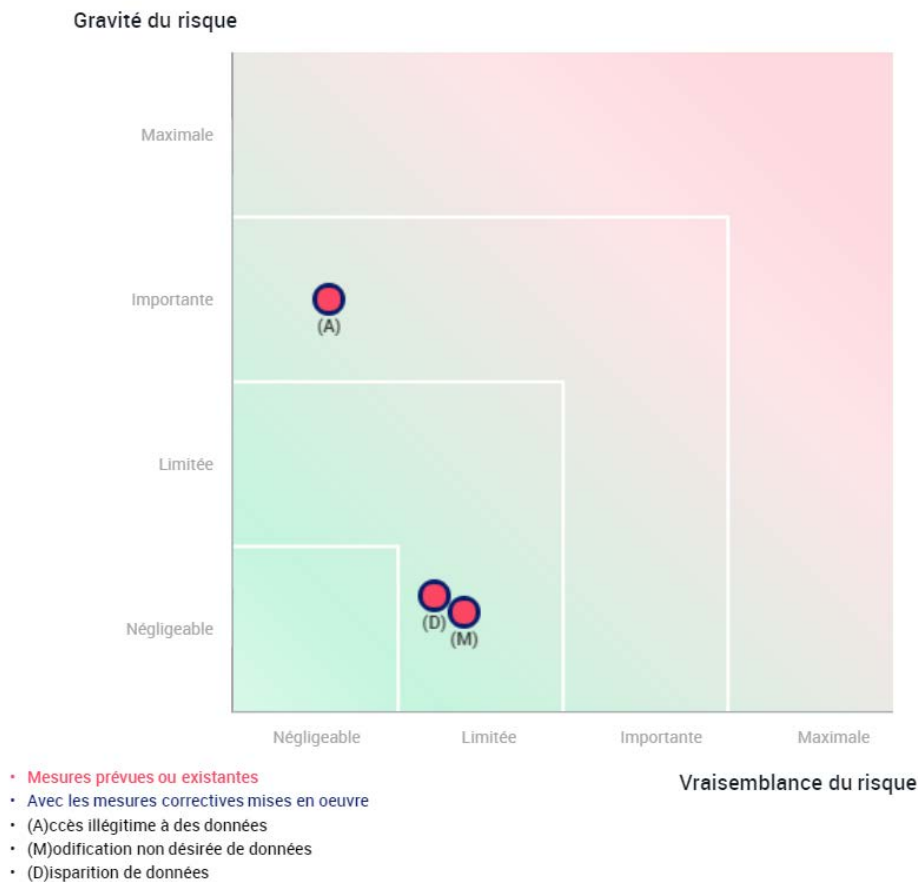


Figure 2 : Cartographie des risques



10/04/2019

▪ **Traitements des données**

Les données des enquêtes Générations sont traitées par le logiciel *Statistical Analysis System (SAS)*. Ce logiciel permet la création, la gestion de bases de données de très grande taille, le traitement analytique de bases de données, la création et la diffusion de rapports. Ce logiciel est utilisable uniquement via le réseau du Céreq et les bases traitées doivent être obligatoirement stockées sur les serveurs sécurisés du Céreq. Ainsi les agents traitant ces bases ne peuvent les copier sur leur ordinateur en local. Cela est une sécurité supplémentaire pour le traitement des bases de données.

▪ **Anonymisation des données d'enquête**

Les bases de données mises à disposition des chargés d'études suivent une procédure d'anonymisation. Les variables désignant directement un individu sont supprimées (nom, prénom,

adresse, numéro de téléphone...). Un identifiant permet uniquement à l'équipe en charge de la production de données de rapprocher l'identifiant à un individu « physique » (identifiant nécessaire pour nos enquêtes longitudinales). Une étude sera réalisée pour déterminer si le croisement de plusieurs variables ne permet pas d'identifier de manière indirecte un individu. Ainsi, les informations indirectement nominatives ou trop fines seront supprimées ou agrégées des bases d'études mises à disposition des chargés d'études du Céreq et du Centre Quételet.

Dans le cadre exceptionnel de travaux d'approfondissement par des chargés d'études, où un accès à des données nominatives est demandé (ré-interrogations en post-enquête), un engagement moral listant les grands principes de confidentialité des données est signé. Ces procédures reçoivent la validation du Correspondant informatique et libertés du Céreq.

Plusieurs prestataires sont amenés à travailler sur des données nominatives en plus du prestataire de collecte : enrichissement des coordonnées téléphoniques, RNVP (restructuration, normalisation et validation postale) et rachat des déménagés, routage des lettres-avis papier. Ces prestataires sont choisis par appel d'offre et sont donc obligatoirement soumis à des obligations de confidentialité mentionnées dans les clauses administratives. Les transferts de fichiers entre le Céreq et les prestataires se font par FTP, avec un cryptage des fichiers avec mot de passe.

▪ **Gestion des demandes de droits des individus**

L'adresse dpo_generation@cereq.fr a été créée dans le but de gérer les demandes de droits des individus. Le Céreq répondra à ces demandes et fera intervenir le DPO du prestataire si nécessaire.

4.3 - Lettres-avis définitives, notices ou plaquettes d'informations

Les individus échantillonnés seront informés de l'enquête :

- Soit par mail avis pour ceux qui disposent d'une adresse e-mail dans la base de sondage
- Soit par lettre avis papier envoyée par courrier postal pour ceux qui disposent d'une adresse postale valide et qui n'ont pas de mail.

Pour certaines populations, notamment les individus rentrant dans le champ d'extension d'échantillon, un double envoi mail et courrier pourra être effectué. Pour ceux qui ne disposent pas de coordonnées postales ou électroniques, l'information se fera lors du premier contact téléphonique avec la possibilité d'envoyer le mail avis sur demande.

Les modèles de lettre-avis et de mail-avis sont présentés ci-après.

Une plaquette d'information sur l'enquête sera jointe aux lettres avis. La maquette de cette dernière est également présentée.

▪ **Lettre-avis (recto et verso)**

Marseille, le XX mars 2020

<Civilité> <Nom> <Prenom>
 <Adr1>
 <Adr2>
 <Adr3>
 <CP> <Commune>

Objet : Enquête statistique nationale « Génération 2017 » sur l'insertion professionnelle

Bonjour <Civilité> <Prenom> <Nom>,

Le Centre d'études et de recherches sur les qualifications (Céreq), établissement public national, mène une importante enquête statistique visant à analyser les parcours professionnels des jeunes ayant quitté le système éducatif en 2016-2017.

Cette enquête cherche à décrire le lien entre les parcours de formation et les situations professionnelles des jeunes quel que soit le diplôme préparé ou la formation suivie. Elle constitue une aide précieuse pour ajuster ou mettre en place des politiques de formation ou d'emploi en faveur de l'insertion professionnelle.

Cette enquête est reconnue d'intérêt général et de qualité statistique par le Conseil national de l'information statistique.

Vous êtes sortis du système éducatif en 2016-2017. À ce titre, vous avez été sélectionné(e) à partir des listes de sortants fournies par les établissements de formation. Afin de permettre la prise en compte de la diversité des caractéristiques de la population et d'assurer ainsi la qualité statistique des résultats, il est très important que vous répondiez à cette enquête. Cela ne vous prendra que quelques minutes.

Nous vous invitons à répondre avant le 31 juillet 2020

Soit par internet, sur <https://enquete-generation2017.cereq.fr> (accessible en flashant le QR-code ci-contre), à l'aide des identifiants :

Identifiant : <IDENT>
 Mot de passe : <MDP>



Soit par téléphone, avec le numéro gratuit : 0 800 XXXX XXXX en mentionnant votre identifiant : <IDENT> (appel gratuit, en semaine entre XXh et XXh et le samedi entre XXh et XXh)

Vous pouvez avoir l'assurance que comme la loi en fait la plus stricte obligation, vos réponses resteront confidentielles et serviront uniquement à la réalisation de statistiques qui alimenteront des études sur l'insertion professionnelle.

En vous remerciant par avance de votre participation,

Florence LEFRESNE
 Directrice du Céreq

Vu l'avis favorable du Conseil national de l'information statistique, cette enquête est reconnue d'intérêt général et de qualité statistique, en application de la loi n° 51-711 du 7 juin 1951 sur l'obligation, la coordination et le secret en matière de statistiques. Elle a obtenu le visa n° XXXXXXXXXX du Ministère de l'économie et des finances, valable pour l'année 2020 - Arrêté en date du JJ/MM/AAAA (date de publication au Journal officiel) [ou Arrêté en cours de parution (si visa accordé avant la parution au JO)]. Les réponses à ce questionnaire sont protégées par le secret statistique et destinées au Céreq. Les réponses à l'enquête ainsi que les données obtenues par appariement seront conservées pendant dix ans à compter de la fin de la collecte pour les besoins de l'enquête. Elles seront archivées au-delà de cette durée. À tout moment, leur usage et leur accès seront strictement contrôlés et limités à l'élaboration de statistiques ou à des travaux de recherche scientifique ou historique. Cette enquête n'est pas obligatoire. Le règlement général 2016/679 du 27 avril 2016 sur la protection des données (RGPD) ainsi que la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés s'appliquent à la présente enquête. Le Céreq est seul destinataire des données d'identification (nom et coordonnées), ainsi que le prestataire, habilité par le Comité du secret de la statistique publique. Elles seront conservées par le service producteur pendant la durée de la collecte et au-delà, afin d'assurer un suivi de cohorte ou la réalisation de post-enquêtes qualitatives. Les personnes enquêtées peuvent exercer un droit d'accès, d'opposition, de rectification, d'effacement ou de limitation de traitement pour les données les concernant pendant la période de conservation des données d'identification. Ces droits peuvent être exercés auprès du Céreq, que vous pouvez contacter à l'adresse dpo-generation@cereq.fr. Vous pouvez si vous l'estimez nécessaire adresser une réclamation à la Cnil.

Centre d'études et de recherches sur les qualifications - www.cereq.fr - Établissement sous tutelle du ministère de l'Éducation nationale et du ministère du Travail.

DOCUMENT A CONSERVER – NE PAS ENVOYER

Votre parcours depuis la sortie des études en 2016-2017

- Nous vous invitons à remplir ce calendrier d'activité afin de faciliter l'entretien téléphonique. L'enquêteur disposera sur écran d'un calendrier vierge qu'il remplira avec vous oralement. Vous trouverez dans la plaquette d'informations, accompagnant ce courrier, un guide pour le compléter.

	2017												2018												2019												2020											
	Janv	Fév.	Mars	Avril	Mai	Juin	Juill.	Août	Sept	Oct	Nov	Déc.	Janv	Fév.	Mars	Avril	Mai	Juin	Juill.	Août	Sept	Oct	Nov	Déc.	Janv	Fév.	Mars	Avril	Mai	Juin	Juill.	Août	Sept	Oct	Nov	Déc.	Janv	Fév.	Mars	Avril	Mai	Juin	Juill.					
Emploi																																																
Recherche d'emploi																																																
Reprise d'études ou formation																																																
Autres situations																																																

Exemple de parcours :

- Dans cet exemple, un jeune ayant quitté le système éducatif en juin 2017, a recherché un emploi de juillet à octobre 2017. Il a travaillé de novembre 2017 à novembre 2018 dans l'entreprise 1 à Paris. Il a ensuite effectué une formation de trois mois, puis de mars à août 2019, il a retrouvé un emploi dans l'entreprise 2 à Paris. Enfin, depuis septembre 2019, il travaille dans l'entreprise 3 toujours à Paris.

	2017												2018												2019												2020											
	Janv	Fév.	Mars	Avril	Mai	Juin	Juill.	Août	Sept	Oct	Nov	Déc.	Janv	Fév.	Mars	Avril	Mai	Juin	Juill.	Août	Sept	Oct	Nov	Déc.	Janv	Fév.	Mars	Avril	Mai	Juin	Juill.	Août	Sept	Oct	Nov	Déc.	Janv	Fév.	Mars	Avril	Mai	Juin	Juill.					
Emploi																																																
Recherche d'emploi																																																
Reprise d'études ou formation																																																
Autres situations																																																

Embauche
Fin

▪ Mail- avis

De : Florence LEFRESNE | Céreq

Objet : Enquête statistique nationale auprès de votre Génération



Bonjour <Civilite> <Prenom> <Nom>,

Le Centre d'études et de recherches sur les qualifications (Céreq), établissement public national, mène une importante enquête statistique visant à analyser les parcours professionnels des jeunes ayant quitté le système éducatif en 2016-2017.

Les objectifs de cette enquête sont de produire des données sur l'insertion professionnelle, selon les formations, à destination des acteurs publics. Elle constitue une aide précieuse pour ajuster ou mettre en place des politiques de formation ou d'emploi en faveur de l'insertion des jeunes.

Vous avez été sélectionné(e) à partir des listes fournies par les établissements de formation.

Votre participation est essentielle. Il est important que vous répondiez à cette enquête, afin de permettre la prise en compte de la diversité des caractéristiques de la population et d'assurer ainsi la qualité statistique des résultats.

Nous vous remercions de répondre avant le 31 juillet 2020

Par internet, sur <https://enquete-generation2017.cereq.fr>

En vous connectant avec les identifiants :

Identifiant : <IDENT>

Mot de passe : <MDP>

Par téléphone, avec le numéro gratuit : 0 800 XXX XXX

en mentionnant votre identifiant : <IDENT>

(appel gratuit, en semaine entre XXh et XXh et le samedi entre XXh et XXh)



Enquête nationale



Confidentialité des réponses



Respect de la législation

Je vous donne l'assurance que, comme la loi en fait la plus stricte obligation, vos réponses resteront confidentielles et serviront uniquement à la réalisation de statistiques qui alimenteront des études sur l'insertion professionnelle.

Je vous invite également à vous rendre sur notre [site internet](#) pour plus d'informations.

En vous remerciant par avance de votre participation,

Florence LEFRESNE
Directrice du Céreq

En application de la loi n° 51-711 du 7 juin 1951 modifiée sur l'obligation, la coordination et le secret en matière de statistiques, les réponses à ce questionnaire sont protégées par le secret statistique et destinées au Céreq. Les réponses à l'enquête ainsi que les données obtenues par appariement seront conservées pendant dix ans à compter de la fin de la collecte pour les besoins de l'enquête. Elles seront archivées au-delà de cette durée. À tout moment, leur usage et leur accès seront strictement contrôlés et limités à l'élaboration de statistiques ou à des travaux de recherche scientifique ou historique. Le règlement général 2016/679 du 27 avril 2016 sur la protection des données (RGPD) ainsi que la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés s'appliquent à la présente enquête. Le Céreq est seul destinataire des données identifiantes (nom et coordonnées), ainsi que le prestataire, habilité par le Comité du secret de la statistique publique. Elles seront conservées par le service producteur pendant dix ans c'est à dire pendant la durée de la collecte et au-delà, afin d'assurer le suivi de cohorte ou la réalisation de post-enquêtes qualitatives. Les personnes enquêtées peuvent exercer un droit d'accès, d'opposition, de rectification, d'effacement ou de limitation de traitement pour les données les concernant pendant la période de conservation des données d'identification. Ces droits peuvent être exercés auprès du Céreq que vous pouvez contacter à l'adresse dpo-generation@cereq.fr. Pour toute question relative au traitement de vos données, vous pouvez également contacter les délégués à la protection des données à l'adresse dpo-generation@cereq.fr. Vous pouvez si vous l'estimez nécessaire adresser une réclamation à la Cnil.

Pour ne plus recevoir de courrier électronique de la part du Céreq cliquer [ici](#).

- **Plaquette de présentation de l'enquête (recto et verso)**

| Confidentialité des données

Vos réponses au questionnaire sont couvertes par le secret statistique et ne donneront lieu qu'à des exploitations statistiques. Cette enquête est labélisée d'intérêt général et de qualité statistique par le Comité national d'information statistique (CNIS). Vos coordonnées nous ont été communiquées par votre établissement de formation. Celles-ci seront systématiquement supprimées avant mise à disposition des données aux chercheurs.

| Que deviennent les réponses ?

Les résultats de cette enquête donnent lieu à des publications en direction de différents acteurs publics politiques, des professionnels de la formation et de l'emploi ainsi qu'au grand public. Vous pouvez consulter celles des précédentes Génération sur notre site :

www.cereq.fr/enquetes-et-donnees



DEPUIS 1971

Mieux connaître les liens formation - emploi - travail.
Un collectif scientifique au service de l'action publique.

SAVOIR-FAIRE
Pôle d'expertise,
études, enquêtes
et recherches.

AU SERVICE
Des acteurs de la
formation, du travail
et de l'emploi.

CHAMPS

Formation initiale et continue, parcours, accès
à l'emploi, insertion professionnelle des jeunes,
mobilité professionnelle, organisation du travail,
évolution des métiers, des compétences
et des certifications.

NOUS CONTACTER

www.cereq.fr

generation2017@cereq.fr

0800 XXX XXX

appel gratuit,
du lundi au vendredi de 10h à 21h
et le samedi de 10h et 19h

Suivez-nous sur FB,
Twitter et LinkedIn

Lien du groupe G17 dans FB

<https://www.facebook.com/cereqG17>

Dans tous les cas, pensez à vous munir de votre
identifiant communiqué dans le courrier

 MINISTÈRE DU TRAVAIL
MINISTÈRE DE L'ÉDUCATION NATIONALE



| Enquête 2020 auprès de la Génération 2017

Votre guide d'enquête |





LE DISPOSITIF GÉNÉRATION

De quoi s'agit-il ?

Réalisées par le Céreq, les enquêtes génération observent l'entrée dans la vie active de l'ensemble d'une cohorte ayant quitté le système scolaire une même année. Emploi, chômage, formation, inactivité, ... Elles apportent un éclairage sur la diversité des parcours et les effets du contexte économique et social sur l'insertion professionnelle des jeunes.

Pourquoi vous interroge-t-on ?

L'enquête 2020 cherche à décrire les parcours des sortants du système éducatif en 2016-2017, quel que soit le diplôme préparé ou la formation suivie. Elle constitue une aide précieuse pour ajuster ou mettre en place des politiques en faveur de l'insertion des jeunes. Les résultats permettront également d'améliorer l'information à disposition des jeunes et de leur famille au moment de l'orientation scolaire.

... par téléphone ou par internet

Vous avez la possibilité de répondre à l'enquête par téléphone avec l'aide d'un enquêteur, ou par internet en vous connectant directement à votre questionnaire.

... sur quels sujets ?

→ sur votre parcours de formation afin de vérifier que vous faites bien partie de la Génération 2017.

→ sur les diplômes obtenus, séjours à l'étranger, expériences de travail en cours d'études.

→ sur votre parcours professionnel depuis votre sortie du système éducatif, vos perspectives professionnelles, vos mobilités résidentielles ainsi que sur votre environnement familial.



PRÉPARER L'ENTRETIEN

Étape 1 | Rassemblez si possible les documents suivants

- CV.
- Contrat de travail.
- Bulletin de salaire (celui du 1^{er} mois entièrement travaillé et le dernier de chaque emploi).

Étape 2 | Souvenez-vous de votre situation en vous aidant d'évènements marquants

Décembre 2017 :

Mort de Johnny Hallyday

Juillet 2018 :

La France remporte la Coupe du monde de football en Russie

Mai 2019 :

Ultimes épisodes de la série à succès Game of Throne.

Exemple de parcours d'un jeune

- a quitté le système éducatif en juin 2017,
- a recherché un emploi de juillet à octobre 2017.
- a travaillé de novembre 2017 à novembre 2018 dans l'entreprise 1 à Paris.
- a ensuite effectué une formation de 3 mois,

Étape 3 | Remplissez le calendrier d'activité vierge

Repérez votre mois de sortie de formation en 2016-2017 et rayez les mois qui précèdent. Pour les mois suivants, encadrez vos différentes situations comme dans l'exemple.

Pour remplir le calendrier, déterminez la situation principale pour chaque mois.

Exemple : 10 jours de travail et 20 jours de recherche d'emploi, considérez le mois entier en recherche d'emploi.

Quelques précisions :

« *Emploi* » = Toute activité rémunérée y compris les congés payés, congés maladie de moins de 6 mois, congés maternité/paternité, stage post-doctoral.

→ Chaque changement de lieu de travail donne lieu à une nouvelle situation d'emploi.

→ Situations d'intérim/intermittent, saisir une seule période si vous avez eu un ou plusieurs contrats dans une ou plusieurs entreprises.

« *Recherche d'emploi* » = en situation de recherche d'emploi, inscrit ou non, rémunéré ou non, par Pôle emploi ou un autre organisme.

« *Autre situation* » = congé parental, homme/femme au foyer s'il/elle ne recherche pas d'emploi, congés de longue durée ou longue maladie et invalidant (sans activité professionnelle), bénévolat, volontariat, ect.

→ puis de mars à aout 2019, il a retrouvé un emploi dans l'entreprise 2 à Paris.

→ enfin, depuis septembre 2019, il travaille dans l'entreprise 3 toujours à Paris.

	2017												2018												2019												2020											
	Janv	Fev	Mars	Avril	Mai	Juin	Juil	Août	Sept	Oct	Nov	Dec	Janv	Fev	Mars	Avril	Mai	Juin	Juil	Août	Sept	Oct	Nov	Dec	Janv	Fev	Mars	Avril	Mai	Juin	Juil	Août	Sept	Oct	Nov	Dec	Janv	Fev	Mars	Avril	Mai	Juin	Juil					
Emploi	X	X	X	X	X	X	X	X	X	X	X	Entreprise 1 à Paris (12***)												Entreprise 2 à Paris(11**)												Entreprise 1 à Paris (12***)												
Recherche d'emploi	X	X	X	X	X	X	X	X	X	X	X																																					
Reprise d'études ou formation	X	X	X	X	X	X	X	X	X	X	X																																					
Autres situations	X	X	X	X	X	X	X	X	X	X	X																																					

Embauche

Fin

5. Questionnaire

5.1 - Compte-rendu détaillé des tests

À l'issue de la validation du questionnaire multimode entre le Céreq et le prestataire, un pilote a été organisé pour un test en réel. Il s'est déroulé du 16 au 29 septembre 2019.

Ce test avait les objectifs suivants :

- valider le bon fonctionnement du questionnaire (filtres, webservices, soucis éventuels de programmation...)
- vérifier la fluidité du questionnement et la bonne compréhension des questions par les enquêtés
- effectuer une estimation du temps de passation du questionnaire, au global et par module

Par ailleurs, ce pilote a permis de tester en réel le système d'information de Médiamétrie, qui assure notamment le passage d'un mode de recueil à un autre de façon transparente pour l'enquêté et enregistre l'ensemble des données et para-données (réponses au questionnaire, nombre et modes de connexions utilisés...).

▪ Protocole du test

Nous nous étions fixé comme objectif le recueil de 300 questionnaires complétés, dont 150 en CATI pur (pour nos estimations de durée), le reste en CAWI ou multimode. L'échantillon envoyé en production était de 7 998 individus. Le fichier de réserve n'a pas été utilisé.

Pour ce premier test, nous avons procédé à un protocole simplifié en deux phases, correspondant aux phases 1 et 3 de l'enquête prévue en 2020. La première semaine était dédiée à la collecte internet et la deuxième à la relance téléphonique. Les enquêtés pouvaient continuer à répondre sur internet lors de la semaine téléphone. Des individus qui avaient commencé à répondre sur internet ont été rappelés (en fin de plateau), afin de vérifier le fonctionnement de la reprise des questionnaires par téléphone.

Le calendrier a été le suivant :

- 16/09 - 18/09 : envoi des mails-avis
- Du 16/09 au 23/09 : recueil en Cawi uniquement
- 23/09 matin : formation des enquêteurs
- Du 23/09 après-midi au 28/09 soir : collecte CATI, en laissant le CAWI ouvert

Le premier jour de plateau, ont été formés : 15 enquêteurs (dont 14 mobilisés), 4 superviseurs (2 mobilisés) et 2 contrôleurs (non mobilisés). 4 personnes de l'équipe EIGE étaient présentes sur le plateau. Elles ont assuré en partie la formation et étaient sur le plateau durant les appels pour réaliser des écoutes et répondre aux éventuelles questions des enquêteurs/superviseurs.

Les plages d'horaires d'appels étaient les suivantes : 11h-14h et 15h-21h (arrêt des nouveaux appels à 20h30).

▪ **Bilan du test : chiffres-clés**

Sur les 7998 individus de notre échantillon, 241 individus ont intégralement complété leur questionnaire : 141 par téléphone, 87 sur internet et 13 en multimode.

On note un taux très élevé de hors champ : parmi les individus ayant répondu à l'enquête par téléphone ou sur internet, 868 individus ne sont pas allés au-delà du questionnaire filtre car ils se sont révélés hors champ (pour 241 individus qui ont terminé le questionnaire). Le taux de hors champ dans les enquêtes Génération est en général proche de 50%. Cette proportion anormalement élevée de hors champ s'explique par le fait qu'au moment du pilote, l'opération d'élimination de la base de sondage des individus repérés comme poursuivant en 2017-2018 n'avait pas été réalisée.

Tableau 11 - Répartition des répondants lors du Pilote 1

	Internet uniquement	Téléphone uniquement	Multimode	Total enquête
Questionnaires complets	87	141	13	241
Questionnaires partiels (individus dans le champ)	62	68	6	136
Individus Hors Champ	329	532	7	868
Abandons en cours de Questionnaire Filtre *	40	89	2	131
Total	518	830	28	1 376

* En cas d'abandon en cours de Questionnaire Filtre, on ne peut savoir si l'individu faisait ou non partie du champ.

Bilan des appels

Au cours de la 2^{ème} semaine du pilote, 7 186 individus ont été appelés (90% du fichier), et 12 486 appels ont été passés. Les statistiques du nombre d'appels passés sont les suivantes :

Tableau 12 – Statistiques du nombre d’appels passés

Nombre d’appels moyens au total	1,74
Nombre d’appels moyen pour réaliser une enquête dans le champ	1,62
Nombre d’appels moyens pour refus	1,68
Nombre d’appels moyens pour HC	1,42
Nombre d’appels minimum	1
Nombre d’appels maximum	5

Le tableau ci-dessous détaille le statut du dernier appel passé auprès des 7186 individus que les enquêteurs ont cherché à joindre. Parmi eux, 1282 ont pu être joints par les enquêteurs et 1005 individus sont des contacts aboutis au sens où la fiche est classée (soit parce que l’individu a répondu, soit parce qu’il est introuvable, soit parce qu’il a refusé). 74 individus ont exprimé un refus de répondre, soit environ 6% des individus joints et 7% des contacts aboutis.

Tableau 13 – Dernier état des fiches

Statut du dernier appel	N	%
Individu hors cible - personne de ce nom ici	155	2,16
Individu non francophone / Problème physique ou mental	1	0,01
Injoignable permanent	74	1,03
Refus de l’individu	46	0,64
Refus RGPD/CNIL de l’individu	28	0,39
Abandon en cours avec refus d’être rappelé	19	0,26
Individus hors champ - questionnaire terminé	532	7,4
Individus dans le champ - questionnaire terminé	150*	2,09
Total contact aboutis	1005	14
Demande de répondre par internet - envoi d’un mail	2	0,03
Prise de rdv par l’individu	137	1,91
Faux numéro	587	8,17
Ne répond pas	1156	16,10
Occupé	73	1,02
Répondeur : Individu identifié sur nom prénom	64	0,89
Répondeur sans identification possible de l’individu	3668	51,08
Fax / modem	4	0,06
Pas de sonnerie	79	1,10
Raccroche au nez après 2ème tentative	79	1,10
Refus du tiers de nous transmettre des coordonnées	30	0,42
Prise de rdv par un tiers avec validation ou modification du téléphone	23	0,32
Envoi d’un e-mail suite enrichissement par un tiers sans enrichissement tel	52	0,72

Envoi d'un e-mail suite enrichissement par un tiers + enrichissement tel	67	0,93
Validation d'un téléphone par un tiers sans prise de rdv et sans envoi de mail	7	0,10
Refus du tiers de nous transmettre des coordonnées avec mail enrichi sur un appel précédent	1	0,01
Interruption avec rappel (Pour finir l'enquête)	49	0,68
Abandon en cours de Questionnaire Filtre	89	0,92
Classement non identifié	3	0,05
Total contact non aboutis	6181	85,40
Nombre total d'individus contactés	7186	90
Nombre total d'individus	7998	

*Parmi ces 150 individus, 9 sont des répondants en multimode.

Bilan de l'envoi des mails

Pour ce premier pilote, nous avons souhaité tester deux maquettes de mails-avis légèrement différentes. Sur le fond, les éléments étaient sensiblement les mêmes, mais les avions disposés dans un ordre différent, avec quelques variantes de formulation pour voir l'effet produit.

Parmi les 7 998 individus de notre échantillon, 7 941 disposaient d'une adresse mail. Nous avons adressé aléatoirement pour la moitié de l'échantillon le mail avis 1 et pour l'autre moitié le mail avis 2. Ces mails ont été envoyés entre le 16 et le 18 septembre.

À noter que des difficultés ont été rencontrées avec le prestataire pour la réalisation de ces mails avis : il a été impossible pour eux de nous proposer des versions responsive, c'est-à-dire qui s'adaptent au type de support de lecture (smartphone, ordinateur). Nous avons dû accepter l'envoi des mails en l'état pour pouvoir démarrer le pilote. Cela a pu nuire au taux de réponse par internet, car certains individus ont pu douter du sérieux de l'expéditeur. Pour la suite, et dès le prochain pilote, une agence de communication travaillera sur la réalisation de ces mails-avis pour le compte de Médiamétrie.

Tableau 14 – Mails avis et questionnaires recueillis

	Nombre de mails envoyés	Nombre de mails classés en erreur	Nombre de mails ouverts	Nombre de clics sur le lien du questionnaire	Nombre de répondants CAWI pur dans le champ	Nombre de répondants CATI pur dans le champ	Nombre de répondants multimode dans le champ
Mail 1	3967	382	2021	294	44	65	4
Mail 2	3974	352	1809	250	43	71	9

Le tableau 14 résume l’aboutissement des mails-avis envoyés. Les deux modèles ont conduit à un taux de questionnaires Cawi de l’ordre de 1,1%.

Au cours de la semaine de CATI, deux relances ont été effectuées auprès des individus qui avaient commencé le questionnaire. Le tableau 15 donne les statistiques liées à ces relances.

Tableau 15 – Mails de relance et questionnaires recueillis

	Nombre de mails envoyés	Nombre de mails classés en erreur	Nombre de mails ouverts	Nombre de clics sur le lien du questionnaire	Nombre de répondants CAWI pur dans le champ	Nombre de répondants CATI dans pur le champ	Nombre de répondants multimodes dans le champ
Relance - 24/09	346	3	271	54	8	0	10
Relance - 27/09	116	2	63	4	3	1	5

▪ **Bilan qualitatif du test**

Ce test s’est globalement bien déroulé. Il n’y a pas eu de blocage majeur du questionnaire : les principaux dysfonctionnements observés sont ceux liés aux recherches d’établissements employeurs avec l’outil Cap adresse, cependant, une solution temporaire a été mise en place pour contourner le problème.

Les jeunes que nous sommes parvenus à joindre ont manifesté une attitude globalement très positive vis-à-vis de l’enquête et du questionnaire. Même lorsque la durée a dépassé la demi-heure, ils se sont montrés volontaires pour aller au bout. Il semble que l’argumentaire présentant les objectifs de cette enquête ait été convaincant, les jeunes interrogés semblaient participer volontiers à l’éclairage des politiques publiques en matière de formation et d’insertion des jeunes. En outre, il ressort que ce temps passé à faire un bilan de leurs premières années de vie active est un moment apprécié, qui leur offre la possibilité de prendre du recul sur leur parcours. À noter que l’attitude des parents a été souvent moins favorable : lorsque les enquêteurs se sont trouvés avec les parents des jeunes au téléphone, ces derniers ont souvent manifesté une grande méfiance notamment pour fournir un numéro où joindre leurs enfants.

Le Youth Proofing

Pour ce pilote, Médiamétrie a fait venir une intervenante spécialiste des focus group, diplômée en psychologie sociale. Un recrutement avait été effectué pour sélectionner des jeunes faisant partie de notre champ et résidant à Amiens actuellement. Six jeunes sont ainsi venus dans les locaux de

Médiamétrie et ont complété le questionnaire (par téléphone ou sur internet), suite à quoi nous avons recueilli leurs impressions.

Cela a permis de confirmer certains points à améliorer perçus lors des écoutes :

- Quelques redondances
- Des formulations de questions à alléger pour écourter et éviter un côté répétitif
- Proposer davantage de réponses spontanées
- Effort de mémoire et redondance des questions liées aux adresses (pas de possibilité d'indiquer que l'adresse n'a pas changé d'une date à l'autre)
- Nécessité d'introduire certaines questions/modules pour en comprendre l'utilité : pourquoi demander les adresses ? pourquoi poser des questions sur les parents ?

Par ailleurs, ces jeunes ont également eu la possibilité de donner leur opinion sur les mails-avis, et de comparer les deux versions proposées en termes de présentation du texte, de logos, d'ordre des éléments présents, d'objet du mail. Cela nous a aidé à faire une nouvelle proposition de mail-avis à partir des points forts de chacun des deux modèles (voir 4.3).

▪ **Minutage du questionnaire**

Les durées sont calculées sur les enquêtés ayant répondu sur un seul mode de collecte. Les effectifs correspondent au nombre d'enquêtés ayant répondu à au moins une question du module.

La durée cible de ce questionnaire est de 30 minutes. Les durées ont été calculées pour le questionnaire dans sa globalité mais aussi par module (voir Tableau 16) et par question : cela permet de détecter des modules ou questions qui seraient trop chronophages et peut conduire à des coupes dans le questionnaire.

Tableau 16 - Durée de passation du questionnaire avec détail par module

Module		Effectif		Durée Moyenne	
		Internet	Téléphone	Internet	Téléphone
QF	Questionnaire Filtre	518	830	03:24	02:45
Total Parcours de formation		153	213	09:16	09:45
PS	Parcours scolaire	153	213	03:42	04:09
PHD	Plus haut diplôme	144	198	01:11	01:16
ETR	Séjour à l'étranger	143	195	01:26	01:04
LOG	Logement	143	202	02:05	02:16
PSS	Post-Initiaux Sport				
FE	Financement des études	139	188	00:50	01:03
EXP	Expériences professionnelles	51	58	01:10	01:37
Total Parcours professionnel		131	178	18:24	15:56
CAL	Calendrier	131	178	00:30	00:18
EP	Pop-Up Emploi Passé	117	169	01:56	01:04
EA	Pop-Up Emploi actuel	117	167	01:32	00:51
REP	Pop-Up Recherche d'Emploi du passé	101	155	00:51	00:28
REA	Pop-Up Recherche d'Emploi du présent	101	151	00:54	00:28
FREP	Pop-up Formation et reprise d'études du passé	116	166	00:59	00:34
FREA	Pop-up Formation et reprise d'études du présent	111	165	00:53	00:29
AUP	Pop-Up Autre situation du passé	98	149	00:39	00:19
AUA	Pop-Up Autre situation du présent	95	143	00:40	00:20
EMPLOI PASSE					
EP	Description emplois du passé	102	128	07:29	07:18
EPD	Départ de l'entreprise	102	128	00:28	00:28
EMPLOI ACTUEL					
EA	Description emploi du présent	73	119	03:14	04:23
EAO	Opinion sur l'emploi	72	115	00:44	01:05
EAM	Multi activité	72	115	00:10	00:15
AUTRES MODULES EMPLOI					
IMT	Intermédiaire du marché du travail	104	158	01:56	01:20
SC	Service Civique	12	13	00:28	00:28
FP	Attractivité de la fonction publique	100	152	01:04	01:41
RSQ	Risque Physiques et Chimique	25	64	01:49	02:29
DIP	Diplômes obtenus en post-initial	97	149	00:26	00:38
Total Caractéristiques de l'individu		97	149	06:16	05:39
PP	Perspectives professionnelles	97	149	00:17	00:24
SD	Caractéristiques sociodémographiques	97	149	03:02	02:28
DI	Discrimination	96	147	00:21	00:19
CC	Calendrier de mode de cohabitation	95	145	00:54	00:39
MR	Mobilité résidentielle	92	144	00:45	00:36
AC	Actualisation des coordonnées	91	142	01:04	01:18
Durée Totale du questionnaire (sans contact)		90	141	36:04:00	36:34:00

La durée moyenne de passation du questionnaire est de 36min 34s par téléphone.

Divers éléments ont pu contribuer à rallonger cette durée moyenne :

- Nous avons donné pour consigne aux enquêteurs de systématiquement lister l'ensemble des modalités des questions. Certaines questions présentent un nombre élevé de modalités et

peuvent revenir à plusieurs reprises dans le questionnaire (description des divers emplois) : suite à un échange avec la DMCSI de l'Insee sur ce sujet, nous allons revenir sur cette consigne pour certaines questions ;

- L'utilisation des webservices (recherche dynamique dans des menus d'adresses, d'entreprises) n'a pas été optimale et a contraint les enquêteurs à mettre en attente les enquêtés pour nous demander comment gérer les cas posant problème. Des modifications vont être apportées sur cet aspect dans le questionnaire (détaillé ci-après) ;
- Le Céreq faisant partie du groupe de travail de refonte de la PCS, nous avons incité les enquêteurs à effectuer des saisies en clair lors des questions liées à la profession exercée dès lors que la profession ne collait pas mot pour mot à celle proposée dans la liste : ainsi dans 40% des cas, les enquêteurs ont effectué une saisie manuelle. Cette consigne va évoluer pour la suite.
- Pour assoir les propositions de modalités de certaines questions, des modalités « autre » avec un ouvert ont été ajoutées pour le test. Ces déclarations ont été analysées et ont permis de réaliser quelques ajustements. Ces modalités « Autre pour le test » seront toutes supprimées pour la suite.
- La formation n'avait pas permis de balayer le questionnaire dans son intégralité. De ce fait, les premiers jours de plateau ont contribué à parfaire la formation des enquêteurs. Ils ont ainsi gagné en rapidité au fil de la semaine, comme en atteste le tableau suivant :

Tableau 17 - Durée de passation selon la date de collecte

	Nombre d'enquêtes réalisées dans le champ	Nombre d'appels	Durée moyenne des enquêtes dans le champ
23/09/2019	15	1 589	35:39
24/09/2019	23	2 294	41:21
25/09/2019	25	1 910	37:56
26/09/2019	22	2 169	32:19
27/09/2019	20	1 855	36:35
28/09/2019	36	2 669	32:44
Total	141	12 486	36:34

- Le questionnaire est très filtré et tous les individus n'empruntent pas les mêmes chemins. En particulier, le fait d'être sorti de niveau secondaire ou supérieur, et le fait de sortir d'une voie générale ou professionnelle vont avoir un impact sensible sur le nombre de modules parcouru

et la durée passée dans ces modules. Or la structure des répondants du pilote ne reflète pas la structure des répondants d'une enquête : il n'y a pas eu d'échantillonnage mais une sélection simple d'individus ; par ailleurs nous avons cherché à forcer la structure des répondants en fermant des adresses en cours de plateau pour nous assurer d'avoir des profils variés. En opérant un redressement du temps de réponse par application de la structure des répondants de la Génération 2013, on aboutit à un temps moyen de 34mn 36s.

Ces remarques étant faites, nous avons opéré un certain nombre de changements dans le questionnaire, avant tout pour chercher à gagner du temps, mais aussi pour gagner en fluidité. Les principaux changements sont détaillés dans la partie suivante.

- **Les enseignements du test : les modifications effectuées dans le questionnaire**

- Gestion de la collecte des adresses individuelles

Le questionnaire prévoit la collecte d'adresses postales à différents moments du parcours. Bien que l'outil de collecte soit élaboré et intuitif (normalisation à la volée), la demande d'adresse ponctuée fait défaut à la mémoire de l'enquêté. De plus, une nouvelle thématique a été introduite autour des mobilités résidentielles. Par conséquent, la proposition est de regrouper l'ensemble des demandes d'adresses au sein de ce module qui a donc été adapté par rapport à la version initiale.

Un premier tableau permet de répertorier les adresses durant le parcours de formation : adresse en sixième, adresse au bac et adresse à la fin des études (MR020). Un second qui complète ces premières informations mais qui cible le parcours professionnel. Il s'agit dans ce tableau de lister l'ensemble des déménagements intervenus depuis la fin des études en collectant les dates associées et les communes de résidence (MR030).

À noter que certaines questions sur les adresses ou communes qui concernent l'individu n'ont pas été déplacées par souci de cohérence. C'est le cas de la question sur la commune de l'établissement de formation (QF050) qui est posée seulement aux individus qui invalident l'établissement de formation issu de la base de sondage (faible taux). N'a pas non plus été déplacée la question sur la commune de résidence de la première interruption des études des post-initiaux (qui ne passent pas le module sur les mobilités résidentielles).

- Collecte des salaires et primes

Le questionnaire intègre le changement opéré sur les salaires, avec l'introduction du prélèvement à la source effectif depuis de janvier 2019. Plusieurs formulations de question ont été proposées sur la collecte des salaires, en fonction de cette date, pour recueillir la bonne information.

À l'issue du test, des remarques ont été formulées au sujet des primes notamment des difficultés à déclarer certains types de primes : prime macron, primes ponctuelle... Cette confusion a conduit les enquêtés à demander des précisions sur ce qui était attendu et cela a pu induire une perte de temps. Il s'agit donc d'améliorer le questionnement sur ces aspects. Dans le cadre du groupe de travail sur la

refonte de l'EEC, des propositions de questions sur les salaires et primes ont été formulées. La nouvelle version du questionnement s'appuie sur ces exemples. En effet, les questions sur les primes ont été regroupées (exemple : EP630, EP640 et EP650) pour ne former qu'une seule question plus générale mais intégrant une périodicité (Exemple : *EP630 - Au cours des 12 derniers mois, avez-vous perçu un 13ème mois, une prime d'intéressement, une prime exceptionnelle ?*). Des exemples de primes seront indiqués en infobulle ainsi qu'une question sur le montant global perçu sur la période. Pour compléter, les primes ne seront désormais abordées qu'une seule fois (au lieu de deux précédemment pour les séquences longues : à la date d'embauche et en fin de séquence). Cette information sur les primes perçues sera récoltée au moment du départ de l'emploi (séquence passée), ou à la date de l'enquête.

- Suppression ou fusion de questions

Deux questions du module séjour à l'étranger ont été supprimées : *ETR030A - Étiez-vous parti à l'étranger ?* car selon la définition du séjour les réponses seront majoritairement « Dans le cadre scolaire », mais aussi car l'information paraît redondante par rapport à ce qui précède.

ETR180 - L'année de ce séjour, touchiez-vous une bourse sur critères sociaux ou au mérite ? est supprimée car elle est déjà posée dans le module financement des études (FE010).

Deux questions ont été fusionnées dans le module portant sur l'attractivité de la fonction publique : *FP080 - Cet emploi vous a-t-il permis de,* résulte de la fusion des anciennes FP080 et FP090 pour mieux identifier l'apport de l'emploi au niveau personnel et professionnel.

- Autres ajustements visant à réduire le temps de passation

Nous avons décidé de filtrer davantage de questions pour réduire le temps de passation. Par exemple, ne poser la question sur la spécialité qu'à certains diplômes (création d'une indicatrice : SANSPE). Nous avons également ajouté une question « filtre » pour vérifier l'existence d'un employeur pour les jeunes déclarant un contrat de professionnalisation ou d'apprentissage (FREA061), elle permet notamment de filtrer le module logement (i.e. ne poser les questions d'accès à l'entreprise qu'à ceux qui l'ont trouvée à la date d'enquête).

Pour fluidifier la passation et gagner du temps, des consignes pour les enquêteurs ont été ajoutées, telles que : « Attendre la réponse spontanée ». Exemples : PS040, PS060, PS220, (...). Cela évite de lister l'ensemble des modalités ou de s'arrêter lorsque l'enquêté valide une modalité (pour des questions à choix unique).

Enfin, le nombre de questions ouvertes ou modalités « autres » donnant lieu à des champs ouverts a été réduit. Comme indiqué plus haut, certaines modalités « autres » avaient été ajoutées pour ce test pour valider les modalités proposées.

- Reformulations de questions

Certaines questions ont été reformulées car il est apparu qu'elles étaient mal comprises par les enquêtés.

Par exemple, la question sur le nombre d'enfants a été modifiée : *SD200 - Actuellement, combien d'enfants avez-vous ? Comptez également vos enfants ou beaux-enfants qui vivent dans votre ménage en garde partagée/alternée.* La version initiale portait à confusion. Lors du test, les jeunes, sortants du secondaire notamment, ont demandé s'ils devaient compter leurs frères et sœurs.

La question *FP040 – Que pensez-vous des affirmations suivantes relatives à la Fonction publique* a été revue avec proposition de modalités à la négative.

La question *ETR010 - Depuis votre entrée au CP et jusqu'à la fin de vos études en <FINETU>, avez-vous réalisé un séjour à l'étranger (hors vacances ou activités de loisirs) ?* a été reformulée pour lever l'ambiguïté sur la définition du type de séjour à déclarer et sa temporalité dans le parcours.

- Collecte de l'établissement employeur

Enfin, un dernier point en cours mais non visible dans le questionnaire concerne les questions liées à la recherche de l'établissement employeur des individus (qui intervient dans chaque séquence d'emploi). Le questionnaire fait appel à un webservice qui permet de rechercher une entreprise dans la base Sirene via son adresse, et de récupérer tout un ensemble d'informations (dont le code NAF). Quelques difficultés ont été rencontrées et sont en cours de discussion avec le prestataire en charge du développement de l'outil, afin d'améliorer le processus de recherche.

Actuellement la recherche d'entreprise se fait par commune ou par code postal. Lorsqu'une commune regroupe plusieurs codes postaux, il faut que l'enquêté se souvienne du code postal de son entreprise pour pouvoir la retrouver, ce qui a parfois posé souci. Nous allons donc étudier la possibilité de proposer une recherche par commune « tous codes postaux confondus ».

Par ailleurs, certaines entreprises des enquêtés n'ont pas été retrouvées : il s'agissait d'entreprises étrangères (absentes de Sirene), mais pas seulement. Un échange est en cours avec l'INSEE pour déterminer quels sont précisément les types d'entreprises/administrations présentes dans la base SIRENE disponible gratuitement. Il faudra prévoir de laisser la possibilité de saisir nom d'entreprise et un code postal dans un champ en clair (avec récupération du code INSEE).

5.2 – Questionnaire définitif (incluant les modifications résultant des tests)

Le questionnaire de l'enquête 2020 auprès de la Génération 2017 est présenté en annexe 8 (dans un fichier pdf séparé). La version fournie comprend les principales évolutions suite au test effectué en Septembre. Toutefois, il est mené actuellement un arbitrage pour éventuellement supprimer de nouvelles questions.

6. Coût des enquêtes

Six personnes (6 ETP niveau A) composent de manière permanente l'équipe ingénierie et gestion d'enquête (EIGE) du Département des entrées et évolutions dans la vie active (DEEVA) du Céreq. Cette équipe est chargée de la réalisation des enquêtes depuis la constitution de la base de sondage jusqu'à la mise à disposition des fichiers de détail et de leur documentation.

Certaines phases de l'enquête sont confiées à des sociétés externes choisies selon la procédure officielle des marchés publics. Les tâches suivantes sont sous-traitées :

- Collecte des fichiers d'inscrits auprès des établissements délivrant de la formation initiale (et ne figurant pas dans les bases centralisées fournies par les ministères)
- Restructuration, normalisation et validation des adresses postales pour les individus tirés dans l'échantillon
- Recherche de coordonnées téléphoniques pour les individus tirés dans l'échantillon
- Édition et envoi des lettres avis
- Développement du questionnaire CATI/CAWI et collecte (administration des questionnaires par téléphone)
- Mise à disposition de webservices pour la recherche d'adresses et d'entreprises
- Suivi du plateau durant toute la durée de l'enquête par une personne désignée comme « référent qualité »

Chacune de ces opérations est suivie et contrôlée par l'EIGE.

Le tableau 18 détaille les coûts externes liés à l'enquête, et affiche les recettes perçues par les partenaires d'extension.

Tableau 18 – Coût anticipé de l'enquête Génération 2017

DEPENSES		RECETTES	
Collecte de la base de sondage	43 725 €	Extension formations environnementales (échantillon et questionnement) - CGDD	82 845 €
Développement du questionnaire et collecte (Médiamétrie) *	934 255 €	Extension de questionnement sur le service civique	4 568 €
Normalisation des adresses postales *	10 500 €	Extension de questionnement sur les séjours à l'étranger des jeunes en cours d'études - INJEP, Erasmus+, OFQJ	14 225 €
Enrichissement de coordonnées téléphoniques *	21 000 €	Extension d'échantillon et de questionnement sur les sortants de l'enseignement supérieur - DGESIP	99 616 €
Impression et routage des lettres-avis *	28 000 €	Extension d'échantillon destiné à produire des indicateurs sur les diplômés des secteurs du sport et de l'animation - INJEP	10 000 €
Mise à disposition de Webservices pour les recherches d'adresses et d'entreprises (Cap Adresse) *	12 500 €	Extension de questionnement sur l'attractivité de la fonction publique - DGAFP	40 236 €
Frais postaux pour l'expédition des lettres avis *	50 000 €	Extension de questionnement sur les risques physiques et chimiques - DARES	
Rémunération du personnel vacataire (réfèrent(s) qualité) *	24 000 €	Extension d'échantillon et questionnement sur les sortants de la région Bretagne - Région Bretagne	142 924 €
		Extension de questionnement sur le logement - Injep, Sdes/DHUP	
		Extension d'échantillon sur les jeunes sortants de formation en région Pays de la Loire -PDLL	43 000 €
		Extension d'échantillon sur les jeunes sortants de formations automobiles - ANFA	43 000 €
		Extension d'échantillon sur les jeunes issus des QPV - CGET	99 487 €
Total DEPENSES	1 123 980 €	Total RECETTES	579 901 €

* La plupart des dépenses ont été approximées sur la base d'une prévision de la taille de l'échantillon tiré, du nombre de répondants à l'enquête, de la durée moyenne des questionnaires et de la répartition entre CATI et CAWI (en faisant des hypothèses hautes à chaque fois).

L'extension d'échantillon de la région Pays de la Loire est à confirmer, mais a été prise en compte dans les estimations de coûts.

Les montants sont hors taxes.

Liste des annexes

A.1- Avis d'opportunité du Cnis

A.2- Fiche de présentation de l'enquête présentée au Cnis pour examen d'opportunité

A.3 – Compte-rendu du Comité de concertation du 14 Novembre 2018

A.4 – Publications récentes

A.5 - Cahier des clauses techniques particulières (CCTP)

A.6 – Inscription au registre des traitements de données à caractère personnel du Céreq

A.7 – Analyse d'impact (PIA) de l'enquête Génération 2017

A.8 – Questionnaire de l'enquête 2020 auprès de la Génération 2017 (dans un document séparé)

A.1- Avis d'opportunité du Cnis



Montrouge, le 14 mai 2019 n°88 / H030

AVIS D'OPPORTUNITÉ

Enquête de cheminement tous niveaux de sortie du système éducatif. Génération 2017

Type d'opportunité : réédition d'enquête déjà réalisée

Périodicité : environ tous les 3 ans

Demandeur : Centre d'études et de recherches sur les qualifications (Céreq)

Au cours de sa réunion du 12 avril 2019, la commission « Emploi, qualification et revenus du travail » a examiné la demande de reconduction de l'opportunité pour l'enquête de cheminement tous niveaux de sortie du système éducatif, Génération 2017.

Le dispositif Génération a été mis en place à la fin des années 1990. Les deux ministères de tutelle du Céreq, c'est-à-dire le Ministère en charge de l'Education et le Ministère en charge de l'Emploi, sont à l'origine de la demande.

Jusqu'à présent, sept générations de sortants du système éducatif ont été interrogées : les Générations 1992, 1998, 2001, 2004, 2007, 2010 et 2013. Jusqu'à la Génération 2013, une cohorte de sortants sur deux faisait l'objet de ré-interrogations : la Génération 1998 a été enquêtée quatre fois (trois ans, cinq ans, sept ans et dix ans après la sortie du système éducatif) ; les Générations 2004 et 2010 ont été enquêtées trois fois (trois ans, cinq ans et sept ans après la sortie du système éducatif).

Après presque 20 ans d'existence, le Céreq a rénové son dispositif des enquêtes Génération. L'objectif est d'en réduire le coût afin de l'inscrire dans le long terme et de mieux saisir les mutations à l'œuvre dans le système scolaire et sur le marché du travail. Les principaux éléments de cette rénovation ont été présentés et discutés devant cette même commission du Cnis (séance du 9 novembre 2017). A partir de la Génération 2017, une nouvelle génération de sortants sera interrogée tous les quatre ans. Dans ce dispositif rénové, chaque cohorte, de taille identique, fera l'objet d'une première interrogation trois ans après la sortie du système éducatif puis d'une ré-interrogation à six ans. Le mode de collecte a également évolué : les jeunes auront la possibilité de répondre au questionnaire par Internet en plus de l'habituel mode de collecte par téléphone.

Le dispositif des enquêtes Génération permet d'étudier l'accès à l'emploi et les premières transitions professionnelles des jeunes à l'issue de leur formation initiale. L'objectif principal est d'assurer une production régulière d'indicateurs standards sur l'insertion professionnelle des jeunes trois ans après la sortie du système éducatif, par grands niveaux de formation et diplôme, et par grands domaines de formation. Cette enquête permet également d'obtenir des indicateurs de cadrage sur l'insertion au niveau régional. Par ailleurs, le suivi longitudinal, jusqu'aux six premières années de vie active (grâce à la ré-interrogation six ans après la sortie du système éducatif) permet l'étude des débuts de carrières et des mobilités sur moyen terme. Enfin, ces enquêtes offrent la possibilité d'examiner la place et le rôle des mesures pour l'emploi ou la formation en début de vie active.

Le questionnaire de tronc commun, après une première partie destinée à valider l'identification de l'individu et à vérifier les critères d'éligibilité, aborde successivement les thèmes suivants : le parcours scolaire et les diplômes obtenus, le calendrier mensuel d'activité sur les trois années suivant la sortie du système éducatif avec une description détaillée des situations et conditions d'emploi, les caractéristiques individuelles, l'environnement familial et les mobilités résidentielles.

A ce questionnaire s'ajoutent des modules de questionnement thématiques. Ainsi, en décrivant la situation professionnelle des jeunes mois par mois depuis la sortie de formation initiale en 2017 jusqu'à la date de l'interrogation au printemps 2020, l'enquête permet une description des trajectoires en termes d'emploi, de chômage, de reprise d'études, de formation ou d'autres situations d'inactivité. Les parcours professionnels des jeunes au cours de leurs trois premières années de vie active sont analysés en tenant compte de leur parcours scolaire, du(des) diplôme(s) obtenu(s) ainsi que de leurs caractéristiques sociodémographiques.

L'enquête a pour champ géographique les individus sortant d'un établissement de formation situé en France métropolitaine ou dans un département ou région d'outre-mer. Les individus enquêtés peuvent résider en France ou à l'étranger au moment de l'enquête. Elle aura lieu en 2020 entre avril et juillet. La réinterrogation de cette Génération aura lieu en 2023.

La taille de l'échantillon envoyé en production sera comprise entre 200 000 à 300 000 individus. L'échantillon enquêté comprendra entre 25 000 et 45 000 répondants

La réalisation des enquêtes sera confiée à un prestataire sous-traitant choisi selon la procédure officielle des marchés publics. Conformément au cahier des charges et des clauses techniques rédigé par le Céreq, il sera chargé de concevoir le développement informatique du questionnaire en passation multimode (téléphonique, avec saisie simultanée des réponses sur ordinateur – CATI par un enquêteur, et Internet avec saisie en ligne du questionnaire par l'enquêté – CAWI), d'effectuer les contacts par téléphone (premier contact, relances et passation des questionnaires selon la phase de collecte) et d'effectuer les envois de mail (premier contact et relances).

Le temps de passation du questionnaire sera de l'ordre de 30 minutes en moyenne.

Des post-enquêtes pourront être réalisées par internet sur la base des adresses électroniques demandées aux individus enquêtés ou par entretien (approche qualitative) auprès des individus ayant donné leur consentement pour une telle opération. Dans une visée exploratoire, des appariements avec des sources administratives pourront être envisagés notamment avec les DADS de l'Insee dans l'objectif de mesurer dans le cadre de la description des périodes d'emploi, si, à terme, cette source administrative peut permettre de réduire la taille du questionnaire, d'enrichir les informations recueillies et d'améliorer la qualité de l'information en évitant les effets de mémoire et de reconstruction biographique.

Le comité de concertation des enquêtes Génération se réunit deux fois par an. Il est composé des représentants des principaux partenaires institutionnels du Céreq (Direction de l'évaluation, de la prospective et de la performance – Depp, Direction de l'animation, de la recherche, des études et des statistiques – Dares, Direction générale de l'enseignement scolaire – Dgesco, Direction générale de l'enseignement supérieur et de l'insertion professionnelle – Dgesip, Délégation générale à l'emploi et de la formation professionnelle – Dgefp), de la Direction Générale de l'administration et de la fonction publique (Dgafp), de France Stratégie, de Pôle emploi, de l'Institut national de la jeunesse et de l'éducation populaire (Injep), de l'Institut national de la statistique et des études économiques (Insee), des partenaires sociaux (Medef, CFDT, CGT, FO, UNSA, etc.), des représentants d'observatoires, d'organismes d'études et de recherche, des associations d'étudiants et de familles, de l'union nationale des missions locales et des représentants de la Politique de la Ville.

Les premières publications sont prévues à partir de début 2021, notamment les premiers résultats présentant les indicateurs standards de l'insertion professionnelle des jeunes (collection Céreq Bref).

Les fichiers anonymisés d'enquête seront accessibles sur demande par la communauté des chercheurs via le réseau Quetelet et l'Adisp (Archives de Données issues de la Statistique Publique).

Le Président, après avoir entendu la commission, émet un **avis d'opportunité favorable** à cette enquête. L'opportunité est accordée pour quatre années, soit de 2020 à 2023.



**Conseil national
de l'information statistique**

**Commission
« Emploi, qualification et revenus du travail »**

Réunion du 12 avril 2019

**Fiche descriptive pour
une demande d'avis d'opportunité**

**Enquête de cheminement tous niveaux de sortie du
système éducatif. Génération 2017.**

1. Titre

Intitulé de l'enquête *	Enquête de cheminement tous niveaux de sortie du système éducatif. Génération 2017.
-------------------------	---

2. Statut de l'enquête

Initiative	Enquête d'initiative nationale ou régionale
Contenu du questionnaire	Sans lien avec un règlement européen

3. Champs de l'enquête

Unité statistique enquêtée	Individu ou ménage
Champ statistique couvert	<p>Les unités enquêtées sont des individus sortants de tous niveaux de formation initiale.</p> <p>La Génération 2017 concerne les « primo sortants » de formation initiale en 2016-2017 (année scolaire), qu'ils soient diplômés ou non. Dans les « primo-sortants » sont inclus les individus ayant interrompu leurs études par le passé pendant une durée comprise entre 1 et 16 mois. Tous les niveaux et domaines de formations sont concernés.</p> <p>De façon plus précise, les critères d'éligibilité pour être dans le champ retenu, nommé ensuite « champ Céreq », sont les suivants :</p> <ul style="list-style-type: none">- Avoir été inscrit dans un établissement de formation en France métropolitaine ou dans un DROM durant l'année scolaire 2016-2017 ;- Avoir quitté le système éducatif entre octobre 2016 et octobre 2017 ;- Ne pas avoir interrompu ses études durant plus de 16 mois avant l'année scolaire 2016-2017 (sauf pour raison de santé) ;- Ne pas avoir repris ses études au cours des 16 mois qui ont suivi l'entrée sur le marché du travail;- Avoir 35 ans ou moins en 2017. <p>Quelques points particuliers concernent l'application de ces critères :</p> <p>Bien qu'il s'agisse de contrats de travail les contrats d'apprentissage et les contrats de professionnalisation à visée diplômante et qui interviennent dans la continuité du parcours scolaire initial sont considérés comme relevant de la formation initiale. Une personne repérée comme sortant de formation en 2016-2017 qui poursuit par un contrat d'apprentissage ou un contrat de professionnalisation en 2017-2018 est donc considérée en poursuite d'études, et donc hors champ.</p> <p>Une personne sortie d'un établissement de formation en 2016-2017 qui poursuit des cours par correspondance ou des cours du soir en 2017-2018 est considérée comme en poursuite d'études, donc hors champ, si elle n'a pas d'emploi en parallèle. Si elle a un emploi en parallèle, sa situation d'emploi prime.</p>

Champ géographique	L'enquête a pour champ géographique les individus sortant d'un établissement de formation situé en France métropolitaine ou dans un département ou région d'outre-mer. Les individus enquêtés peuvent résider en France ou à l'étranger au moment de l'enquête.
--------------------	---

4. Présentation de l'enquête

Titre du(es) questionnaire(s)	Première interrogation auprès d'un échantillon de sortants du système éducatif français en 2017
URL de consultation du questionnaire	https://drive.cereq.fr/d/f/475898307948847479
Objectifs	Le dispositif des enquêtes Génération permet d'étudier l'accès à l'emploi et les premières transitions professionnelles des jeunes à l'issue de leur formation initiale. L'objectif principal est d'assurer une production régulière d'indicateurs standards sur l'insertion professionnelle des jeunes trois ans après la sortie du système éducatif, par grands niveaux de formation et diplôme, et par grands domaines de formation. L'enquête Génération permet également d'obtenir des indicateurs de cadrage sur l'insertion au niveau régional. Par ailleurs, le suivi longitudinal, jusqu'aux six premières années de vie active (grâce à la ré-interrogation six ans après la sortie du système éducatif) permet l'étude des débuts de carrières et des mobilités sur moyen terme. Enfin, ces enquêtes permettent d'examiner la place et le rôle des mesures pour l'emploi ou la formation en début de vie active. Cette enquête sera la première interrogation, trois ans après leur sortie du système éducatif, des sortants de formation initiale de l'année scolaire 2016/2017.
Thèmes du (ou des) questionnaires	Le questionnaire de tronc commun (hors extensions thématiques), après une première partie destinée à valider l'identification de l'individu et à vérifier les critères d'éligibilité, aborde successivement les thèmes suivants : le parcours scolaire et les diplômes obtenus, le calendrier mensuel d'activité sur les trois années suivant la sortie du système éducatif avec une description détaillée des situations et conditions d'emploi, les caractéristiques individuelles, l'environnement familial et les mobilités résidentielles. A ce questionnaire de tronc commun s'ajoutent des modules de questionnement thématiques permettant d'approfondir certaines thématiques de recherche privilégiées. Ainsi, en décrivant la situation professionnelle des jeunes mois par mois depuis la sortie de formation initiale en 2017 jusqu'à la date de l'interrogation au printemps 2020, l'enquête permet une description des trajectoires en termes d'emploi, de chômage, de reprise d'études, de formation ou d'autres situations d'inactivité. Les parcours professionnels des jeunes au cours de leurs trois premières années de vie active sont analysés en tenant compte de leur parcours scolaire, du(des) diplôme(s) obtenu(s) ainsi que de leurs caractéristiques sociodémographiques.

Historique	<p>Le dispositif Génération a été mis en place à la fin des années quatre-vingt-dix. Sept générations de sortants du système éducatif ont été interrogées jusqu'à présent : les Générations 1992, 1998, 2001, 2004, 2007, 2010 et 2013.</p> <p>Jusqu'à la Génération 2013, une cohorte de sortants sur deux faisait l'objet de ré-interrogations : la Génération 1998 a été enquêtée quatre fois (trois ans, cinq ans, sept ans et dix ans après la sortie du système éducatif) ; les Génération 2004 et 2010 ont été enquêtées trois fois (trois ans, cinq ans et sept ans après la sortie du système éducatif) .</p> <p>Après presque 20 ans d'existence, le Céreq a rénové son dispositif des enquêtes Génération. L'objectif est d'en réduire le coût afin de l'inscrire dans le long terme et de mieux saisir les mutations à l'œuvre dans le système scolaire et sur le marché du travail. Les principaux éléments de cette rénovation ont été présentés et discutés lors de la séance du 9 novembre 2017 de la commission « Emploi, qualification et revenus du travail » du CNIS.</p> <p>A partir de la Génération 2017, une génération nouvelle de sortants sera interrogée tous les quatre ans. Dans ce dispositif rénové, chaque cohorte, de taille identique, fera l'objet d'une première interrogation trois ans après la sortie du système éducatif puis d'une ré-interrogation à six ans. Le mode de collecte a également évolué : les jeunes auront la possibilité de répondre au questionnaire par internet en plus du traditionnel mode de collecte par téléphone.</p>
------------	--

Faits marquants précédentes éditions	<p>Les principales publications issues de la précédente édition de l'enquête Génération (enquête 2016 auprès de la Génération 2013) sont listées ci-dessous :</p> <ul style="list-style-type: none"> - Calmand, J. et A. Robert (2019), « Séjours des jeunes à l'étranger : des objectifs européens partiellement atteints, mais un accès encore inégal à la mobilité », Céreq Bref, n°371. - Gaubert E., V. Henrard, A. Robert, P. Rouaud (2017), « Pas d'amélioration de l'insertion professionnelle pour les non-diplômés », Céreq Bref, n°356 & Training and Employment, n° 129. - Henrard V. et V. Ilardi (coord.) (2017), « Quand l'école est finie, premiers pas dans la vie active de la Génération 2013 », Céreq Enquêtes, n°1. - Ilardi V., O. Joseph et E. Sulzer (2018), « L'entrée sur le marché du travail des jeunes de la voie professionnelle renouvelée », Céreq Bref, n°365. <p>Par ailleurs, plusieurs publications, dont un ouvrage collectif (collection Céreq Essentiel), ont été réalisées à l'occasion des vingt ans de l'enquête Génération, interrogeant l'évolution de l'insertion professionnelle des jeunes sur longue période et reposant sur l'exploitation de plusieurs enquêtes Générations (1992, 1998, 2004 et 2010) :</p> <ul style="list-style-type: none"> - T. Couppié, A. Dupray, D. Épiphanie, V. Mora (coord.), « 20 ans d'insertion professionnelle des jeunes : entre permanences et évolutions », Céreq Essentiels, n°1. - Mora V. (2018), « En vingt ans, des parcours d'insertion dégradés mais des jeunes moins inquiets pour leur avenir professionnel », Formations et emploi, Insee Références, édition 2018. - Dupray A. et C. Barret (2019), « Que gagne-t-on à se former ? Zoom sur 20 ans d'évolution des salaires en début de vie active », Céreq Bref n°372. - Cart, B., A. Léné et M.H. Toutin (2018), « A l'aube de la réforme de la formation professionnelle, retour sur 20 ans d'insertion des apprentis », Céreq Bref n°370.
Concertation	<p>Le dispositif des enquêtes Génération dispose d'un comité de concertation (cf partie Comitologie). Lors de la rénovation du dispositif, les ministères de tutelles ainsi que d'autres ministères partenaires d'extensions à l'enquête, le conseil scientifique du Céreq, le comité de concertation ainsi que des chercheurs extérieurs au Céreq ont été consultés. Les principaux éléments de cette rénovation ont été présentés et discutés lors de la séance du 9 novembre 2017 de la commission « Emploi, qualification et revenus du travail » du CNIS.</p>
Origine de la demande	<p>Les deux ministères de tutelle du Céreq, c'est-à-dire le Ministère en charge de l'Éducation et le Ministère en charge de l'Emploi, sont à l'origine de la demande.</p>

Place dans un dispositif statistique	<p>Cette enquête s'inscrit dans le dispositif des enquêtes Génération du Céreq, dans lequel tous les quatre ans auprès une nouvelle Génération de sortants du système éducatif est interrogée.</p> <p>Une enquête Génération comprend :</p> <ul style="list-style-type: none"> - Une première interrogation trois ans après la sortie du système éducatif - Une ré-interrogation six ans après la sortie du système éducatif. <p>Des post-enquêtes sont envisageables sur des thèmes et publics ciblés.</p> <p>L'opportunité est demandée pour les deux interrogations de la Génération 2017 : la première interrogation en 2020 et la seconde interrogation qui aura lieu en 2023.</p> <p>Lors de la constitution de la base de sondage, certaines sources administratives sont mobilisées (cf. partie sur le plan de sondage). Les principales sources statistiques extérieures à l'enquête Génération du Céreq qui permettent d'appréhender les entrées dans la vie active des jeunes sont l'enquête emploi de l'Insee, les enquêtes IVA (Insertion dans la Vie Active) et IPA (Insertion Professionnelle des Apprentis) de la Depp, le panel EVA (Entrées dans la Vie Active) conduit par l'Insee et l'enquête nationale sur l'insertion professionnelle des diplômés de l'Université (Master, Licence professionnelle ou DUT) que le Sies (SSM du ministère de l'Enseignement supérieur) a mis en place.</p> <p>Au niveau européen, les sources d'information proviennent des enquêtes Labour Force Survey qui sont une harmonisation des enquêtes nationales sur l'emploi (Enquête Emploi pour ce qui concerne la France).</p> <p>L'apport des enquêtes Génération réside dans une large couverture du champ des sortants de formation initiale (diplômés ou non) et dans une taille d'échantillon suffisante pour permettre des analyses fines par diplôme et spécialité. Contrairement à d'autres enquêtes d'insertion qui visent des publics segmentés, le dispositif Génération propose un questionnement, une méthodologie et un cadre d'analyse homogène pour tous, quels que soient leur parcours scolaire, les diplômes obtenus, les domaines et voies de formation. Par rapport aux enquêtes basées sur une notion de « classe d'âge », le dispositif Génération est fondé sur une notion de « promotion » (sortie du système éducatif une année donnée). Cela permet notamment d'effectuer une comparaison de l'insertion des jeunes selon leur niveau d'étude à conjoncture « constante » puisqu'ils entrent sur le marché du travail une même année.</p>
Extensions géographiques	Des extensions régionales sont possibles.

Utilisateurs	<p>Les utilisateurs potentiels des résultats sont divers : les deux ministères de tutelles, les ministères ou autres institutions partenaires d'extensions (voir ci-dessous), les acteurs publics de l'orientation et de l'insertion des jeunes, les acteurs de la formation professionnelle, les observatoires de branches, les régions, les observatoires régionaux de l'emploi et de la formation, les chercheurs, le grand public (via les médias). Outre la mise à disposition des résultats via des études et ouvrages de référence accessibles sur le site internet du Céreq, les données sont transmises aux services statistiques ministériels des tutelles du Céreq ainsi qu'aux partenaires d'extensions. En effet, cette enquête permet également de répondre à des demandes d'extensions de champ ciblées. Par exemple sur les jeunes issus de certains niveaux de formation (exemple : extension aux sortants de l'enseignement supérieur), de certaines spécialités de formation (exemple : formations environnementales) ou de certains établissements de formation (exemple : universités). L'échantillon sera élargi en conséquence, comme cela a été fait pour chacune des enquêtes Génération depuis 1998, afin d'obtenir une meilleure précision sur les indicateurs standards. Enfin, un nombre limité de modules d'extensions de questionnement sur des thématiques de recherche privilégiées peuvent être réalisés à la demande de partenaires. Elles concernent par exemple des questionnements comme les séjours à l'étranger durant les études, le rapport au travail, l'attractivité de la fonction publique. Les partenaires du Céreq (Ministères, Régions, branches professionnelles) à l'origine des demandes d'extensions seront aussi utilisateurs des données et des résultats de cette enquête. Les fichiers individuels d'exploitation qui leur seront transmis selon une procédure informatique sécurisée ne permettront pas l'identification des personnes enquêtées dès que les bases définitives sont disponibles.</p> <p>Ces données sont par la suite accessibles aux chercheurs via la participation à un groupe d'exploitation à vocation recherche organisé par le Céreq. Enfin, les données des enquêtes Génération sont disponibles sur demande via le réseau Quetelet et l'Adisp (Archives de Données issues de la Statistique Publique) environ deux ans après la finalisation des bases.</p>
--------------	--

5. Caractéristiques techniques

Périodicité de l'enquête	Ponctuelle ou pluri-annuelle
Période de collecte	L'enquête aura lieu en 2020 entre avril et juillet. La réinterrogation de cette Génération aura lieu en 2023.

<p>Mode de collecte</p>	<p>Il s'agit d'une enquête multimode : téléphonique, avec saisie simultanée des réponses sur ordinateur (CATI) par un enquêteur, et internet avec saisie en ligne du questionnaire par l'enquêté (CAWI). L'enquête sera précédée d'une lettre-avis courrier et/ou électronique envoyée aux enquêtés pour les informer de la réalisation et des objectifs de l'enquête.</p> <p>Des post-enquêtes pourront être réalisées par internet sur la base des adresses électroniques demandées aux individus enquêtés ou par entretien (approche qualitative) auprès des individus ayant donné leur consentement pour une telle opération.</p> <p>Dans une visée exploratoire, des appariements avec des sources administratives pourront être envisagés notamment avec les DADS de l'Insee dans l'objectif de mesurer dans le cadre de la description des périodes d'emploi, si, à terme, cette source administrative peut permettre de réduire la taille du questionnaire, d'enrichir les informations recueillies et d'améliorer la qualité de l'information en évitant les effets de mémoire et de reconstruction biographique.</p> <p>Lors de la constitution de la base de sondage, certaines sources administratives sont mobilisées (cf. partie sur le plan de sondage).</p>
<p>Comitologie</p>	<p>Le comité de concertation des enquêtes Génération se réunit deux fois par an. Il est composé des représentants des principaux partenaires institutionnels du Céreq (Direction de l'évaluation, de la prospective et de la performance - Depp, Direction de l'animation, de la recherche, des études et des statistiques - Dares, Direction générale de l'enseignement scolaire - Dgesco, Direction générale de l'enseignement supérieur et de l'insertion professionnelle - Dgesip, Délégation générale à l'emploi et de la formation professionnelle - Dgefp), de la Direction Générale de l'administration et de la fonction publique (Dgafp), de France Stratégie, de Pôle emploi, de l'Institut national de la jeunesse et de l'éducation populaire (Injep), de l'Institut national de la statistique et des études économiques (Insee), des partenaires sociaux (Medef, CFDT, CGT, FO, UNSA, etc.), des représentants d'observatoires, d'organismes d'études et de recherche, des associations d'étudiants et de familles, de l'union nationale des missions locales et des représentants de la Politique de la Ville.</p>
<p>Contraintes pour l'enquêté</p>	<p>Le temps de passation du questionnaire sera de l'ordre de 30 minutes en moyenne.</p>

Coût de l'enquête	<p>Cinq personnes (5 ETP niveau A) composent de manière permanente l'équipe ingénierie et gestion d'enquête (EIGE) du Département des entrées et évolutions dans la vie active (DEEVA) du Céreq. Cette équipe est chargée de la réalisation des enquêtes depuis la constitution de la base de sondage jusqu'à la mise à disposition des fichiers de détail et de leur documentation. Certaines phases de l'enquête sont confiées à des sociétés externes choisies selon la procédure officielle des marchés publics. Pour cette enquête, le coût externe est estimé à environ 500 000 euros HT (hors coût de constitution de la base de sondage). Pour la passation du questionnaire, les tâches suivantes sont sous-traitées :</p> <ul style="list-style-type: none"> - Développement du CATI/CAWI - Restructuration, normalisation et validation des adresses postales - Recherche de coordonnées téléphoniques - Edition et envoi des lettres avis - Collecte (passation des questionnaires par téléphone). <p>Chacune de ces opérations est suivie et contrôlée par l'EIGE. Une personne est également recrutée par le Céreq pour assurer le suivi qualité de la collecte lors de la passation des questionnaires chez le sous-traitant. Elle est présente sur le plateau des télé-enquêteurs tout au long de la phase de collecte.</p> <p>A ces moyens humains et financiers dédiés à la réalisation des enquêtes s'ajoute la contribution des chercheurs et chargés d'études du DEEVA qui participent à la conception du questionnaire, contrôlent les résultats et assurent les travaux d'étude et de recherche.</p>
Organisme collecteur	<p>La réalisation des enquêtes sera confiée à un prestataire sous-traitant choisi selon la procédure officielle des marchés publics. Conformément au cahier des charges et des clauses techniques rédigé par le Céreq, il sera chargé de concevoir le développement informatique du questionnaire en passation multimode, d'effectuer les contacts par téléphone (premier contact, relances et passation des questionnaires selon la phase de collecte) et d'effectuer les envois de mail (premier contact et relances)</p>

Plan de sondage	<p>Il n'existe pas de base nationale d'élèves nominative couvrant l'ensemble des sortants du système éducatif une année donnée. Le Céreq constitue donc cette base de sondage pour l'enquête, à partir de différentes sources et en opérant quelques retraitements.</p> <p>Pour cela, deux principales opérations de collecte ont lieu pour récupérer des listes nominatives d'élèves ou d'étudiants inscrits, sortants d'un établissement français, incluant quand c'est possible les coordonnées postales, téléphoniques et électroniques :</p> <ul style="list-style-type: none"> - Dans un premier temps, une collecte de fichiers de données nominatifs centralisés au niveau national sur un champ spécifique est réalisée. Plus précisément, sont mobilisées les données du ministère en charge de l'Education avec, d'une part les élèves du second degré à l'aide du Système d'Information Statistique Consolidé Académique (SYSCA) et d'autre part les apprentis à partir du Système d'Information de la Formation des Apprentis (SIFA) ; les données du ministère en charge de la Culture sur les étudiants des écoles d'architecture et d'arts uniquement à partir de la plateforme TAIGA ; les données du ministère de l'Agriculture sur les élèves inscrits dans les lycées agricoles ou équivalents ; les données de la Direction des Sports sur les diplômés jeunesse et sport à partir de l'outil de gestion FORMATIONS et diplÔMES de l'animation et du sport (FORÔMES) les données sur les contrats de professionnalisation de la base Extrapro de la direction de l'Animation de la recherche, des Études et des Statistiques (Dares). - Dans un second temps, une collecte auprès de tous les autres établissements de formation (universités, écoles d'ingénieurs, diplômés d'écoles de la santé et du social, etc.) est réalisée spécifiquement pour le Céreq avec l'aide d'un prestataire externe pour le contact des différents établissements. La liste des établissements à enquêter est issue de la BCE (Base centrale des établissements - ministère en charge de l'Education). La collecte se déroule entre janvier et avril 2019. <p>La précédente base de sondage (Génération 2013) était composée de 1 300 000 jeunes présumés sortants du système éducatif. La taille de l'échantillon envoyé en production sera comprise entre 200 000 à 300 000 individus. L'échantillon enquêté comprendra entre 25 000 et 45 000 répondants.</p>
Taille de l'échantillon	30000

6. Service(s) producteur(s)

Service producteur principal	Centre d'études et de recherches sur les qualifications (Céreq)
Autres services producteurs	
Partenariat	Aucun

<p>Date prévisionnelle de la première publication</p>	<p>L'exploitation des résultats sera principalement réalisée par le Céreq et son réseau de Centres Associés régionaux. Les résultats feront l'objet de publications dans les collections habituelles du Céreq (Céreq Bref, Céreq Études, Céreq Enquête, Céreq échanges) et sur le site Internet du Céreq (www.cereq.fr).</p> <p>Les premières publications sont prévues à partir de début 2021, notamment un quatre pages de premiers résultats présentant les indicateurs standards de l'insertion professionnelle des jeunes (collection Céreq Bref).</p> <p>Les fichiers anonymisés d'enquête seront accessibles sur demande par la communauté des chercheurs via le réseau Quetelet et l'Adisp (Archives de Données issues de la Statistique Publique).</p> <p>Les individus enquêtés seront informés, lors de la passation de l'enquête, sur les possibilités de se procurer les premiers résultats. La collecte des adresses électroniques ainsi que la mise en place d'une page Facebook dédiée à l'enquête (et relayée sur le compte twitter du Céreq) faciliteront la diffusion des premiers résultats auprès des individus interrogés.</p> <ul style="list-style-type: none"> - Début 2021, 4 pages de premiers résultats dans la collection Céreq Bref. - 2e semestre 2021, plusieurs Bref thématiques (enseignement secondaire, supérieur, ...) - 2e semestre 2021, Ouvrage de synthèse « Quand l'Ecole est finie » dans la collection Céreq Enquêtes. - Mi-2021, Les Ministères de tutelle ainsi que les partenaires d'extension reçoivent les fichiers détaillés. - La mise à disposition à Quetelet et Adisp se fera à l'été 2023.
---	--

Liste des participants

NOM	PRENOM	ORGANISME
GARNER	Hélène	Premier ministre - France stratégie
CROGUENNEC	Yannick	Ministère des solidarités et de la santé - DREES
PICHAVANT	Anne-Sophie	Ministère du travail - DARES
REIST	Cindy	Ministère du travail – DARES
DUSSOUET	Isabelle	Ministère de l'éducation nationale et de la jeunesse/ministère de l'enseignement supérieur, de la recherche et de l'innovation/ONISEP
TESTAS	Anna	Ministère de l'éducation nationale et de la jeunesse – DEPP
HAGEGE	Claire	Ministère de l'action et des comptes publics – DGAFP/DESSI
TARDIEU	Frédéric	Ministère de l'action et des comptes publics – DGAFP/DESSI
PONCEAU	Juliette	Ministère de l'enseignement supérieur, de la recherche et de l'innovation – DGEISIP/SIES
ROBIC	Hélène	Ministère de l'enseignement supérieur, de la recherche et de l'innovation – DGEISIP/SIES
DIEUSAERT	Pascal	Ministère de la cohésion des territoires et des relations avec les collectivités territoriales - CGET
MININ	Inès	Confédération française démocratique du travail – CFDT
LEFEBVRE	Geoffrey	Institut national de la jeunesse et de l'éducation populaire – INJEP
TIMOTEO	Joaquim	Institut national de la jeunesse et de l'éducation populaire – INJEP
RAZAFINDRANOVONA	Tiaray	Institut national de la statistique et des études économiques – INSEE
BRIFFAUX	Amélie	Réseau des observatoires de l'enseignement supérieur - RESOSUP
BOUDESSEUL	Gérard	Centre associé régional au Céreq - CAR Université Caen
DURIEUX	Éric	Centre d'études et de recherches sur les qualifications – CEREQ
EPIPHANE	Dominique	Centre d'études et de recherches sur les qualifications – CEREQ
GAUBERT	Émilie	Centre d'études et de recherches sur les qualifications – CEREQ
HENRARD	Valentine	Centre d'études et de recherches sur les qualifications – CEREQ
MAZARI	Zora	Centre d'études et de recherches sur les qualifications – CEREQ
OLARIA	Manon	Centre d'études et de recherches sur les qualifications – CEREQ
ROBERT	Alexie	Centre d'études et de recherches sur les qualifications - CEREQ

En annexe, la liste complète des membres du comité de concertation.

Introduction

Le Céreq remercie l'Injep pour l'accueil dans ses locaux de ce comité de concertation des enquêtes Génération.

Pour rappel, le comité est important pour les enquêtes Génération du Céreq afin de bénéficier des avis et expertises des principaux acteurs mobilisés sur la question de la relation formation-emploi. Il s'agit d'une étape essentielle dans le processus de labélisation des enquêtes. Ce comité a également pour but de tenir ses membres informés des principales évolutions et actualités du dispositif des enquêtes Génération.

Nous souhaitons la bienvenue aux nouveaux membres. Dans la plupart des cas, il s'agit de nouveaux représentants d'organismes déjà membres du comité :

Liste des nouveaux membres

Hélène GARNER (Premier ministre/France stratégie)

Benjamin NEFUSSI, Marine GUILLERM, Anne-Sophie PICHAVANT et Cindy REIST (Ministère du travail/DARES)

Boris MÉNARD (Ministère de l'enseignement supérieur, de la recherche et de l'innovation/DGESIP/SIES)

Hélène ROBIC (Ministère de l'enseignement supérieur, de la recherche et de l'innovation)

Frédéric TARDIEU (Ministère de l'action et des comptes publics/DGAFF)

Tiaray RAZAFINDRANOVA (Institut national de la statistique et des études économiques – INSEE)

Frédéric LAINÉ (Pôle emploi)

Ines MININ (Confédération française démocratique du travail - CFDT)

Amélie BRIFFAUX (Réseau des observatoires de l'enseignement supérieur – RESOSUP)

Nous remercions les participants à cette réunion ainsi que les membres n'ayant pu y assister mais qui nous ont fait parvenir leurs remarques sur le questionnaire ainsi que ceux qui le feraient *a posteriori*.

ORDRE DU JOUR

Introduction et tour de table

1. Présentation de l'enquête 2020 auprès de la Génération 2017.
 - 1.1. Les principales évolutions de l'architecture, du mode de collecte et du champ (V. Henrard)
 - 1.2. Les principales évolutions du questionnaire (M. Olaria, Z. Mazari)
2. Présentation des partenariats d'extensions possibles pour la prochaine enquête et de l'appel à projets pour les modules thématiques (A. Robert, D. Epiphane)

Dans le cadre de la rénovation du dispositif Génération, cette réunion du comité est consacrée essentiellement à la présentation de l'enquête 2020 auprès de la Génération 2017, notamment les principales évolutions de l'architecture, du champ, du mode de collecte et du questionnaire. Le Céreq a envoyé aux membres du comité, en amont de la réunion, une première version du questionnaire de l'enquête 2020 auprès de la Génération 2017 (1^{ère} interrogation de la Génération 2017).

V. Henrard, cheffe du Département des entrées et des évolutions dans la vie active (Deeva), annonce l'arrivée d'Éric Durieux, nouveau chef adjoint du Deeva et responsable de l'équipe Ingénierie et gestion d'enquête.

Les membres du Céreq présentent les principales évolutions de l'architecture, du mode de collecte et du champ de l'enquête ainsi que les principales nouveautés du questionnaire et sollicitent l'avis des membres du comité sur ces évolutions (ajouts, modifications ou suppressions). Les remarques du comité sont importantes. Certaines pourront conduire à une modification du questionnaire mais elles ne pourront cependant pas toutes être prises en compte pour cette enquête. En effet, il y a une contrainte forte, celle de la maîtrise de la durée globale du questionnaire (dont l'objectif est de 30 minutes en moyenne, extensions comprises).

1. Présentation de l'enquête 2020 auprès de la Génération 2017

Lors du dernier comité de concertation qui s'est réuni le 10 mai 2017, Christophe Barret, ancien chef de l'équipe ingénierie et gestion d'enquête, avait présenté l'avancée des différents chantiers autour de la rénovation du dispositif. Plus particulièrement, il avait évoqué la révision du questionnaire et la définition du socle commun.

1.1. Les principales évolutions de l'architecture, du champ et du mode de collecte

V. Henrard présente en introduction les principales évolutions en termes d'architecture, de champ et de mode de collecte. Ces changements impactent profondément le dispositif.

Voici les principales permanences (caractéristiques du dispositif actuel qui sont conservées) :

- Génération de primo sortants de formation initiale, de tous niveaux de formation ;
- Première interrogation trois ans après la sortie ;
- La dimension longitudinale via le calendrier mensuel d'activités.

Les principales modifications envisagées peuvent être résumées ainsi :

- **Au niveau de l'architecture du dispositif**

L'alternance « Génération pleine » et « Génération légère » est abandonnée pour aller vers une architecture plus simple où se succèdent, à 4 années d'intervalle, des générations « équivalentes » interrogées à 3 et 6 ans.

C'est donc une génération sur 4 (chaque Génération est une génération « pleine » dans la mesure où il n'y a plus de petite génération) qui sera interrogée (voir support de présentation). La première interrogation aura lieu 3 ans après la sortie du système éducatif et il n'y aura qu'une seule ré-interrogation à 6 ans.

- **Au niveau du questionnaire**

Le questionnaire se décompose en deux parties, une fixe et une variable. Chaque enquête aura, ainsi, un « tronc commun » fixe auquel s'ajoutera un nombre limité de « modules » variables. Ce tronc commun est plus réduit que celui d'une génération pleine du dispositif actuel et les modules variables permettront d'approfondir des thématiques de recherche qui seront abordées de manière spécifique ou privilégiée à chaque enquête. Ces modules pourraient être assez conséquents, mais en nombre limité et différents à chaque enquête en fonction des priorités de recherche définies en amont. La définition de ce tronc commun s'est faite conjointement à la révision du questionnaire afin de prendre en compte certaines thématiques insuffisamment couvertes ou couvertes trop en détails.

- **Au niveau du champ**

Le champ a évolué afin que les critères définissant la formation initiale correspondent davantage à la période d'investissement initiale dans les études ou la formation.

L'**année de césure** est à présent incluse dans le champ de l'enquête, comme faisant partie de la période de formation initiale (accompagné d'un questionnaire spécifique) : pour cela, le critère sur l'interruption de scolarité a été modifié pour ne plus considérer hors-champ les jeunes qui ont fait une année de césure en changeant la condition d'interruption des études de « au moins un an » à « au moins 16 mois ». Depuis la rentrée 2015, les étudiants peuvent réaliser une année de césure (de 6 mois à un an) dans leur parcours tout en gardant le statut d'étudiant. L'objectif est de pouvoir capter ces années de césure ou d'interruption et d'étudier leur rôle dans le parcours des jeunes. Il s'agit également de s'adapter à une évolution de fond concernant la définition de la formation initiale et son achèvement pour les jeunes générations.

Le **contrat de professionnalisation** venant prolonger les études initiales (ou l'apprentissage) a été intégré dans le champ de l'enquête, comme relevant d'une période « en étude » et non « en emploi ». En effet, certains contrats de professionnalisation s'effectuent directement dans la continuité de la formation initiale avec une visée diplômante. Il a donc été décidé de les considérer, au même titre que l'apprentissage, comme une composante intégrante du parcours scolaire initial. Le suivi statistique de ces contrats est assuré par la Dares *via* la base Extrapro qui sera fournie au Céreq pour la constitution de notre base de sondage (une expertise de cette base a été réalisée en amont).

Les jeunes **résidant à l'étranger** au moment de l'enquête sont intégrés dans le champ de l'enquête. Il y a un intérêt évident à l'inclusion des jeunes résidents à l'étranger dans le champ de l'enquête, notamment pour les sortants de l'enseignement supérieur (docteurs, grandes écoles). Cela nous permettra d'obtenir une meilleure représentativité des parcours d'insertion.

- **Au niveau du mode de passation**

La collecte historiquement menée en monomode téléphone évolue vers une **collecte multimode internet-téléphone**. L'analyse des trois expérimentations réalisées en 2013, 2015 et 2016 permettent aujourd'hui de définir un protocole multimode adapté à l'enquête.

- **Innovations dans l'adaptation de l'outil de collecte multimode**

Il y a également des innovations dans l'adaptation de l'outil de collecte multimode sur des questions spécifiques de l'enquête. L'objectif était à la fois de collecter une information de qualité mais également de parvenir à réduire la durée de l'enquête en allégeant le questionnaire. Plusieurs pistes ont été envisagées.

La première concerne la **collecte de l'activité économique de l'établissement employeur**. La nomenclature utilisée est la NAF révision 2 sur 2 positions (88 divisions). Il semble impossible de collecter une information plus fine que cela via le questionnaire sans l'intégration d'un outil spécifique. Sur le terrain et lors du codage de l'activité *a posteriori*, il semble que les enquêtés ont des difficultés à répondre aux questions. Un nombre non négligeable d'informations en clair sont collectées pour définir précisément ce que faisait/vendait ou fabriquait l'employeur. On observe également des difficultés dans la distinction entre l'activité de l'employeur et leur profession.

C'est pourquoi, une nouvelle solution est envisagée via l'intégration d'un moteur de recherche de type « Google » à l'enquête. Une recherche de l'adresse de l'entreprise sera demandée à partir de la saisie du nom de l'entreprise et de la commune. Cette recherche s'effectue par complétion automatique. Cette fonctionnalité informatique permet à l'utilisateur de limiter la quantité d'informations qu'il saisit avec son clavier, en se voyant proposer un complément qui pourrait convenir à la chaîne de caractères qu'il a commencé à taper. Enfin, la recherche se ferait directement dans la base de données SIRENE qui

rassemble des informations économiques et juridiques sur environ 10 millions d'établissements appartenant à tous les secteurs d'activité.

Plusieurs avantages se dégagent : obtention de données potentiellement plus fiables que celles recueillies auprès des enquêtés, collecte d'une information précise sur l'activité avec un code NAF sur 5 positions, d'autres données supplémentaires notamment l'effectif de l'établissement, le chiffre d'affaire, ... et notamment un gain de temps de passation du questionnaire.

La deuxième concerne la **collecte de la profession de l'individu**. Une révision de la nomenclature PCS a été initiée par le CNIS en 2018. Des groupes de travail ont ainsi été formés : niveau fin, codage, agrégat et documentation. Le but étant de définir la nouvelle PCS 2020 avec une mise à jour de la nomenclature mais aussi de faire évoluer le système de codification automatique en supprimant certaines informations annexes (qui sont en fait des questions posées aux individus). Le Céreq participe à ces groupes de travail. Nous envisageons d'intégrer cette nouvelle méthode de collecte dans l'enquête 2020. Elle consiste à intégrer un moteur de recherche avec un menu des libellés de professions précis facilitant la codification (outil développé par l'Insee). Les avantages, comme la collecte de l'activité économique, sont la réduction du nombre de questions et la plus grande fiabilité des données collectées.

La troisième concerne la **collecte des diplômes**. Une adaptation du questionnement permettra d'être plus précis dans la définition du plus haut diplôme obtenu (PHD) en formation initiale afin de consolider les choix de codification *a posteriori*. Notamment, en interrogeant les diplômés intermédiaires et en s'assurant de la cohérence des diplômes déclarés (ou non) et du niveau de sortie à la fin des études. Pour cela, un outil de collecte embarqué (SICORE) développé par l'Insee pourrait être intégré à l'enquête. Des échanges sont en cours avec le Pôle diplômes de l'INSEE.

Enfin, afin de vérifier la compatibilité de ces applications avec l'outil de collecte, des tests sont prévus courant 2019 pour attester de la faisabilité.

Un calendrier de l'enquête détaillé présentant les points d'étapes de sa mise en œuvre est disponible dans le support de présentation pour tenir informé le comité des échéances à venir.

1.2. Les principales évolutions du questionnaire

M. Olaria et Z. Mazari présentent les principales évolutions du questionnaire tant au niveau du questionnement que sur le volet technique et ergonomique de l'outil de collecte multimode.

Pour résumer, ci-dessous un tableau récapitulatif des principales évolutions en termes de thématiques de questionnement :

	THEMATIQUES MAINTENUES	NOUVELLES THEMATIQUES	THEMATIQUES SUPPRIMEES*
Parcours de formation initiale	Arrêt des études	Financement des études	Professionnalisation dans l'enseignement supérieur (extension)
	Diplôme de sortie, plus haut diplôme obtenu, diplôme(s) intermédiaire(s)	Double diplomation	Séjours à l'étranger (extension)
	Orientation		Stage(s) en cours d'études
	Expérience(s) de travail en cours d'études		Modules post initiaux (extension)
			Module thèse (extension)
		Développement durable (extension)	
Parcours professionnel	Calendrier d'activité mensuel (description de tous les emplois)	Multi-activité (indépendant, auto-entrepreneuriat, intermittent du spectacle, etc.)	Calendrier d'activité mensuel : description des situations de non emploi (seules quelques informations sont conservées dans le cadre du pop-up)
	Opinion sur l'emploi	Intermédiaires marché du travail	Formation dans l'entreprise
		Acquisition de nouveaux diplômes, certifications (PHD à 3 et 6 ans)	
Caractéristiques de l'individu	Origine sociale	Mobilités résidentielles	Secteur d'activité des parents
	Discrimination		Caractéristiques de la fratrie
	Perspectives professionnelles		Module sur le handicap (extension)
	Calendrier du mode de cohabitation		

* Une priorisation des thématiques a été réalisée pour définir le tronc commun de l'enquête. Les modules supprimés à ce stade peuvent toutefois faire l'objet d'une commande par des partenaires d'extensions.

Échanges avec les membres du comité sur le questionnaire de l'enquête 2020 :

G. Boudesseul pose la question du choix de la durée de l'interruption des études, notamment liée à l'année de césure, impactant le champ de l'enquête.

L'analyse des suspensions d'études dans Génération 2013 à travers l'exploitation du questionnaire filtre (avant la présumée primo-sortie) et l'examen des reprises d'études (après la primo-sortie) nous a conduit à élargir la fenêtre de la primo sortie à 17 mois (au lieu de 12 mois auparavant) afin d'observer les suspensions d'études équivalentes à une année scolaire/universitaire (dont les années de césure).

G. Boudesseul questionne sur l'adaptation de l'enquête sur les smartphones et les tablettes notamment au niveau des menus déroulants.

L'enquête sera bien adaptée à la réponse sur smartphones et tablettes notamment avec des menus en auto-complétion et non des menus déroulants classiques.

G. Boudesseul questionne sur la collecte des MFR (Maisons Familiales Rurales) qui forment chaque année 70 000 jeunes et adultes.

Les maisons familiales rurales d'éducation et d'orientation font partie du champ Céreq (environ 350 maisons sont concernées).

G. Boudesseul pose également la question de la représentativité de cette enquête renouvelée par région.

L'enquête Génération n'a pas pour but d'être représentative au niveau régional, mais au niveau national. Seuls quelques grands agrégats peuvent être calculés régionalement. Le financement extérieur d'extensions régionales est néanmoins possible afin d'étendre la représentativité.

G. Boudesseul pose la question de savoir si le code INSEE des communes sera disponible dans le tableau des mobilités résidentielles.

L'enquête ne remplira dans ce tableau que le code postal et le nom de la commune mais en post-enquête le code INSEE sera bien disponible.

F. Tardieu regrette la suppression des questions sur le secteur d'activité (public / privé) du père et de la mère dans leur dernier emploi au moment de la fin des études de l'enquête.

Pour définir le socle commun, des arbitrages ont été réalisés pour alléger le module « Caractéristiques individuelles et environnement familial ». Ces questions étaient trop peu exploitées et d'autres sont apparues plus pertinentes à maintenir. Ces deux questions étaient déjà absentes du questionnaire de la première interrogation de la Génération 2013 et l'étaient donc également du module « Concours et attractivité de la fonction publique ». Ces questions peuvent toutefois être intégrées à un module d'extension thématique.

F. Tardieu souligne la disparition des questions sur la fratrie.

Comme pour le secteur d'activité des parents, cette série de questions peu utilisée dans les études menées au Céreq sont supprimées du tronc commun. Elles peuvent être intéressantes dans le cadre d'un module d'extension thématique sur la famille où combinées avec d'autres, elles permettraient de mieux appréhender ce sujet.

F. Tardieu demande pourquoi il n’y a pas de questions sur le handicap.

Le module présent dans le questionnaire à 3 ans de l’enquête Génération est commandé par un partenaire historique d’extension. Il n’a pas été conservé dans le questionnaire de tronc commun notamment car il s’agit de données sensibles trop éloignées de notre cœur de questionnaire. Seules, deux questions ont été préservées :

SD200

ASK ALL

Comment est votre état de santé en général ?

Très bon	1
Bon.....	2
Mauvais	3
Très mauvais.....	4
Ne veut pas dire.....	5

SD210

ASK IF SD200=(1-4)

Avez-vous néanmoins un problème de santé durable ou un handicap ?

Oui	1
Non	2
Ne veut pas dire.....	3

J. Timotéo se réjouit du retour du module « Financement des études ». Il ajoute également que le module sur les mobilités résidentielles permettra d’approfondir les dimensions territoriales (demande thématique très forte).

Une remarque toutefois sur la première question du module sur le financement des études :

FE010 - QCM

ASK ALL

De quels types de ressources avez-vous pu bénéficier pendant vos études par exemple pour votre logement, vos dépenses quotidiennes, vos frais d’inscription etc. ?

Bourses sur critère sociaux.....	Oui=1/Non=0
Bourses sur critère scolaire\universitaire, bourse au mérite.....	Oui=1/Non=0
Prêts bancaires	Oui=1/Non=0
Emploi(s) rémunéré(s) en cours d’études, déclarés ou non	Oui=1/Non=0
Contrats en alternance	Oui=1/Non=0
Aides financières ou matérielles de la famille.....	Oui=1/Non=0
Autre, précisez.....	Oui=1/Non=0

A. Testas fait remarquer que la période « pendant vos études » est assez vague et des difficultés peuvent apparaître dans l’analyse pour déterminer à quel moment du parcours de formation les ressources sont perçues. Une suggestion serait de dater la question.

J. Timotéo fait remarquer que les modalités « Bourses sur critères sociaux » et « Bourses sur critère scolaire\universitaire » semblent proches. La question est de savoir si les enquêtés sauront faire la différence en enquête. Une proposition serait de définir plus précisément ces bourses dans des infobulles.

Il est demandé des précisions concernant la définition de l'emploi régulier : première question du module expérience(s) de travail en cours d'études.

La définition reste inchangée par rapport au questionnaire actuel : au moins 8 heures par semaine.

A. Testas, tout comme pour les ressources perçues pendant les études, suggère de dater cette activité. Pour finir, elle propose de repositionner la question EXP130 sur le nombre d'heures par semaine à la suite de la question EXP010.

Suggestion acceptée.

J. Timotéo formule une autre remarque sur les situations de volontariat. Il propose de dater l'engagement volontaire dans le cadre du service civique, pour la rattacher éventuellement à un épisode du calendrier d'activité.

Il souligne également, pour les prochaines enquêtes, l'arrivée du service national universel (SNU), il faudra s'interroger sur son impact sur les trajectoires des jeunes.

H. Garner demande des informations concernant les groupements d'employeurs dans le cadre du repérage des nouvelles formes d'emploi.

EXP120

ASK IF CFANEW=0 AND EXP010=1

Étiez-vous rémunéré par un groupement d'employeurs ou par l'intermédiaire d'une plateforme (de type Uber ou Deliveroo) ?

Oui1
Non.....2

H. Garner souligne que la notion de groupement d'employeurs est un cas spécifique et concerne plutôt l'emploi salarié. Peut-être ne mentionner que « par le biais d'un intermédiaire type plateforme (ex : Uber, Deliveroo...) ».

T. Razafindranovona précise qu'il y a un problème d'écriture du champ de la question (non identique partout). Il propose également de se renseigner auprès de la division emploi pour nous apporter des éléments de précision sur ce sujet.

Il est proposé d'ajouter, dans le module de questions sur la multi activité, le temps passé dans cette autre activité et de faire référence à un montant de salaire perçu plutôt qu'à un pourcentage du revenu global (difficile à calculer).

T. Razafindranovona demande si des tests de certaines parties du questionnaire sont prévus.

Le lancement du marché d'appel d'offre européen sur la collecte est avancé par rapport au calendrier habituel (15 décembre prochain). Ce changement a plusieurs objectifs : réaliser 3 tests au lieu de 2 pour notamment fournir une durée moyenne de questionnaire au prochain comité du label (en septembre 2019), mais aussi pour permettre au prestataire sélectionné de développer un outil de collecte multimode et de mettre en place le protocole dans un délai raisonnable.

Une réflexion est en parallèle menée pour réaliser un test *Youthproofing*, toujours avec le prestataire de collecte, qui ferait intervenir les chargés d'études et des volontaires pour tester le questionnaire et recueillir des avis et difficultés de compréhension rencontrées au moment de l'enquête. Ce test peut aussi contenir un volet en auto-administré.

Des précisions sont demandées quant à la mise en place de l'application de recherche de l'activité de l'établissement employeur.

Avec notre expérience dans la collecte d'informations via des menus par complétion automatique et la récente mise à disposition des bases de données entreprises SIRENE, le développement d'une application externe pourrait enrichir les données de Génération par appariement. L'objectif premier est de récupérer l'activité de l'entreprise employeur. En effet, dans le questionnaire actuel, il existe une batterie de questions pour tenter d'identifier l'activité principale de l'employeur. De grandes difficultés pour se positionner sont rencontrées par les enquêtés.

Cette application externe ferait l'objet d'un marché. L'idée est de faciliter la recherche de l'établissement employeur via le nom de l'entreprise et le code postal et enfin d'en sélectionner l'adresse précise.

J. Ponceau pose une question plus générale sur le changement du mode de collecte, avec l'introduction d'internet comme mode alternatif. À savoir, quelles sont les méthodes d'agrégation pour les futures analyses.

Il n'existe pas de consensus dans la littérature sur la manière d'agrèger les données provenant de plusieurs modes de collecte. Certains méthodologues pensent qu'il n'y a pas de mode de collecte idéal et que par conséquent une agrégation simple des données des différents modes de collecte donne une meilleure image de la réalité, les erreurs de mesure d'un mode de collecte étant compensées par celles de l'autre. Pour d'autres, cette stratégie peut s'avérer dangereuse si les erreurs ne se compensent pas mais s'ajoutent. Dans ce cas, il serait préférable de considérer qu'il y a un mode de collecte prépondérant et de corriger les effets de mesure pour faire correspondre les réponses du mode de collecte alternatif.

L'équipe Ingénierie et gestion d'enquête, sur la base des enquêtes expérimentales réalisées en parallèle des dernières enquêtes, a analysé les possibilités d'agrégation et compare les différentes approches pour tenter de corriger le biais lié à l'effet de mesure¹.

Il est demandé des explications sur le protocole de collecte qui sera mis en œuvre pour la prochaine enquête. Est-ce qu'il y a un mode de collecte prioritaire ?

Les différentes expérimentations mise en place depuis 2013 ont permis de tester différents protocoles de collecte². Tout d'abord, monomode internet puis multimode internet-téléphone. Elles ont aussi permis de mesurer l'efficacité d'un protocole multimode séquentiel et/ou concurrentiel.

¹ Barret C. et Cissé M. (2018), Agrégation de données multimode : impact sur la modélisation des variables présentant un effet de mesure, Journées de Méthodologie Statistique (JMS), Paris.

Barret C. et Dzikowski C. (2016), Évaluation d'un protocole multimode avec échantillon embarqué et agrégation des données en présence d'effet de mesure, 9e Colloque Francophone sur les sondages, Gatineau, Canada.

² Barret C., Cissé M., Gaubert E., Mazari Z. et Olaria M. (2018), Efficacité d'un protocole multimode (téléphone et internet), Journées de Méthodologie Statistique (JMS), Paris.

Le protocole de collecte multimode retenu est construit autour de plusieurs phases de collecte. Dans un premier temps, l'objectif est d'obtenir des questionnaires par internet uniquement avec l'envoi de lettre-avis et de relances mails. À l'issue de cette première phase, les relances téléphoniques démarrent avec une incitation à répondre par internet tout en continuant les relances par mail. Enfin, une dernière phase dite de « rattrapage » vise à contacter tous les individus n'ayant pas encore répondu pour réaliser l'enquête par téléphone (en cas de refus, la possibilité de répondre par internet était toujours offerte). Pendant toute la collecte, les enquêtés ont la possibilité de reprendre un questionnaire commencé en changeant ou non de mode de réponse.

J. Timotéo fait remarquer que les comparaisons avec les enquêtes de l'ancien dispositif ne seront pas simples notamment sur certaines thématiques.

Il y aura une rupture de série avec notamment l'introduction du mode de collecte internet.

H. Garner demande si les analyses en termes d'analyses de trajectoires seront toujours possibles ? Par exemple, calculer un temps d'accès au premier EDI ?

Avec le nouveau calendrier d'activité, la description détaillée des emplois est maintenue (contrat de travail, profession, salaire...). Il sera toujours possible de faire des analyses de trajectoires.

La question du caractère obligatoire de l'enquête (label Cnis) est abordée, ainsi que l'impact que cela pourrait avoir sur le taux de réponse.

Le Céreq réfléchit à l'opportunité du caractère obligatoire pour les enquêtes Génération. À ce stade, s'il est envisagé de le demander, ce sera essentiellement pour l'effet espéré lors de la collecte des fichiers d'élèves auprès des établissements de formation en vue de la constitution de la base de sondage.

2. Présentation de l'appel à projets sur les modules thématiques

A. Robert et D. Epiphane présentent les possibilités d'extensions de l'enquête ; extension de champ, d'échantillon et/ou de questionnement au travers d'un appel à projets sur des modules thématiques en lien avec la jeunesse et l'insertion professionnelle. Ci-dessous, le PowerPoint de présentation. Vous pouvez retrouver les informations concernant les extensions ainsi que l'appel à projets pour les modules de questionnement thématique sur cette page : [lien](#).

Discussions autour des possibilités d'extensions et de l'appel à projets :

F. Tardieu demande si la base de sondage est exhaustive dans la collecte des établissements. Quel est le taux de couverture ?

Il n'existe pas en France de base nationale recensant l'ensemble des sortants une année donnée. Le Céreq doit donc la constituer en collectant soit les bases centralisées ministérielles quand elles existent, soit en passant par une collecte des fichiers d'élèves par le biais d'un prestataire. L'enquête mise en place n'est actuellement pas obligatoire. Elle présente un défaut de couverture dû aux établissements qui n'ont pas fourni de base d'élèves lors de la constitution de la base de sondage. La structure des établissements répondants fait l'objet d'un suivi précis en fonction des caractéristiques des établissements (c'est à dire selon les variables de stratification utilisées pour le tirage des individus). Les calculs des taux de couverture sur la base de sondage de l'enquête Génération 2017 ne sont pas encore disponibles mais à titre indicatif, la couverture en termes d'effectifs d'établissements de la base de sondage de la Génération 2013 était de l'ordre de 87%, par rapport à une base de sondage pour laquelle l'ensemble des établissements de formation nous aurait fait parvenir leurs fichiers d'élèves.

Nous réfléchissons à l'opportunité de demander le caractère obligatoire de l'enquête notamment pour pallier cette incertitude de transmission des fichiers.

V. Henrard clôt la séance du comité en remerciant l'ensemble des membres présents pour la qualité et la richesse des échanges. La date du prochain comité sera fixée au premier semestre 2019.

Annexe : Liste des membres du comité de concertation

NOM	ORGANISME
Catherine BEAUVOIS	Premier ministre/CNEFOP
Maria MELICIANO	Premier ministre/CNEFOP
Hélène GARNER	Premier ministre/France stratégie
Benoit BOURGES	Ministère de la transition écologique et solidaire/CGDD
Sophie MARGONTIER	Ministère de la transition écologique et solidaire/CGDD
Eric PAUTARD	Ministère de la transition écologique et solidaire/CGDD
Hélène CHAPUT	Ministère des solidarités et de la santé/DREES
Yannick CROGUENNEC	Ministère des solidarités et de la santé/DREES
Claire MARBOT	Ministère des solidarités et de la santé/DREES
Rémy MARQUIER	Ministère des solidarités et de la santé/DREES
Marine GUILLERM	Ministère du travail/DARES
Benjamin NEFUSSI	Ministère du travail/DARES
Anne-Sophie PICHAVANT	Ministère du travail/DARES
Cindy REIST	Ministère du travail/DARES
Elodie GUIRCHOUN	Ministère du travail/DGEFP
Séverine LEGUILLON	Ministère du travail/DGEFP
Roselyne KERJOSSE	Ministère de l'éducation nationale et de la jeunesse/DEPP
Hélène MICHAUDON	Ministère de l'éducation nationale et de la jeunesse/DEPP
Anna TESTAS	Ministère de l'éducation nationale et de la jeunesse/DEPP
Arnaud LACOURT	Ministère de l'éducation nationale et de la jeunesse/DGESCO
Nicolas TARIEL	Ministère de l'éducation nationale et de la jeunesse/DGESCO
Brigitte TROCME	Ministère de l'éducation nationale et de la jeunesse/DGESCO
Christine REBIERE	Ministère de l'éducation nationale et de la jeunesse/DGESCO
Isabelle DUSSOUET	Ministère de l'éducation nationale et de la jeunesse/Ministère de l'enseignement supérieur, de la recherche et de l'innovation/Onisep
Adrien FRIEZ	Ministère de l'action et des comptes publics/DGAFP
Claire HAGEGE	Ministère de l'action et des comptes publics/DGAFP
Frédéric TARDIEU	Ministère de l'action et des comptes publics/DGAFP
Christine BRUNIAUX	Ministère de l'enseignement supérieur, de la recherche et de l'innovation/DGESIP
Catherine THIBault	Ministère de l'enseignement supérieur, de la recherche et de l'innovation/DGESIP
Didier ROUX	Ministère de l'enseignement supérieur, de la recherche et de l'innovation/DGESIP
Jérôme HARNOIS	Ministère de l'enseignement supérieur, de la recherche et de l'innovation/DGESIP/SIES
Isabelle KABLA-LANGLOIS	Ministère de l'enseignement supérieur, de la recherche et de l'innovation/DGESIP/SIES
Boris MÉNARD	Ministère de l'enseignement supérieur, de la recherche et de l'innovation/DGESIP/SIES
Hélène ROBIC	Ministère de l'enseignement supérieur, de la recherche et de l'innovation/DGESIP/SIES
Pascal DIEUSAERT	Ministère de la cohésion des territoires et des relations avec les collectivités territoriales/CGET
Joaquim TIMOTEO	Institut national de la jeunesse et de l'éducation populaire (INJEP)
Geoffrey LEFEBVRE	Institut national de la jeunesse et de l'éducation populaire (INJEP)
Christel ALIAGA	Institut national de la statistique et des études économiques (INSEE)
Olivier CHARDON	Institut national de la statistique et des études économiques (INSEE)
Tiaray RAZAFINDRANOVONA	Institut national de la statistique et des études économiques (INSEE)
Chloé TAVAN	Institut national de la statistique et des études économiques (INSEE)
Céline GASQUET	Observatoire régional des métiers (ORM PACA)
Anita BONNET	Pôle emploi
Stéphane DUCATEZ	Pôle emploi
Frédéric LAINÉ	Pôle emploi
Murielle MATUS	Pôle emploi
Nicolas PROKOVAS	Pôle emploi

Sylvain DESOIGNIES	Confédération française démocratique du travail (CFDT)
Patricia FERRAND	Confédération française démocratique du travail (CFDT)
Inès MININ	Confédération française démocratique du travail (CFDT)
Magali BOURDON	Confédération générale du travail (CGT)
Sylvia VEITL	Force ouvrière (FO)
Audrey BASLY	Mouvement des entreprises de France (Medef)
Vincent GOMEZ	Union nationale des syndicats autonomes (UNSA)
Jean-Marie TRUFFAT	Union nationale des syndicats autonomes (UNSA)
Alexandre LEROY	Fédération des associations générales étudiantes (FAGE)
Amélie BRIFFAUX	Réseau des observatoires de l'enseignement supérieur (RESOSUP)
Sylvie VILTER	Réseau des observatoires de l'enseignement supérieur (RESOSUP)
Myriam BOBBIO	Union nationale des associations familiales (UNAF)
Yvon SERIEYX	Union nationale des associations familiales (UNAF)
Catherine BEDUWE	Centre associé régional au Céreq (CAR Université Toulouse I)
Gérard BOUDESSEUL	Centre associé régional au Céreq (CAR Université Caen)
Jean-François GIRET	Centre associé régional au Céreq (CAR Université de Bourgogne)

A.4 – Publications récentes

Voici une liste des publications mobilisant l'enquête 2013 auprès de la Génération 2010, classées par thématique.

- **Premiers résultats, travaux de synthèse :**

Enquête 2013 auprès de la Génération 2010 - Face à la crise, le fossé se creuse entre niveaux de diplôme, Christophe Barret, Florence Ryk, Noémie Volle, Céreq Bref, n°319, 2014, 8 p.

2013 survey of the 2010 cohort: The gap between levels of qualification widens as the crisis deepens, Christophe Barret, Florence Ryk et Noémie Volle, Training and Employment, n° 109-110, 2014, 8 p.

Entrer dans la vie active, Alternatives Economiques, hors-série poche n°71, janvier 2015.

Quand l'école est finie. Premiers pas dans la vie active, Pascale Rouaud et Olivier Joseph (coordination), Céreq Ouvrages, 2014, 92 p.

Quand l'école est finie en Paca, Etudes n°26, ORM Paca, 2015

Quand l'école est finie en Nord-Pas-de-Calais – Génération 2010, C2RP, 2015

Parcours de formation, d'insertion et de réussites : nouveaux regards à partir de Génération 2010, Julien Calmand, Thomas Couppié et Valentine Henrard, Formation Emploi 2018/2 (n° 142), p. 7- 14

Un nouveau panorama de l'insertion des jeunes, Yannick L'Horty, Formation Emploi 2018/2 (n° 142), p. 187-198

Rendement éducatif, parcours et inégalités dans l'insertion des jeunes. Recueil d'études sur la Génération 2010, Julien Calmand, Thomas Couppié, Valentine Henrard (coord.), Céreq Echanges, n°5, 2017, 386p.

20 ans d'insertion professionnelle des jeunes, entre permanences et évolutions, coord. Thomas Couppié, Arnaud Dupray, Dominique Epiphane, Virginie Mora, Céreq Essentiels n°1, avril 2018, 196 p.

Des débuts de carrière plus chaotiques pour une génération plus diplômée - Génération 2010, Dominique Epiphane, Zora Mazari, Manon Olaria, Emmanuel Sulzer, Céreq Bref, n°382, 2019, 4p.

Enquête 2016 auprès de la Génération 2013 - Pas d'amélioration de l'insertion professionnelle pour les non-diplômés, Emilie Gaubert, Valentine Henrard, Alexie Robert, Pascale Rouaud, Céreq Bref, n°356, 2017, 4p.

L'insertion professionnelle des jeunes, Valentine Henrard et Alexie Robert, L'état de l'école 2018, Direction de l'évaluation, de la prospective et de la performance, Ministère de l'Education nationale, de l'Enseignement supérieur et de la Recherche, 2015.

▪ **Sur les sortants de l'enseignement supérieur :**

Sortants du supérieur : la hausse du niveau de formation n'empêche pas celle du chômage Boris Ménard Céreq Bref, n°322, 2014, 4 p.

Leavers from higher education: increased levels of education have not prevented a rise in unemployment, Boris Ménard, Training and Employment, n° 111, 2014, 4 p.

L'expansion scolaire a-t-elle modifié les conditions d'insertion des diplômés du supérieur ? Philippe Lemistre, Boris Ménard. In 20 ans d'insertion professionnelle etc. op. cit.

Faire des études supérieures, et après ? Enquête Génération 2010 - Interrogation 2013, Julien Calmand, Boris Ménard, Virginie Mora, Nef , n° 52 , 2015 , 60 p.

La place du doctorat sur le marché du travail a-t-elle évolué durant ces 20 dernières années ? Julien Calmand, In 20 ans d'insertion professionnelle etc. op. cit.

Les enquêtes sur le devenir professionnel des docteurs, état des lieux et usages, Julien Calmand, in Magali Jaoul-Grammare, Simon Macaire, (coord) Étudier le devenir professionnel des docteurs, Cereq Echanges, n°2, 2016.

Parcours d'étude et insertion professionnelle des docteurs : quand doctorat n'est pas synonyme d'emploi, Magali Jaoul-Grammare, in Magali Jaoul-Grammare, Simon Macaire, (coord) Étudier le devenir professionnel des docteurs, Cereq Echanges, n°2, 2016.

L'insertion à trois ans des docteurs diplômés en 2010, Julien Calmand, Net.Doc , n° 144, 2015, 109 p.

Les débuts de carrière des docteurs : une forte différenciation des trajectoires professionnelles, Calmand Julien, Prieur Marie-Hélène , Wolber Odile, Céreq Bref, n°354, 2017, 4p.

Les cinq premières années de vie active des docteurs diplômés en 2010, Calmand Julien, Céreq Etudes, n°9, 2017, 53p.

(In)fortunes professionnelles des femmes à la sortie d'une filière universitaire masculine, les STAPS,

Carine Érard et Christine Guégnard, Formation Emploi 2018/2 (n° 142), p. 79-98

Ombres et lumières à propos des parcours d'insertion des jeunes de la filière STAPS Christine Guégnard, Carine Erard, in Rendement éducatif, parcours et inégalités dans l'insertion etc., op. cit.

Alice au pays du Sport, Christine Guégnard, Carine Erard in Les transitions professionnelles tout au long de la vie etc. op. cit.

Formations environnementales : qualité de l'insertion et ressenti sur les débouchés professionnels, Sophie Margontier, Eric Pautard, Le point sur, n°199, avril 2015.

Les sortants sans diplôme de l'enseignement supérieur en 2010 : tous décrocheurs ? Une analyse à l'aune de l'approche des capacités, Boris Ménard in Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc., op. cit.

Quitter l'enseignement supérieur sans diplôme : de la filière d'inscription au parcours d'insertion, Fanette Merlin, in Jeunesse(s) et transitions vers l'âge adulte etc. op. cit.

Le « décrochage » en STS : l'autre échec dans l'enseignement supérieur, Fanette Merlin, Céreq Bref, n°366, 2018, 4p.

Évolution des inégalités de rendements éducatifs des diplômés du supérieur. Une comparaison Génération 1998 - Génération 2010, Magali Jaoul-Grammare, in Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc., op. cit.

L'état de l'Enseignement supérieur et de la Recherche, MENESR, n°8, 2015 (participation à 3 fiches)

L'état de l'Enseignement supérieur et de la Recherche, MENESR, n°9, 2016 (participation à 3 fiches)

L'état de l'Enseignement supérieur et de la Recherche, MENESR, n°10, 2017 (participation à 2 fiches)

L'état de l'Enseignement supérieur et de la Recherche, MENESR, n°11, 2018 (participation à 2 fiches)

▪ **Sur les sortants de l'enseignement secondaire :**

L'entrée sur le marché du travail des jeunes de la voie professionnelle rénovée, Ilardi Valérie, Joseph Olivier, Sulzer Emmanuel, Céreq Bref, n°365, 2018, 4p.

Le « bac pro ». Élargissement des publics et permanences des fonctions sociales. Emmanuel, Sulzer, Valérie Ilardi. In 20 ans d'insertion professionnelle des jeunes etc. op. cit.

S'insérer dans l'emploi et s'engager dans le travail : parcours types de réussite des sortant.es de l'enseignement professionnel, Estelle Bonnet, Karine Pietropaoli et Élise Verley, Formation Emploi 2018/2 (n° 142), p. 143-166

De la « qualité de l'emploi » au « rapport au travail » des jeunes. Estelle Bonnet, Zora Mazari, Elise Verley. In 20 ans d'insertion professionnelle des jeunes etc. op. cit.

« Demain, c'est loin ». Transitions socioprofessionnelles des jeunes non diplômés. Dominique, Maillard, Pascale, Rouaud. In 20 ans d'insertion professionnelle des jeunes etc. op. cit.

Parcours d'insertion et de réussite des diplômés de l'enseignement secondaire professionnel, Estelle Bonnet, Karine Pietropaoli, Elise Verley, in Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc., op. cit.

CAP-BEP : des difficultés d'insertion encore aggravées par la crise, Valérie Ilardi, Emmanuel Sulzer, Céreq Bref, n°335, 2015, 4 p.

2013 survey of the 2010 cohort: Crisis makes school-to-work transition for CAP-BEP holders even more difficult, Valérie Ilardi, Emmanuel Sulzer, Training and Employment, n° 116, 2015, 4 p.

Etude sur le niveau V dans les champs du sport et de l'animation - Le Brevet d'aptitude professionnelle d'assistant animateur technicien, Mickaële Molinari-Perrier, Net.Doc, n°147, 2016, 94p.

Les métiers de la médiation et de l'intervention sociale et la place du baccalauréat professionnel Services de proximité et vie locale, Françoise Kogut-Kubiak, Valérie Gosseume, Net.Doc, n° 143, 2015, 92 p.

L'évolution de l'enseignement professionnel : des segmentations éducatives et sociales renouvelées ?, Éric Verdier, Noémie Olympio, Vanessa Di Paola, Stéphanie Moullet et Aziz Jellab, rapport pour le conseil national d'évaluation du système scolaire, octobre 2015.

▪ **Sur la professionnalisation des formations :**

La professionnalisation peut-elle ignorer l'adéquation formation emploi ?, Catherine Béduwé,, in Catherine Béduwé, Catherine Soldano, Sandrine Croity-Belz La professionnalisation peut-elle ignorer l'employabilité ?, Céreq Echanges, n°8, 2018, 134p.

Ségrégations et parcours professionnalisés des sortants du supérieur. Philippe Lemistre, Populations vulnérables, 5, 2018, 115-135.

À qui profite la professionnalisation des parcours ?, Philippe Lemistre in Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc., op. cit.

De la professionnalité des étudiants à leur employabilité, n'y a-t-il qu'un pas ?, Catherine Béduwé, Virginie Mora, Formation Emploi, 138, 2017, p. 59-77.

Professionnalisation des publics et des parcours à l'université. Groupe de travail sur l'enseignement supérieur, Philippe Lemistre, Virginie Mora, Céreq Echanges, n°3, 2016, 166p.

Accompagner les étudiants dans leur professionnalisation : le rôle de l'université, Philippe Lemistre, Boris Ménard, in Professionnalisation des publics et des parcours à l'université, etc., op. cit

Comment se professionnalisent les étudiants ?, Catherine Béduwé, Virginie Mora, in Professionnalisation des publics et des parcours à l'université, etc., op. cit

Alternance et professionnalisation : des atouts pour les parcours des jeunes et les carrières ?. Gérard Boudesseul, Benoit Cart, Thomas Couppié, Jean-François Giret, Philippe Lemistre, Marie-Hélène Toutin, Patrick Werquin (Editeurs), Relief n°50, Céreq, 2015.

De bac moins 3 à bac plus 3 : le sens de la professionnalisation des filières, au vu des parcours et de l'insertion, Philippe Lemistre in Alternance et professionnalisation etc. op. cit.

Comment se professionnalisent les étudiants de l'enseignement supérieur ? Catherine Béduwé, Virginie Mora in Alternance et professionnalisation etc. op. cit.

▪ **Sur l'apprentissage :**

Insertion des apprentis : un avantage à interroger, Lopez Alberto, Sulzer Emmanuel, Céreq Bref, n°346, 2016, 4p.

Les apprentis de l'enseignement supérieur : diversité des usages et des logiques, Benoit Cart, Nathalie Jacob, Alexandre Léné, in Alternance et professionnalisation etc. op. cit.

A l'aube de la réforme de la formation professionnelle, retour sur 20 ans d'insertion des apprentis, Benoit Cart, Alexandre Léné, Marie-Hélène Toutin-Trelcat, Céreq Bref, n°370, 2018, 4p.

L'apprentissage favorise-t-il toujours l'insertion professionnelle ? Benoit Cart, Alexandre Léné, Marie-Hélène Toutin-Trelcat. In 20 ans d'insertion professionnelle des jeunes etc. op. cit.

Analyse des écarts d'insertion des apprentis et des scolaires issus de l'enseignement secondaire professionnel dans trois régions, Hélène Couprie, Arnaud Dupray, in Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc., op. cit.

Apprentissage et voie scolaire : comment expliquer les différences d'insertion observées ?, Thomas Couppié, Céline Gasquet, in Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc., op. cit.

L'accès à la formation en apprentissage : une question de ressources régionales ?, Manon Garrouste, Francis Kramarz et Carmelo Zizzo, Formation Emploi 2018/2 (n° 142), p. 15-33

Comment l'apprentissage favorise-t-il l'insertion professionnelle des CAP-BEP ?, Thomas Couppié et Céline Gasquet, Formation Emploi 2018/2 (n° 142), p. 35- 56

Le choix de la formation en apprentissage : une question de ressources régionales ?, Manon Garrouste, Francis Kramarz, Carmelo Zizzo, in Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc., op. cit.

L'apprentissage est-il un accélérateur du passage à l'âge adulte ?, Benoit Cart, Alexandre Léné, Marie-Hélène Toutin, in Jeunesse(s) et transitions vers l'âge adulte etc. op. cit.

- **Sur les parcours en formation, professionnels et de vie :**

À diplôme du supérieur donné, des inégalités d'accès à la qualification renforcées par le parcours d'études ?, Philippe Lemistre, in Bernard Pierre-Yves, Caillaud Pascal, Ghaffari Sarah, Gosseume Valérie, Houdeville Gérald, Hugree Cédric, Michaut Christophe, Poullaouec Tristan, Roupnel-Fuentes Manuella, Boudesseul Gérard, Couppié Thomas, Epiphane Dominique, Giret Jean-François, Werquin Patrick, Qualifications et parcours - Qualification des parcours, Céreq Echanges, n°10, 2019, 380p.

Jeunesse(s) et transitions vers l'âge adulte : quelles permanences, quelles évolutions depuis 30 ans ?, Gérard Boudesseul, Julien Calmand, Thomas Couppié, Arnaud Dupray, Dominique Epiphane, Jean-François Giret, Valérie Ilardi, Olivier Joseph, Emmanuel Sulzer, Patrick Werquin, Céreq Echanges, n°6, 2018, 384p.

Influence du parcours d'études sur l'insertion des sortants de l'université, Fanette Merlin, Philippe Lemistre, in Jeunesse(s) et transitions vers l'âge adulte etc. op. cit.

Le décrochage dans l'enseignement supérieur à l'aune de l'approche par les capacités, Boris Ménard, Formation Emploi 2018/2 (n° 142), p. 119-141

L'usage que font les étudiants du travail étudiant a-t-il évolué en 20 ans ?, Catherine Béduwé, Jean-François Giret, in Jeunesse(s) et transitions vers l'âge adulte etc. op. cit.

S'insérer dans une entreprise où l'on travaillait pendant les études, Laurence Lizé et Géraldine Rieucou, *Formation Emploi* 2018/1 (n° 141), p. 153-173

Le travail salarié des étudiants en France. Catherine Béduwé, Jean-François Giret. In *20 ans d'insertion professionnelle des jeunes etc.* op. cit.

Travailler pendant ses études et s'insérer dans la vie active : premières tendances et résultats (Génération 1998, 2004 et 2010), Laurence Lizé, Géraldine Rieucou, in *Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc.*, op. cit.

Partir travailler à l'étranger après les études, un atout pour l'insertion professionnelle des jeunes ?, Julien Calmand, in *Jeunesse(s) et transitions vers l'âge adulte etc.* op. cit.

Séjours à l'étranger en cours d'études, une plus-value sur l'insertion en France ?, Julien Calmand, Pascale Rouaud, Emmanuel Sulzer, *Céreq Bref*, n°348, 2016, 4p.

Expériences à l'étranger en cours d'études et insertion : des liens complexes, pour quelle plus-value ?, Julien Calmand, Stéphanie Condon, Karine Pietropaoli, Pascale Rouaud et Emmanuelle Santelli, *Formation Emploi* 2018/2 (n°142), p. 57-77.

Les liens entre expériences à l'étranger en cours d'études et insertion. Une approche par le concept de « capital de mobilité », Julien Calmand, Stéphanie Condon, Karine Pietropaoli, Pascale Rouaud, Emmanuelle Santelli, in *Rendement éducatif, parcours et inégalités dans l'insertion etc.*, op. cit.

S'installer chez soi : avec quel emploi et quel salaire ?, Alexie Robert, Emmanuel Sulzer, in *Jeunesse(s) et transitions vers l'âge adulte etc.* op. cit.

Trajectoires résidentielles et professionnelles des jeunes : quand l'accès à l'emploi ne signe pas la décohabitation, Philippe Cordazzo, *Formation Emploi* 2018/2 (n° 142), p. 167-185

Les trajectoires d'entrée dans la vie adulte des jeunes en période de crise, Philippe Cordazzo, in *Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc.*, op. cit.

Quels sont les profils des jeunes qui ne s'inscrivent pas dans des situations d'emploi durable et progressif ?, Anne-Sophie Dumortier, Cécile Réveille-Dongradi, Camille Stephanus, in *Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc.*, op. cit.

Retours précoces sur la voie des diplômés : vers une formation « tout au long du début de la vie » ?, Virginie Mora, Alexie Robert, *Céreq Bref*, n°360, 2017, 4p.

Obtenir un nouveau diplôme après avoir arrêté ses études : logiques hétérogènes et résultats inégaux, Virginie Mora, Alexie Robert in *Les transitions professionnelles tout au long de la vie etc.* op. cit.

Le profil des NEET a-t-il évolué en 20 ans ?. Christine Guégnard, Magali Danner, Olivier Joseph. In *20 ans d'insertion professionnelle des jeunes etc.* op. cit.

Les situations de NEET dans les parcours d'insertion des jeunes en France, Christine Guégnard, Jean-François Giret, Olivier Joseph, Jake Murdoch, in *Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc.*, op. cit.

Alice au Pays des NEET : la traversée du miroir sur 20 ans, Magali Danner, Christine Guégnard, Olivier Joseph, in *Jeunesse(s) et transitions vers l'âge adulte etc.* op. cit.

Quelques trimestres de mauvaise conjoncture au départ ne suffisent pas à dégrader durablement les conditions d'insertion des jeunes. Stéphane Jugnot, Claude, Minni. In 20 ans d'insertion professionnelle des jeunes etc. op. cit.

Les transitions professionnelles tout au long de la vie. Nouveaux regards, nouveaux sens, nouvelles temporalités ?, Baslé Maurice, Beaupère Nathalie, Guéguen Chantal, Issehnane Sabina, Moizeau Fabien, Moysan-Louazel Anne, Podevin Gérard, Urasadettan Jennifer, Boudesseul Gérard, Couppié Thomas, Giret Jean-François, Lemistre Philippe, Werquin Patrick, Céreq Echanges, n°1, 2016, 526p.

▪ **Sur les contextes territoriaux et les mobilités géographiques :**

Déménager pour accéder aux études supérieures et débiter sa vie active : des parcours façonnés par les ressources territoriales et sociales, Mélanie Vignale et Arnaud Dupray, in CGET : Les mobilités résidentielles en France : Tendances et impacts territoriaux - Rapport 2018 de l'Observatoire des territoires, 2018, p. 88-95

La force du local, la part du rural. Les parcours d'insertion des jeunes sans diplôme, Joël Zaffran, in Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc., op. cit.

Emploi des jeunes sans diplôme : la prime au rural, Joël Zaffran, Formation Emploi 2018/2 (n° 142), p. 99-117

20 ans d'évolutions régionales de l'insertion. Patrice Caro, Gérard Boudesseul, Agnès Checcaglini., In 20 ans d'insertion professionnelle etc. op. cit.

Quel rôle pour le marché du travail local sur le risque de décrochage scolaire ?, Vanessa di Paola et Stéphanie Moullet, Formation Emploi, 2018/4 (n° 144), p. 95-116

Mobilités interrégionales de jeunes diplômés - Analyse par niveau de sortie, Mélanie Vignale, Net.Doc, n° 160, Juin 2016, 60 p.

Mobilités interrégionales de jeunes diplômés du supérieur : qui forme pour qui ?, Mélanie Vignale, Céreq Bref n°347, 2016.

Partir travailler à l'étranger après les études, un atout pour l'insertion professionnelle des jeunes ?, Julien Calmand, in Jeunesse(s) et transitions vers l'âge adulte etc. op. cit.

Les liens entre expériences à l'étranger en cours d'études et insertion. Une approche par le concept de « capital de mobilité », Julien Calmand, Stéphanie Condon, Karine Pietropaoli, Pascale Rouaud, Emmanuelle Santelli, in Rendement éducatif, parcours et inégalités dans l'insertion etc., op. cit.

Expériences à l'étranger en cours d'études et insertion : des liens complexes, pour quelle plus-value ? J. Calmand, S. Condon, K. Pietropaoli, P. Rouaud, E. Santelli, Formation Emploi, 142, 2018.

Séjours à l'étranger en cours d'études, une plus-value sur l'insertion en France ? Julien Calmand, Pascale Rouaud, Emmanuel Sulzer, Céreq Bref, n°348, 2016, 4p.

Portrait des jeunes en mobilité internationale, Analyses de l'enquête Génération 2010 (1/2), Les notes de l'observatoire Erasmus+, n°1, L'Observatoire Erasmus +, 2015, 18 p.

(sur l'enquête 2016 auprès de la Génération 2013) Séjours des jeunes à l'étranger : des objectifs européens partiellement atteints, mais un accès encore inégal à la mobilité Julien Calmand, Alexie Robert, Céreq Bref, n° 371, Janvier 2019, 4 p.

▪ **Sur les conditions d'emploi et de travail :**

La place des emplois atypiques dans les trajectoires d'entrée dans la vie active, Hélène Couprie, Xavier Joutard, Revue française d'économie 2017/1 (Vol. XXXII), p. 59-93.

Des parcours différenciés par le motif de recours au CDD, Catherine Béduwé, Brigitte Reynès, in Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc., op. cit.

Le déclassement : un phénomène enraciné. Vanessa di Paola, Stéphanie Moullet. In 20 ans d'insertion professionnelle des jeunes etc. op. cit.

Une progression contrastée des salaires en 20 ans, source de réduction des inégalités. Christophe Barret, Arnaud Dupray. In 20 ans d'insertion professionnelle des jeunes etc. op. cit.

Comment les conditions d'insertion des jeunes se sont-elles transformées en 20 ans ? Virginie, Mora. In 20 ans d'insertion professionnelle des jeunes etc. op. cit.

La mobilité professionnelle des cadres débutant dans une petite entreprise du secteur privé, Marina Bourgain, in Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc., op. cit.

En vingt ans, des parcours d'insertion dégradés mais des jeunes moins inquiets pour leur avenir professionnel, Virginie Mora, in Insee Références, Formations et emploi, Édition 2018

D'une Génération à l'autre : l'inquiétude des jeunes en question, Catherine Béduwé, Arnaud Dupray, Céreq Bref, n°361, 2018, 4p.

Perceived Job Insecurity in early careers, Catherine Béduwé, Arnaud Dupray, Assâad El Akremi, Working paper, n°6, 2018, 20p.

Que gagne-t-on à se former ? Zoom sur 20 ans d'évolution des salaires en début de vie active, Christophe Barret, Arnaud Dupray, Céreq Bref n°372, 2019, 4 p.

Les jeunes diplômés de bac+5 s'estiment-ils compétents pour occuper leurs emplois ?, Julien Calmand, Jean-François Giret, Philippe Lemistre, Boris Ménard, Céreq Bref, n°340, 2015, 4p.

Évaluation des compétences par des diplômés de Bac + 5 et emploi : quelles incidences sur les politiques de professionnalisation ?, Philippe Lemistre, Boris Ménard in Alternance et professionnalisation etc. op. cit.

Compétences acquises et requises des diplômés de bac+5, Julien Calmand, Jean-François Giret, Philippe Lemistre, Boris Ménard, Net.Doc, n°142, 2015, 62p.

L'effet des soft-skills sur la rémunération des diplômés, Inès Albandea, Jean-François Giret, Net.Doc, n°149, 2016, 31p.

▪ **Sur les inégalités sociales et leurs liens avec l'insertion professionnelle et les parcours :**

Et les femmes devinrent plus diplômées que les hommes..., Thomas Couppié, Dominique Epiphane, Céreq Bref n°373, 2019, 4 p.

La relation genre-insertion a-t-elle évolué en 20 ans ? Thomas Couppié et Dominique Epiphane. In 20 ans d'insertion professionnelle etc. op. cit.

De l'école aux premiers emplois : analyse des ségrégations sexuées, Alban Jacquemart, Danièle Trancart, in Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc., op. cit.

Accès des jeunes femmes et des jeunes hommes aux emplois cadres, une égalité trompeuse, Vanessa di Paola, Arnaud Dupray, Dominique Epiphane, Stéphanie Moullet, Céreq Bref, n°359, 2017, 4p.

Accès des femmes et des hommes aux positions de cadres en début de vie active : une convergence en marche ?, Vanessa di Paola, Arnaud Dupray, Dominique Epiphane, Stéphanie Moullet, in Insee Femmes et hommes, l'égalité en question, Insee Références, Édition 2017

Les jeunes hommes dans des formations très féminisées : quelles destinées professionnelles ?, Thomas Couppié, Dominique Epiphane in Alternance et professionnalisation etc. op. cit.

Les débuts de carrières des femmes diplômées des filières scientifiques - « Rien ne se perd, rien ne se crée, tout se transforme » ?, Thomas Couppié, Dominique Epiphane Net.Doc, n°155, 2016, 25p.

Les ZUS, 20 ans après. Valentine Henrard, Mélanie Vignale, In 20 ans d'insertion professionnelle etc. op. cit.

Quelles évolutions de l'insertion professionnelle des jeunes issus de l'immigration ? Jean-Luc Primon. In 20 ans d'insertion professionnelle etc. op. cit.

Les trajectoires d'insertion professionnelle des jeunes portent-elles la marque de leurs origines socio-économique et culturelle ? Céline Goffette, Louis-André Vallet. In 20 ans d'insertion professionnelle etc. op. cit.

D'une génération à l'autre : enseignement professionnel et inégalités sociales, Vanessa di Paola, Stéphanie Moullet, Noémie Olympio, Eric Verdier, in Rendement éducatif, parcours et inégalités dans l'insertion des jeunes etc., op. cit.

De la génération 1992 à la génération 2010, comment le lien entre la qualité de l'insertion professionnelle des jeunes et leur origine socio-économique et culturelle a-t-il évolué ?, Céline Goffette, Louis-André Vallet in Jeunesse(s) et transitions vers l'âge adulte etc. op. cit.

▪ **Travaux méthodologiques**

Bilan méthodologique - Enquête Génération 2010 Interrogation 2013, Christophe Barret, Christophe Dzikowski, Zora Mazari, Pascale Rouaud, Florence Ryk, Noémie Volle, Céreq Etudes, n°25, 2019, 151p.

Bilan méthodologique - Enquête Génération 2010 Interrogation 2015, Manon Olaria (coord.) Christophe Barret, Mady Cissé, Christophe Dzikowski, Émilie Gaubert, Zora Mazari, Manon Olaria, Florence Ryk, Céreq Etudes, n°26, 2019 (à paraître).

Bilan méthodologique - Enquête Génération 2010 Interrogation 2017, Manon Olaria (coord.), Mady Cissé, Émilie Gaubert, Zora Mazari, Manon Olaria, Céreq Etudes, n°27, 2019 (à paraître)

Plan de sondage des enquêtes Génération : utilisation d'un calage pour suréchantillonner les extensions, Christophe Barret, Mady Cissé, Christophe Dzikowski, XIIIèmes Journées de Méthodologie Statistique, 2018.

Efficacité d'un protocole multimode (téléphone et internet), Christophe Barret, Mady Cissé, Emilie Gaubert, Zora Mazari, Manon Olaria, XIIIèmes Journées de Méthodologie Statistique, 2018.

Agrégation de données multimode : Impact sur la modélisation des variables présentant un effet de mesure, Christophe Barret, Mady Cissé, XIIIèmes Journées de Méthodologie Statistique, 2018.

Note sur l'attrition et la déformation de l'échantillon des enquêtes - Enquête Génération 2010, Mady Cissé, Céreq Etudes, n°28, 2019, 14p

A.5 – CCTP du marché public « Enquête multimode » (téléphone, internet)



Marché public de fournitures courantes ou de services Marché d'appel d'offre européen

Première interrogation auprès d'un échantillon de sortants
du système éducatif français en 2017.

CAHIER DES CLAUSES TECHNIQUES PARTICULIÈRES

**Enquête multimode (téléphone, internet) 2020
auprès de la Génération 2017**

Sommaire

1.	Objectif de l'opération.....	120
2.	Champ de l'enquête « Génération »	122
3.	L'échantillon et le fichier d'import.....	123
4.	Minutage et effectif prévisionnel de répondants.....	125
5.	La phase « préparatoire »	126
5.1	<i>L'envoi de lettres avis par mail</i>	<i>126</i>
5.2	<i>Ouverture du numéro gratuit et réception d'appels.....</i>	<i>127</i>
5.3	<i>Ouverture d'une boîte email de contact et traitement des emails</i>	<i>128</i>
5.4	<i>Émission d'appels</i>	<i>128</i>
6.	Les fichiers « historiques de suivi »	130
6.1	<i>Historique de contact et de tentatives de contact.....</i>	<i>130</i>
6.1.1	<i>Un fichier historique d'appels téléphoniques entrants sur le numéro gratuit</i>	<i>130</i>
6.1.2	<i>Un fichier historique d'appels téléphoniques sortants</i>	<i>131</i>
6.1.3	<i>Un fichier historique de la boîte mail Céreq</i>	<i>131</i>
6.1.4	<i>Un fichier historique des mails de contact.....</i>	<i>132</i>
6.1.5	<i>Un fichier historique des SMS de contact</i>	<i>132</i>
6.1.6	<i>Un fichier historique de connexions sur la plateforme CAWI</i>	<i>132</i>
6.1.7	<i>Un fichier historique récapitulatif</i>	<i>133</i>
6.2	<i>Historique des réponses à l'enquête.....</i>	<i>133</i>
7.	La phase « Enquête multimode ».....	134
7.1	<i>Informations générales sur l'enquête</i>	<i>134</i>
7.2	<i>Le protocole de collecte multimode internet-téléphone.....</i>	<i>135</i>
7.2.1	<i>Phase 1 - Incitation à répondre par internet.....</i>	<i>135</i>
7.2.2	<i>Phase 2 - Choix de répondre par internet ou téléphone.....</i>	<i>136</i>
7.2.3	<i>Phase 3 – Rattrapage pour répondre par téléphone</i>	<i>136</i>
7.3	<i>Le protocole de contact</i>	<i>136</i>
7.3.1	<i>Durée totale du terrain</i>	<i>136</i>
7.3.2	<i>Consignes pour le premier contact</i>	<i>136</i>
7.3.3	<i>Numéros de téléphone et tentatives</i>	<i>138</i>
7.4	<i>Le protocole de relances multimode.....</i>	<i>139</i>
7.4.1	<i>Les contacts téléphoniques et message sur le répondeur</i>	<i>139</i>
7.4.2	<i>Les relances mails.....</i>	<i>139</i>
7.4.3	<i>Les relances par SMS.....</i>	<i>140</i>
8.	Spécificités de développement informatique du questionnaire	141
8.1	<i>Contenu du questionnaire de l'enquête multimode</i>	<i>141</i>
8.2	<i>Fourniture du questionnaire multimode par le Céreq.....</i>	<i>142</i>
8.3	<i>Personnalisation du questionnaire et charte graphique.....</i>	<i>142</i>
8.4	<i>Développement du questionnaire et contraintes techniques imposées</i>	<i>144</i>
8.5	<i>Le calendrier interactif.....</i>	<i>147</i>
8.5.1	<i>Le calendrier d'activité</i>	<i>147</i>
8.5.2	<i>Le calendrier du mode de cohabitation</i>	<i>150</i>
8.6	<i>Le tableau dynamique</i>	<i>151</i>
8.6.1	<i>Le tableau des mobilités résidentielles</i>	<i>151</i>
8.6.2	<i>Le tableau des langues parlées</i>	<i>152</i>
9.	Tests du questionnaire de l'enquête.....	153
9.1	<i>Tests manuels et tests automatiques du questionnaire</i>	<i>153</i>

9.2	Tests réels en plateau d'enquête	153
9.3	Test « Youthproofing » du questionnaire	155
10.	Composition des équipes, présences plateaux et formations	157
10.1	Responsable du projet et organisation	157
10.2	Équipes et plateau	157
10.3	Formation des superviseurs et des enquêteurs.....	158
10.4	Suivi du plateau, responsable et personne « qualité »	158
11.	Transmission, suivi et bilans de la phase enquête téléphonique.....	160
12.	Préconisations pour l'application du Règlement général de protection des données (RGPD)	162
13.	Calendrier prévisionnel des opérations	165
14.	Annexes.....	166
	Annexe 1 : Extrait du questionnaire.....	166
	Annexe 2 : Éléments de charte graphique	170
	Annexe 3 : Quelques données de cadrage sur les appels téléphoniques	171
15.	Annexe financière.....	172
	Notice de l'annexe financière.....	175

Pour information :

Tableau 1 – Dessin de fichier d'import (provisoire)	123
Tableau 2 – Descriptif des critères de charte graphique demandés	143
Tableau 3 – Descriptif des techniques de développement demandées	144
Tableau 4 – Description de la variable de synthèse CAL du calendrier d'activité	149
Figure 1 – Logigramme du protocole de l'enquête multimode	135
Figure 2 – Phase de contact.....	138
Figure 3 – Logigramme (provisoire) du questionnaire de l'enquête	141
Figure 4 – Visuel du calendrier d'activité de l'enquête 2020 (exemple)	147
Figure 5 – Visuel du calendrier du mode de cohabitation de l'enquête 2020	150
Figure 6 – Visuel du tableau des mobilités résidentielles.....	151
Figure 7 – Visuel du tableau des langues parlées	152

1. Objectif de l'opération

Le Centre d'études et de recherches sur les qualifications (Céreq), organisme public national opérant sous la double tutelle ministre chargé de l'éducation et du ministre chargé de l'emploi, est producteur des enquêtes Génération, enquêtes statistiques nationales qui portent sur l'insertion et le cheminement des sortants du système éducatif français lors de leurs premières années de vie active.

Ces enquêtes ont pour objectif principal de produire des indicateurs statistiques sur l'insertion professionnelle (taux d'emploi, taux de chômage, taux d'emploi à durée indéterminée, etc.), selon les niveaux de formation, les filières et spécialités de formation, à destination des acteurs publics. Elles produisent ainsi des informations qui contribuent à la compréhension des processus d'insertion et des différenciations des parcours en début de carrière.

Cette enquête s'inscrit dans le nouveau dispositif « Génération » (première enquête du dispositif rénové). Les enquêtes s'enchaînent au rythme d'une génération nouvelle de sortants interrogée tous les quatre ans. Chaque génération de sortants est interrogée trois et six ans après la sortie du système éducatif. La première interrogation, à trois ans, est principalement tournée vers la production d'indicateurs et des usages institutionnels et la ré-interrogation, à six ans, est plutôt centrée sur les analyses des parcours et des mobilités sur moyen terme.

L'objet de l'étude menée par le Céreq est de retracer les itinéraires d'insertion professionnelle d'individus issus de tous les niveaux de formation initiale (diplômés ou non), depuis leur sortie du système éducatif l'année scolaire 2016-2017.

Cette nouvelle enquête, organisée au printemps 2020 (d'avril à juillet), permettra de reconstituer la chronologie des situations professionnelles connues par ces individus entre leur sortie du système éducatif, en 2016-2017, jusqu'au printemps 2020, soit trois années sur le marché du travail.

L'enquête couvre les sortants des établissements de l'ensemble du territoire national, France métropolitaine et départements et régions d'outre-mer - DROM (Guadeloupe, Martinique, Guyane, île de la Réunion et Mayotte).

Les objectifs du présent marché sont :

1. **Envoyer une lettre avis par mail** aux individus disposant d'une adresse mail dans l'échantillon.
2. **Réceptionner des appels sur le numéro gratuit et répondre aux mails envoyés par les enquêtés.** Une hotline devra permettre aux individus, ayant reçu une lettre avis papier ou mail, de se renseigner sur l'enquête, de communiquer leurs coordonnées de contact (téléphone et/ou adresse mail) ou encore de répondre directement à l'enquête. Le prestataire devra également traiter les mails des individus soit en répondant directement par ce biais, soit, si nécessaire, en appelant l'individu.
3. **Construire des historiques d'appels et de connexions complets** afin de suivre très précisément le remplissage et la navigation de l'enquêté quel que soit le mode de collecte. Il s'agit également d'être en mesure de déterminer, pour chacune des réponses collectées, le mode de réponse utilisé. Ces informations sont indispensables pour déterminer la qualité des données produites et permettre leur analyse *a posteriori*. Ce fichier comprendra également l'historique des contacts de toute nature (envois des lettre avis, des relances téléphoniques, mails, sms, ...).
4. **Développer un questionnaire multimode téléphone-internet CATI-CAWI.** L'outil de collecte devra être ergonomique - notamment pour limiter les biais entre modes de collecte - et responsive pour une utilisation sur tablette et smartphone. Le système devra permettre le basculement instantané d'un mode de collecte à un autre. L'enquêté aura la possibilité de commencer son questionnaire dans un mode et le finir sur le mode alternatif (en reprenant à l'endroit où il s'est arrêté).

5. **Administrer par internet et/ou par téléphone un questionnaire** de 30 minutes en moyenne auprès d'un échantillon compris entre 200 000 et 300 000 individus (nombre d'individus envoyés en production). *La durée moyenne cible du questionnaire sera précisément définie une fois le questionnaire stabilisé (après les phases de tests).*
6. **Relancer par téléphone, par mail et par sms les individus** n'ayant pas encore répondu selon le protocole établi par le Céreq. Tous les individus ne seront pas forcément soumis au même protocole de relance. L'incitation à répondre par internet au questionnaire évoluera en cours de d'enquête : très forte en début de plateau puis de plus en plus modérée. Le protocole définit au maximum 10 relances par mail disponible et 20 tentatives d'appel sur chaque numéro non classé définitivement. Le prestataire sera aussi amené à laisser des messages sur le répondeur de manière périodique. Le protocole de relance est décrit dans la partie 7.4.

Le questionnaire de l'enquête Génération s'articule autour d'un calendrier d'activité. Celui-ci permet de suivre mois par mois la situation des jeunes à l'issue de leur formation initiale et de décrire les séquences successives d'emploi et de recherche d'emploi, de formation/reprise d'études ou d'inactivité. Ce module constitue le cœur de l'enquête et commande l'ouverture de pop-up et/ou de modules descriptifs des situations déclarées. Une attention particulière devra donc lui être portée de façon à assurer la cohérence du questionnaire.

Les individus échantillonnés sont issus de tous les niveaux de formation initiale, des jeunes sans qualification aux plus diplômés. L'expérience des collectes précédentes montre que la facilité à joindre les jeunes est variable selon le niveau de formation et la situation d'activité au moment de l'enquête. Il conviendra donc d'assurer un effort approprié au cours de la collecte pour respecter au mieux le protocole de contact et de relance. Cela permettra ainsi d'éviter que l'échantillon des répondants ne soit biaisé.

Les enquêtes Génération, extensions comprises, sont des enquêtes de la statistique publique et relèvent à ce titre de la loi n° 51-711 du 7 juin 1951 sur l'obligation, la coordination et le secret en matière de statistiques. Celle-ci définit le cadre de procédures destinées à garantir l'intérêt, la pertinence et la qualité des enquêtes publiques, ainsi que la confidentialité des informations collectées.

L'opportunité de l'enquête sera ainsi discutée au Comité national de l'information statistique (Cnis) en avril 2019. Le Comité du label de la statistique publique du Cnis l'examinera ensuite, au cours du mois de septembre 2019, afin d'obtenir le label d'intérêt général et de qualité statistique.

Pour plus d'information sur les enquêtes « Génération », à titre d'exemple, les résultats détaillés de la dernière enquête réalisée trois ans après la sortie du système éducatif :

<http://www.cereq.fr/publications/Cereq-Enquetes/Quand-l-ecole-est-finie.-Premiers-pas-dans-la-vie-active-de-la-Generation-2013>

2. Champ de l'enquête « Génération »

Les individus à interroger sont issus d'une base de sondage, constituée par le Céreq, couvrant l'ensemble des formations initiales. Tous les niveaux d'éducation sont représentés par les individus constituant cette base de sondage : des niveaux secondaires – 4^{ème} au baccalauréat – jusqu'aux formations de l'enseignement supérieur. Elle comprend tous les types d'établissements délivrant de la formation initiale en France ou dans un DROM.

La Génération 2017 concerne les « primo sortants » de formation initiale en 2016-2017 (année scolaire). Dans les « primo-sortants » sont inclus les individus ayant interrompu leurs études par le passé pendant une durée comprise entre 1 et 16 mois.

Les individus interrogés sont français ou étrangers au moment de l'interrogation, quel que soit leur niveau de diplôme. Les individus résidant à l'étranger à la date d'interrogation (hors France métropolitaine et DROM) sont également interrogés.

De façon plus précise, les critères d'éligibilité pour être dans le champ retenu, nommé ensuite « champ Céreq », sont les suivants :

- Avoir été inscrit dans un établissement de formation en France métropolitaine ou dans un DROM durant l'année scolaire 2016-2017 ;
- Avoir quitté le système éducatif entre octobre 2016 et octobre 2017, diplômé ou non ;
- Ne pas avoir interrompu ses études durant 17 mois ou plus avant l'année scolaire 2016-2017 (sauf pour raison de santé) ;
- Ne pas avoir repris ses études pendant les 16 mois qui ont suivi l'entrée sur le marché du travail ;
- Avoir 35 ans ou moins en 2017.

Ces conditions sont cumulatives.

Pour résumer, le champ de l'enquête est constitué des jeunes scolarisés en France ou dans un DROM l'année scolaire 2016-2017, n'ayant pas poursuivi d'études en 2017-2018 et n'ayant pas interrompu leurs études plus de 17 mois avant 2017.

3. L'échantillon et le fichier d'import

Le Céreq fournira au prestataire un fichier suffisamment dimensionné pour réaliser entre 20 000 et 45 000 enquêtes. Cette base comprendra entre 200 000 à 300 000 individus à contacter. Ce fichier est construit à partir des informations issues de la base de sondage complétées des coordonnées postales et téléphoniques mises à jour suite aux recherches dans les fichiers d'adresses de La Poste.

Cet échantillon national vise à être représentatif, sur le territoire français, de l'ensemble des sortants du système éducatif l'année scolaire 2016-2017. Ces individus seront soumis au protocole de collecte multimode décrit dans la partie 7.2 de ce marché avec pour objectif principal de maximiser le taux de réponse.

Les informations individuelles présentes dans le fichier d'import seront de différents ordres.

Tableau 1 – Dessin de fichier d'import (provisoire)

Variable	Descriptif	Type	Long	Modalités
IDENT	Matricule individu	char	8	1 lettre + 7 chiffres
NOM	Nom individu	char	29	
PREN	Prénom individu	char	31	
ADR1	Adresse	char	37	
ADR2	Adresse complémentaire	char	36	
ADR3	Adresse complémentaire	char	36	
CP	Code postal	num	5	
VILLE	Ville	char	36	
DROM	Numéro de téléphone issu d'un DROM	num	1	1 = Réunion 2 = Guadeloupe 3 = Martinique 4 = Guyane 6 = Mayotte
SEXE	Sexe de l'individu	char	1	1 = Homme 2 = Femme
TEL	Téléphone principal phase enquête	num	10	
TELB	Téléphone secondaire phase enquête	num	10	
TELC	Téléphone 3 phase enquête	num	10	
TELD	Téléphone 4 phase enquête	num	10	
TEL01	Téléphone 1 d'enrichissement	num	10	
TEL02	Téléphone 2 d'enrichissement	num	10	
TEL03	Téléphone 3 d'enrichissement	num	10	
TEL04	Téléphone 4 d'enrichissement	num	10	
TEL05	Téléphone 5 d'enrichissement	num	10	
TEL06	Téléphone 6 d'enrichissement	num	10	
DEST	Destinataire	char	3	
STRATE	Strates détaillées	num	3	
STRADEF	Strates agrégées	num	2	
ZUS	Appartenance à une zone urbaine sensible	char	1	1 (oui), 2 (non)
EFP	Type d'école de la fonction publique	char	1	1 (oui), 2 (non) ou « »
NETAB	Nom de l'établissement de formation	char	40	
ANAI	Année de naissance	char	4	
MNAI	Mois de naissance	char	2	
CLASS	Classe de sortie	char	80	
DIPLO	Nom du diplôme en cas de classe terminale	char	80	
IMPSP1 à 4	Libellé de spécialité de la classe de sortie	char	320	
AGRI	Lycée agricole	char	1	1 (oui), 2 (non)
CFA	Formation dispensée par voie d'apprentissage	char	1	1 (oui), 2 (non)

CAPBE	Formation de CAP/BEP	char	1	1 = sortant (année terminale ou non) 2 = non sortant
ANTER	Classe terminale	char	1	1 (oui), 2 (non)
SUPER	Sortants du supérieur (dont IV+)	char	1	1 (oui), 2 (non)
INFG3	Classe de sortie inférieure ou égale à la 3ème	char	1	1 (oui), 2 (non)
IMP39	Classe de sortie agrégée	char	1	1 = 1ere et 2eme année DEUG 2 = 1ere et 2eme année DUT 3 = 1ere et 2eme année BTS 4 = Autre
BP3	Sortant de bac+3 ou plus	char	1	1 (oui), 2 (non)
ERA	Niveau de diplôme	char	1	1 = inférieur bac+2 2 = école 3 = université, supérieur ou égal à bac+2
MODTH	Niveau thèse (hors santé)	char	1	1 (oui), 2 (non)
TERRA	Identification par niveau de diplôme	char	1	1 = Bac+2 et plus 2 = Niveau bac et CAP/BEP diplômés 3 = CAP/BEP non diplômés et sans qualification
LA	Type de lettre avis	char	1	1 = Lettre papier 2 = Lettre mail 3 = Les deux
REGETAB	Région de l'établissement de formation	char	2	
MAIL	Adresse mail	char	50	Adresse mail initiale
TYPE	Type d'établissement allant de « A » à « Z »	char	1	
MDP	Mot de passe	char	6	

Ce dessin de fichier est donné à titre indicatif et n'est donc pas définitif. Il est possible que des variables soient modifiées, supprimées ou ajoutées.

Important : il est indispensable que le prestataire respecte l'intégrité du fichier transmis par le Céreq (le nom des variables, le format (texte ou numérique) et la liste et numérotation des modalités). À titre d'exemple, les « 0 » en début de chaînes de caractères devront être conservés (code postal ou numéro de téléphone).

4. Minutage et effectif prévisionnel de répondants

Le nombre de répondants visé sera d'environ 30 000 (au minimum 20 000 et au maximum 45 000 individus). La taille de l'échantillon fourni sera de l'ordre de 200 000 à 300 000 individus.

La durée moyenne cumulée (DMC) est la durée moyenne de l'ensemble des appels téléphoniques effectués pour obtenir des questionnaires complets et validés. La DMC cible est difficilement estimable puisqu'au moment de la rédaction du présent marché, les arbitrages sur le questionnaire n'ont pas tous été effectués. Trois notions sont donc introduites dans le cadre de ce marché :

- La DMC cible, fixée à 30 minutes ;
- La DMC minimum, fixée à 25 minutes ;
- La DMC maximum, fixée à 35 minutes.

Ces données servent de base pour construire l'annexe financière du présent marché.

L'annexe financière jointe devra être complétée par le contractant sur la base des informations pré-remplies et précisera le coût de la minute supplémentaire d'enquête ainsi que la déduction du coût pour une minute en moins.

Le minutage moyen pour une enquête réalisée ne comprend ni les tentatives d'appels, ni les contacts intermédiaires avec des tiers. Autrement dit, le minutage comprend uniquement le temps passé en enquête avec l'individu recherché. La réalisation d'une enquête peut s'opérer en plusieurs fois (comportant un ou plusieurs rappels pour finir l'enquête, permettant de remplir une partie du questionnaire), le cumul des durées ayant amené à une réponse complète pourra être utilisé. De même, le minutage moyen sera un minutage moyen global et ne distinguera pas les éventuelles différences entre temps moyen sur téléphone mobile et temps moyen sur téléphone fixe.

Si la DMC prévisionnelle, estimée à partir des tests, laisse penser que la DMC finale se situera en dehors de la fourchette [25 min, 35 min], une réunion avec un compte rendu écrit entre le prestataire et le Céreq aura lieu avant le démarrage du plateau. Cette réunion permettra de statuer, d'un commun accord, sur la faisabilité financière et technique de l'opération et, le cas échéant, sur les ajustements possibles (réduction de la taille du questionnaire, augmentation de la taille du plateau, etc.) à effectuer pour aboutir à la réalisation de l'enquête dans des conditions acceptables pour les deux parties et dans la limite des règles fixées par les marchés publics.

Le moindre ralentissement du système informatique (ralentissement réseau, temps de latence entre les questions...) ne pourra pas être imputé sur le temps moyen de passation du questionnaire. Dans ce cas, la durée moyenne de ralentissement sera estimée et déduite de la durée moyenne du questionnaire.

Le prestataire devra adapter sa production en continu (notamment le nombre d'enquêteurs alloué à l'opération) pour permettre de respecter le protocole fixé et *in fine*, d'atteindre le nombre de répondants souhaité dans les délais impartis.

5. La phase « préparatoire »

La phase dite « préparatoire » inclut l'ensemble des opérations réalisées avant le lancement de l'enquête à proprement parler.

5.1 L'envoi de lettres avis par mail

(Les envois par courrier ne font pas partie de ce présent marché, seuls les envois par mails sont concernés)

La transmission de la lettre avis est une obligation pour obtenir le label de qualité statistique. L'ensemble des individus de l'échantillon sera contacté par voie postale et/ou numérique avant le démarrage de la collecte. L'objectif principal est d'informer les individus du lancement de l'enquête. Ils sont sollicités pour participer à l'enquête sur l'insertion professionnelle par téléphone ou par internet. Au-delà du rôle informatif sur les objectifs, le contenu et la période d'enquête, l'individu a la possibilité de modifier ses coordonnées via le numéro gratuit fourni. La prise de rendez-vous pour répondre à l'enquête est également possible par ce biais. Il a surtout la possibilité de se connecter directement au questionnaire pour répondre à l'enquête via un lien de connexion.

Le cœur du questionnaire est un calendrier d'activité qui décrit mois par mois la situation professionnelle de l'individu sur les trois dernières années. Pour faciliter son remplissage lors de l'entretien téléphonique ou en auto-administré, un modèle de calendrier accompagne la lettre-avis. Chacun a donc la possibilité de le préparer en amont de l'enquête.

Pour tous les individus disposant d'une adresse mail, une lettre avis électronique devra être transmise (objet de ce présent marché). Ce courriel sera envoyé au démarrage du plateau (à titre d'information, le taux de présence du mail dans l'échantillon est estimé *a minima* à 30 %).

Le prestataire s'engage, à l'aide d'une solution technique adaptée, à ce que chaque adresse électronique fournie dans le fichier d'import ait fait l'objet d'un envoi de lettre avis mail. Il informera le Céreq du nombre de retours dû à des mails erronés et fournira la liste des individus concernés. Il doit également s'engager à garantir la sécurité lors de la connexion au questionnaire, notamment au travers d'un site sécurisé (https).

Durant toute la phase de collecte multimode, des envois de mails en masse seront réalisés que ce soit pour l'envoi du mail avis ou pour les relances par mail (partie 7.4.2). Le prestataire devra montrer sa capacité à envoyer des mails en masse en précisant le logiciel ou la solution retenue ainsi que les réglages effectués : nombre de mails par lot, espacement entre les lots,... Ces mails auront des objets et nom d'expéditeur permettant de reconnaître facilement l'opération Génération du Céreq.

Le candidat devra démontrer sa capacité technique à gérer l'envoi massif de courriers électroniques :

- Décrire et détailler l'outil utilisé ;
- Gestion des blacklists des hébergeurs de mail ;
- Rythme d'envoi pertinent ;
- Gestion des envois de mails en masse automatiques et programmés à une date donnée ;
- Indication sur le nombre de champs maximum à personnaliser dans le mail ;
- Utilisation d'une adresse mail Céreq comme boîte d'expédition, utilisation d'un alias au choix du Céreq ;
- Capacité à utiliser dans le mail des liens web externes (suffixés cereq.fr). Certains de ces liens seront simplement des redirections vers les adresses du prestataire ;
- Conversion en format web du modèle de lettre avis et adaptation aux différents navigateurs ou téléphones mobiles.

5.2 Ouverture du numéro gratuit et réception d'appels

Un numéro gratuit sera mis en place par le prestataire. Il devra être opérationnel dès l'envoi des courriers aux individus à enquêter (mail avis ou lettre avis), c'est-à-dire le jour du démarrage du plateau d'enquête. Ce numéro gratuit devra être communiqué suffisamment tôt au Céreq pour que celui-ci puisse l'inclure dans l'impression des lettres/mails avis (*l'envoi des lettres-avis courrier ne fait pas partie du présent marché*).

Les appels sur ce numéro gratuit permettront de contacter le prestataire toute la journée, du lundi au samedi, pendant tout le déroulement de l'enquête. En dehors des heures de présence des enquêteurs, ou en cas de surcharge de la ligne, un répondeur devra être activé et les messages traités dès que possible.

Les individus pourront ainsi obtenir des informations complémentaires, actualiser leurs coordonnées, mentionner les plages horaires souhaitées pour un contact ultérieur, manifester leur refus d'être interrogé et le cas échéant répondre à l'enquête.

Le Céreq préconise pour cette phase la détection du numéro entrant pour accéder à la fiche de l'individu. Dans le cas où le numéro n'est pas reconnu, il faudra prévoir une phase manuelle à partir d'une interface en lien direct avec la base d'enquête (avec recherche à partir du numéro de téléphone, identifiant Céreq, nom, prénom, date de naissance, code postal et commune).

Le candidat devra démontrer sa capacité de réception et de traitement des appels sur ce numéro gratuit :

- Commande du numéro gratuit et délai d'activation ;
- Faculté à traiter les appels entrants et à identifier des individus (méthode d'identification) ;
- Création d'une interface de recherche de l'appelant ;
- Gestion des enrichissements et lien direct avec la base de données de l'enquête ;
- Répondeur, capacité, délais et traitement des messages reçus.

5.3 Ouverture d'une boîte email de contact et traitement des emails

Suite à l'envoi des lettres avis, l'enquêté aura la possibilité de poser des questions et obtenir de l'information complémentaire sur l'enquête en envoyant un mail sur une boîte Céreq ouverte spécialement pour l'opération (création et hébergement par le Céreq).

Cette boîte mail sera gérée par le prestataire qui devra réceptionner et traiter les mails reçus quotidiennement. Il devra également identifier la raison de l'envoi, par exemple un problème de connexion, un problème de compréhension du questionnaire ou un problème technique et être en mesure d'apporter une solution. Pour chaque mail reçu, le prestataire adressera une réponse par mail ou, si nécessaire, par un appel téléphonique. Si la demande nécessite d'intervenir sur la fiche de l'individu (par exemple un refus de répondre évoquant les règles Cnil), le prestataire mettra à jour les informations directement dans la base d'enquête et en informera le Céreq.

Des échanges avec le Céreq (ou la personne qualité « Céreq » présente dans les locaux du prestataire durant toute la période de collecte) seront établis pour répondre aux questions les plus complexes ou en cas d'hésitation sur le traitement à effectuer.

Il sera demandé au prestataire de mettre en place une procédure avec des délais courts de réponse pour chaque nouveau message reçu.

5.4 Émission d'appels

Le candidat détaillera les caractéristiques techniques de sa plateforme d'appels téléphoniques. Il devra indiquer quel mode d'appel sera utilisé entre Preview, Predictif, Progressif, ou tout autre mode d'appel. Ce choix devra être justifié et détaillé selon la méthode choisie. Le Céreq prendra la décision du mode d'appel conjointement avec le contractant.

Il précisera en particulier si l'outil permet :

- Une alternance entre appel masqué/démasqué en détaillant le protocole (changement tous les n appels ou changement en fonction de la date d'appel). Si un individu, d'office, n'accepte pas les numéros masqués, des appels avec un numéro démasqué devront être effectués sur la totalité des tentatives restantes (afin que les 20 tentatives soient effectives). Le contractant devra donc être en mesure à la fois d'attribuer un mode d'appel inchangeable à un type d'individu, tout en proposant une alternance de mode d'appel au reste de la population ;
- D'afficher un ou plusieurs numéro(s) d'appelant avec un indicatif classique (de type 01, ..., 05) ;
- D'afficher un numéro d'appelant « régionalisé » dans le cas d'un appel sur un numéro fixe (indicatif adapté au lieu de résidence de l'enquêté) ou a minima avec un préfixe non spécial ;
- De rediriger sur le numéro gratuit dans le cas d'un rappel sur le numéro démasqué.

Il devra également préciser si l'outil permet d'enrichir un nouveau numéro de téléphone en direct.

Les données contenues dans la base de sondage datent de trois années et correspondent souvent aux coordonnées téléphoniques des parents. Une phase de recherche de coordonnées est réalisée quelques mois avant le démarrage de l'enquête (marché indépendant). Toutefois, l'enquêteur sera amené fréquemment à recueillir un nouveau numéro de téléphone. Dans ce cas, une mise à jour de la base de données est nécessaire (en direct). Ce numéro devra être classé comme prioritaire vis-à-vis de l'ensemble des numéros de téléphone

disponibles. Il s'agira également de remettre le compteur à 0 tentative d'appel pour respecter le protocole d'appels (20 tentatives).

Le contractant devra démontrer sa capacité technique en détaillant les éléments suivants :

- Gérer des appels en masse
- Indiquer l'usage ou non d'un prédictif d'appels et si usage, détailler son fonctionnement
- Gérer une plateforme de grande envergure
- Gérer l'enrichissement d'un numéro de téléphone et sa position prioritaire par rapport à ceux déjà fournis par le Céreq
- Indiquer les choix techniques pour remettre à zéro le nombre de tentatives d'appels sur ce nouveau numéro enrichi
- Indiquer la méthode choisie pour la composition des numéros de téléphone

6. Les fichiers « historiques de suivi »

Le dispositif des enquêtes « Génération » a été rénové. Au-delà de l'évolution dans son architecture, un changement majeur impacte directement le mode d'interrogation des jeunes. Avec l'introduction du mode de collecte par internet, le fichier des répondants ne suffit plus pour analyser correctement les données collectées. Des évolutions dans la méthode de redressement de la non réponse nécessitent également d'obtenir des informations plus précises sur le processus de contact ainsi que de collecte des données. De fait, des données complémentaires sont demandées apparaissant comme indissociables du fichier d'enquête.

Les fichiers historiques permettent un suivi précis du comportement de chaque individu de l'échantillon. Plusieurs objectifs de traitement en cours ou *a posteriori* de la collecte : déterminer la qualité des coordonnées fournies (numéros de téléphone, mails, etc.) et des informations personnelles, calculer le degré d'instance pour pouvoir joindre un individu,... Il s'agit également de savoir quels sont les différents canaux empruntés par l'individu pour obtenir de l'information sur l'enquête, sa participation, le cas échéant, partielle ou totale à l'enquête (que l'individu soit dans le champ Céreq ou non),...

Les fichiers de données attendus seront construits au fur et à mesure avec le prestataire lors des tests du questionnaire (partie 9). Ces données de suivi devront être fournies à une fréquence hebdomadaire.

Important : chacun des fichiers demandés devra être assorti d'une documentation explicite détaillant, pour chaque variable, la nature de l'information livrée ainsi que la définition de chacune des modalités qui la compose. Ce document, de type dictionnaire des variables, permettra d'accéder rapidement à toutes les informations utiles pour exploiter les données.

6.1 Historique de contact et de tentatives de contact

Afin de suivre exactement le remplissage du questionnaire et la navigation de l'enquêté quel que soit le mode de collecte, le prestataire devra créer plusieurs fichiers historiques détaillés (7 fichiers précisément). Il s'agit d'être en mesure de déterminer, pour chacun des individus de l'échantillon, les médias utilisés afin de l'inciter à répondre à l'enquête, ainsi que les canaux par lesquels il a répondu au questionnaire. Le format des historiques attendu est sous la forme d'un fichier par medium de contact ainsi qu'un fichier récapitulatif de toutes les tentatives de contact effectuées. À noter que ces fichiers concernent l'ensemble des individus composant l'échantillon.

6.1.1 Un fichier historique d'appels téléphoniques entrants sur le numéro gratuit

Chaque appel reçu par le prestataire via le numéro gratuit, et quel qu'en soit le motif, devra être enregistré. À noter qu'une détection automatique ou, à défaut, une recherche manuelle de l'individu est prévue (partie 5.2). L'entrée de ce fichier est l'évènement « appel téléphonique entrant ». Il comprendra une ligne par appel entrant et *a minima* :

- Date et heure de l'appel
- Durée de l'appel
- Motif de l'appel (les modalités seront fournies par le Céreq ultérieurement)
- Identification de l'individu appelant (oui/un tiers/non)
- Identifiant de l'individu appelant
- Numéro de téléphone entrant
- Rang d'appel sur le numéro gratuit (nombre d'appels précédents de l'individu sur le n° gratuit +1)
- Numéro du téléphone enrichi (*le cas échéant*)

- Démarrage de l'enquête (oui/non)
- Date et heure d'un RDV programmé (*le cas échéant*)
- Commentaire éventuel de l'enquêteur
- Identifiant de l'enquêteur ayant reçu l'appel

6.1.2 Un fichier historique d'appels téléphoniques sortants

Pour chaque individu contacté et pour chaque numéro de téléphone disponible, une description détaillée de chaque appel devra être enregistrée. L'entrée de ce fichier est l'évènement « appel téléphonique sortant ». Il comprendra une ligne par appel sortant et *a minima* :

- Identifiant de l'individu appelé
- Numéro de téléphone appelé
- Date et heure de l'appel
- Indicateur du numéro sortant (classique/régionalisé)
- Affichage du numéro sortant (oui/non)
- Rang de l'appel (nombre d'appels sortants précédents pour cet individu +1)
- Rang de l'appel sur ce numéro (nombre d'appels sortants précédents pour ce numéro de téléphone +1)
- Durée de l'appel
- Classement de l'appel (les modalités seront fournies par le Céreq ultérieurement)
- Date et heure d'un RDV programmé (*le cas échéant*)
- Commentaire éventuel de l'enquêteur
- Toutes les variables issues de la phase de contact du questionnaire
- Identifiant de l'enquêteur ayant passé l'appel

6.1.3 Un fichier historique de la boîte mail Céreq

Chaque mail reçu par le prestataire, via la boîte mail Céreq et quel qu'en soit le motif, devra être enregistré. L'individu peut être amené à contacter le Céreq via la boîte mail dédiée à l'enquête (partie 5.3) à défaut du numéro gratuit. Il sera informé de l'existence de cette boîte mail via la lettre avis ou au moment du remplissage du questionnaire. L'entrée de ce fichier est l'évènement « mail entrant ». Il comprendra une ligne par mail reçu et *a minima* :

- Identifiant de l'individu
- Adresse mail
- Date et heure de réception
- Rang du mail reçu (nombre de mails précédents de l'individu sur la boîte Céreq +1)
- Motif du mail reçu (les modalités seront fournies par le Céreq ultérieurement)
- Objet du mail reçu
- Texte du mail reçu
- Traitement du mail (réponse fournie) (oui/non)
- Mode utilisé pour la réponse (mail ou téléphone)
- Date de réponse
- Identifiant de l'enquêteur ayant traité le mail

6.1.4 Un fichier historique des mails de contact

Pour chaque individu contacté par mail dans le cadre de l'envoi du mail avis ou du protocole de relance, un suivi détaillé devra être enregistré (quel que soit le motif de l'envoi).

L'entrée de ce fichier est l'évènement « mail sortant ». Il comprendra une ligne par mail envoyé et *a minima* :

- Identifiant de l'individu
- Motif de l'envoi (lettre avis, relance pour une première connexion ou pour terminer l'enquête)
- Date et heure de l'envoi
- Rang du mail envoyé (nombre de mails précédents pour cet individu +1)
- Numéro du lot (si plusieurs)
- Mail délivré (oui/non)
- Ouverture du mail (oui/non)
- Clic sur le lien de redirection vers le questionnaire (oui/non)
- Mail blacklisté (oui/non)
- Mail en échec (oui/non)
- Code de l'échec
- Motif de l'échec (erroné,...)

6.1.5 Un fichier historique des SMS de contact

Pour chaque individu contacté par SMS dans le cadre du protocole de relance (partie 7.4.3), un suivi détaillé devra être enregistré (quel que soit le motif de l'envoi). L'entrée de ce fichier est l'évènement « SMS sortant ». Il comprendra une ligne par SMS envoyé et *a minima* :

- Identifiant de l'individu
- Motif de l'envoi (relance pour une première connexion ou pour terminer l'enquête)
- Date et heure de l'envoi
- Rang du SMS envoyé (nombre de SMS précédents pour cet individu +1)
- Numéro du lot (si plusieurs)
- Accusé de réception (oui/non/perdu)

6.1.6 Un fichier historique de connexions sur la plateforme CAWI

Pour chaque connexion sur la page d'accueil de la plateforme CAWI, une description détaillée de son parcours dans le questionnaire devra être enregistrée (quel que soit le niveau de remplissage de ce dernier lors de sa déconnexion). L'entrée de ce fichier est l'évènement « connexion à la plateforme CAWI ». Il comprendra une ligne par connexion et *a minima* :

- Date et heure de la connexion sur la plateforme CAWI
- Durée sur la plateforme CAWI
- Identification (oui/non)
- Date et heure d'identification
- Identifiant de l'individu connecté
- Rang de la connexion (nombre de connexions précédentes identifiées +1)
- Date et heure d'un RDV téléphonique programmé (*le cas échéant*)
- Démarrage ou reprise du questionnaire (oui/non)

6.1.7 Un fichier historique récapitulatif

Ce fichier est une compilation de l'ensemble des six fichiers précédemment décrits. Pour chaque individu de l'échantillon, il s'agira de synthétiser l'information sur les contacts et tentatives de contact contenant l'ensemble des événements associés (6 événements possibles). Il comprendra une ligne par individu et par événement et *a minima* :

- Identifiant de l'individu
- Date de l'événement
- Événement (appel téléphonique entrant/appel téléphonique sortant/mail entrant/mail sortant/SMS sortant/connexion à la plateforme CAWI)

6.2 Historique des réponses à l'enquête

Afin d'être en mesure de déterminer, pour chacune des réponses collectées, le mode de réponse utilisé, il conviendra de réaliser un suivi précis du remplissage du questionnaire. Dans le cadre d'une enquête multimode internet-téléphone, ce fichier est indissociable du fichier des réponses final. Il est essentiel afin de permettre le traitement des réponses en post-collecte ainsi que l'analyse *a posteriori* des données par les chercheurs. L'entrée de ce fichier est l'événement « réponse enregistrée dans la base de données » et concerne l'ensemble des réponses apportées par les individus de l'échantillon. Il comprendra une ligne par réponse et *a minima* :

- Identifiant de l'individu
- Nom de la question
- Date/heure/minute/seconde de l'enregistrement
- Mode de réponse (internet/téléphone)
- Rang de la connexion pour internet (celui de l'historique des connexions)
- Utilisation de l'infobulle (oui/non/sans objet). Si plusieurs, utilisation de chacune des infobulles
- Type d'appel pour le téléphone (entrant/sortant)
- Rang de l'appel (celui de l'historique d'appels entrants ou sortants selon la réponse précédente)

7. La phase « Enquête multimode »

7.1 Informations générales sur l'enquête

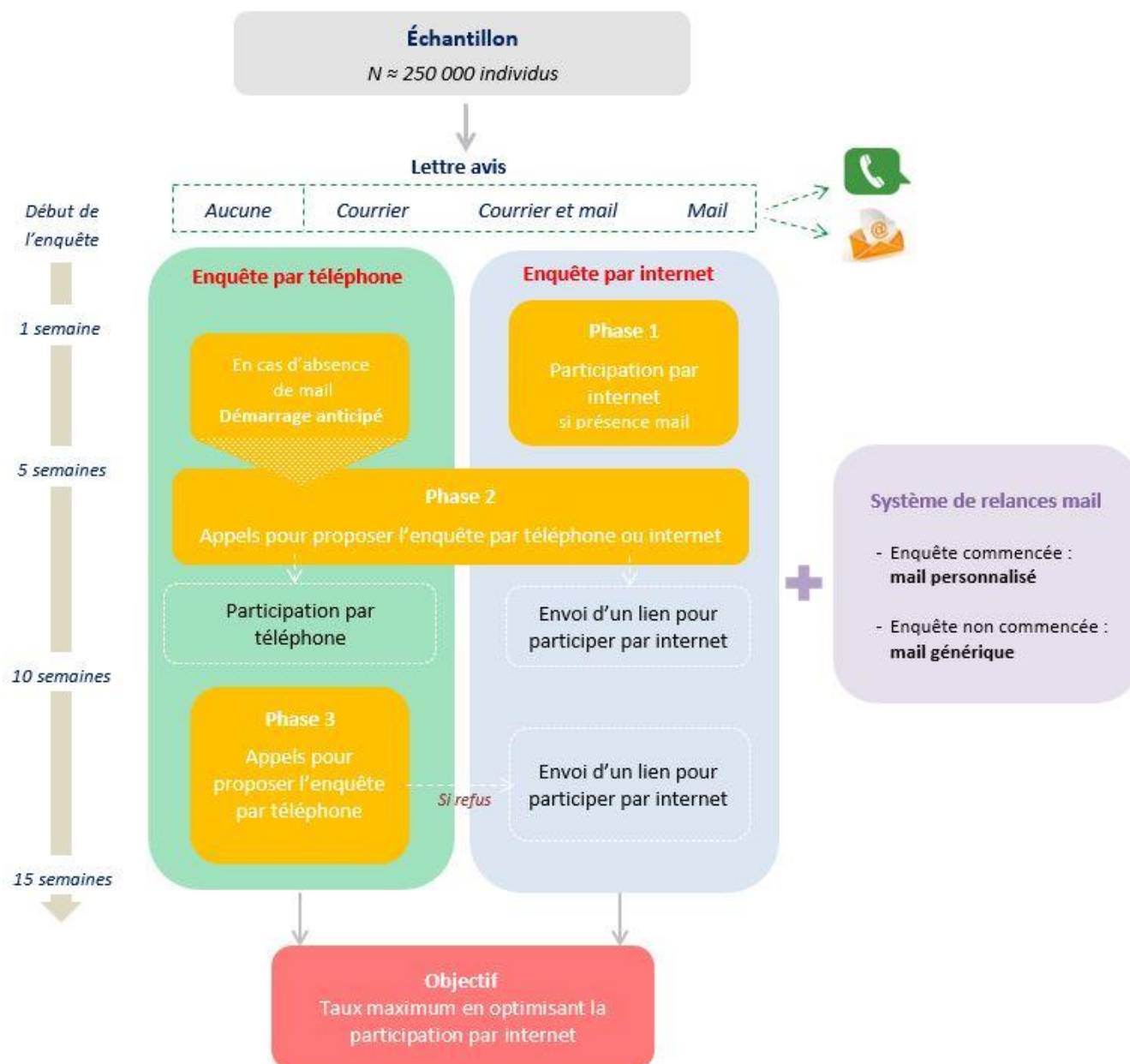
Le Centre d'études et de recherches sur les qualifications (Céreq) mène depuis vingt ans le dispositif des enquêtes Génération qui s'intéresse à l'insertion professionnelle et au cheminement des sortants du système éducatif, tous niveaux de formation confondus, lors de leurs premières années de vie active. Ces enquêtes sont historiquement administrées avec une collecte monomode téléphone.

Depuis 2012, le Céreq est dans une démarche de réflexion pour rénover son dispositif d'enquêtes Génération. Cette initiative résulte d'un travail de refonte à caractère scientifique et d'une adaptation technologique visant à moderniser le questionnaire. L'implémentation d'internet, dans la collecte historiquement téléphonique, permet par ailleurs de toucher une génération de jeunes connectés, peu joignables par téléphone, ou résidant à l'étranger, non couverts actuellement par l'enquête.

Dans la version rénovée de l'enquête Génération, les individus seront contactés par mail et/ou par téléphone et auront la possibilité de répondre par internet ou par téléphone. Les protocoles de contact et les possibilités de réponses sont multimodes. La passation en auto administré devra pouvoir se faire par internet aussi bien sur ordinateur, smartphone et/ou tablette (tout système d'exploitation, tout navigateur et toute taille d'écran).

La collecte reposera sur un échantillon entre 200 000 et 300 000 individus pour obtenir un fichier de répondants dans le champ Céreq entre 20 000 et 45 000 répondants. La participation par internet est estimée à 50 %.

Figure 1 - Logigramme du protocole de l'enquête multimode



7.2 Le protocole de collecte multimode internet-téléphone

Le protocole d'enquête est construit autour de trois phases distinctes dans le temps et privilégie une collecte au préalable internet. Des plans de dialogue adaptés sont associés à chacune des phases et seront intégrés directement dans le questionnaire (partie phase de contact). Au démarrage du plateau, la première semaine sera dédiée à observer l'effet de la lettre avis sur la participation. Le plateau téléphonique devra être dimensionné pour réceptionner les éventuels appels sur le numéro gratuit et répondre aux mails reçus sur la boîte Céreq.

7.2.1 Phase 1 - Incitation à répondre par internet

Une première phase de quatre semaines comprend l'envoi de la lettre-avis destinée à informer du démarrage de l'enquête intégrant un lien personnalisé pour participer directement par internet. Le protocole diffère selon la disponibilité ou non du mail dans le fichier d'import.

- *Présence d'un mail* : À la suite de ce mail avis, des mails de relance seront envoyés en variant les jours et heures d'envoi. Il est prévu de faire au moins une relance par semaine. Si l'individu enquêté a un mail invalide, celui-ci sera rebasculé dans le traitement ci-dessous.

- *Absence de mail* : Pour les individus n'ayant pas de mail dans le fichier d'import, des tentatives d'appels seront réalisées. Le plateau téléphonique sera dimensionné par le prestataire pour qu'au bout de trois semaines au moins un contact par téléphone ait été effectué pour l'ensemble des individus. Il s'agira de qualifier les individus, c'est-à-dire de confirmer que c'est la bonne personne et qu'elle appartient au champ de l'enquête. Si cette confirmation est réalisée, sera proposée la préférence entre les deux modes de collecte. Si le choix se porte sur internet, une adresse électronique sera demandée avec confirmation aux individus qualifiés. Par la suite, un mail contenant le lien pour répondre par internet sera envoyé immédiatement. Si la personne enquêtée n'a pas de mail ou souhaite répondre par téléphone, la passation se fera par téléphone. Le plan de dialogue permettra de s'assurer du consentement de l'enquêté à répondre par internet pour éviter au maximum que ce mode de réponse puisse être une forme de non réponse.

7.2.2 Phase 2 - Choix de répondre par internet ou téléphone

Une seconde phase de cinq semaines où le plateau téléphonique est lancé pour l'ensemble de l'échantillon. Le choix est donné aux enquêtés contactés par téléphone de participer soit par internet soit par téléphone selon leur préférence. Cette phase se caractérise par une alternance entre les relances par mail et les tentatives d'appels téléphoniques. Est prévu un envoi de mail par semaine et l'ensemble de l'échantillon devra être contacté par téléphone au moins une fois (notamment ceux avec présence mail dans la phase 1). Dans cette étape de la collecte, le plan de dialogue est identique à la phase 1.

7.2.3 Phase 3 - Rattrapage pour répondre par téléphone

Une dernière phase de collecte d'une durée de cinq semaines est prévue pour proposer aux enquêtés de participer par téléphone. Le plan de dialogue de l'enquêteur sera adapté pour proposer uniquement de répondre par téléphone. En cas de demande spontanée de l'enquêté pour répondre par internet, l'enquêteur devra tout de même inciter à répondre par téléphone. En cas de refus de réponse par téléphone, le lien du questionnaire sur internet sera envoyé après confirmation de l'adresse mail de l'enquêté.

7.3 Le protocole de contact

7.3.1 Durée totale du terrain

Le protocole d'enquête multimode internet-téléphone prévoit la mise en place d'un plateau téléphonique sur 3 mois et demi, d'avril à juillet 2020. Avec l'introduction du mode de collecte internet, le terrain d'enquêteurs a une fonction multiple et sa taille est variable selon la temporalité de la collecte.

La plateforme de réponse sur internet devra être ouverte le premier jour de réception des lettres-avis postales afin que les individus puissent participer directement à l'enquête. Le prestataire devra envoyer les lettres avis mail le jour du démarrage du terrain.

7.3.2 Consignes pour le premier contact

Il est essentiel que ce soient les individus figurant dans le fichier qui répondent personnellement au questionnaire. Le prestataire veillera à ce que l'individu, contacté par téléphone, ne soit ni un parent, conjoint ou toute autre personne qui réponde à la place de l'enquêté lui-même, sauf en cas de handicap empêchant la communication (exemple : cas d'une personne muette).

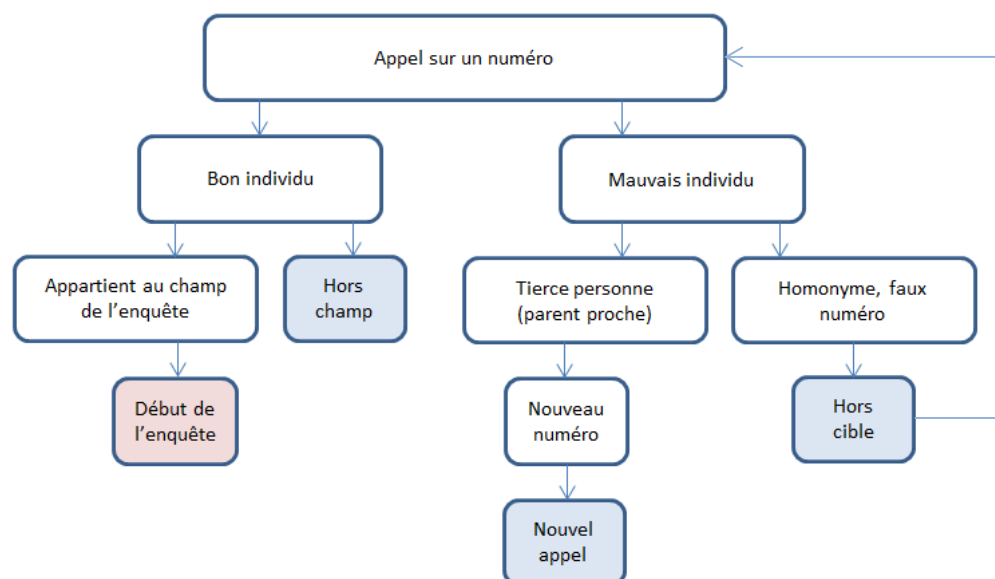
Le questionnaire débutera par un plan de dialogue permettant de gérer le contact avec l'individu ou un tiers. Ce plan de dialogue initial sera conjointement développé par le prestataire et le Céreq. De nouveaux numéros de téléphone pourront être collectés lors de cette phase. Par exemple, si l'individu contacté sur un numéro de

téléphone portable dispose d'un numéro fixe ou si l'individu contacté est un tiers qui accepte de donner un nouveau numéro. Dans ce cas, il faudra positionner ce nouveau numéro de téléphone en première position de la liste des numéros disponibles de l'individu et appliquer la règle de rappel (partie 5.4). Ce plan de dialogue initial doit également permettre de gérer des prises de rendez-vous éventuelles pour un rappel ultérieur.

La partie introductive comportera quelques questions ayant pour objectif de déterminer si l'individu en ligne est celui qui est recherché. Ainsi, il faudra :

- Confirmer le nom et le prénom : si la personne n'est pas l'individu mais la connaît, il faudra essayer d'obtenir le numéro de téléphone de l'individu recherché ;
- Être en mesure d'identifier l'individu recherché : si le nom et le prénom coïncident, quelques questions vont permettre de vérifier si c'est le « bon individu » qui est en ligne ;
- Administrer la suite du questionnaire : proposer le choix du mode de collecte, envoyer un mail si le choix est internet, effectuer la passation du questionnaire dans son intégralité, prendre un rendez-vous téléphonique si nécessaire ou souhaité par l'enquêté ;
- En cas de confirmation par un tiers indiquant que l'individu, correctement identifié, n'est définitivement pas joignable (décès, parti à l'étranger,...), l'entretien sera terminé. Le numéro de téléphone sera classé parmi les non succès avec son motif et aucun autre numéro de téléphone associé au même individu dans le fichier de l'échantillon ne sera composé ;
- En cas de confirmation de l'identité par l'individu lui-même et malgré les arguments livrés sur l'importance de l'enquête, s'il exprime son refus, l'entretien sera terminé. Le numéro de téléphone sera classé en refus de réponse et aucun autre numéro de téléphone associé à cet individu ne sera composé ni aucune relance (mail ou SMS) ;
- En cas de non confirmation de l'identité de l'individu, l'entretien sera terminé et le numéro de téléphone sera classé « hors cible ». Le numéro de téléphone suivant associé au même individu dans le fichier de l'échantillon sera alors utilisé ;
- En cas d'impossibilité de joindre l'individu à un numéro de téléphone donné après 20 tentatives à des moments différents de la journée et à des jours différents de la semaine (non réponse au téléphone ou répondeur), le numéro de téléphone sera classé impossible à joindre. Le numéro de téléphone suivant associé au même individu dans le fichier de l'échantillon sera alors utilisé.

Figure 2 – Phase de contact



Une fois le questionnaire démarré, le prestataire devra tout mettre en œuvre pour que l'interrogation soit dispensée entièrement et sans interruption (en particulier lors du changement d'équipes, des pauses méridiennes, etc.).

Si l'enquête a été interrompue inopinément en cours de passation, l'enquêteur essaiera de joindre immédiatement l'enquêté pour terminer le questionnaire par téléphone. Si l'enquêté souhaite interrompre la passation, un rendez-vous sera fixé pour finir l'enquête. En cas de refus strict de continuer par téléphone et de prendre rendez-vous, l'enquêteur lui proposera le mode alternatif internet. Le système informatique devra prévoir la reprise du questionnement là où il s'est arrêté.

Important : le prestataire devra proposer un calendrier de prise rendez-vous (en l'adaptant pour tenir compte du décalage horaire pour les DROM) et honorer les rendez-vous pris aux jours et heures indiqués par l'individu.

7.3.3 Numéros de téléphone et tentatives

Vingt tentatives d'appels seront effectuées avant de passer au numéro de téléphone suivant. Le système proposé par le contractant devra gérer automatiquement ce passage à un autre numéro de téléphone pour un individu à partir du moment où le précédent numéro de téléphone a été tenté 20 fois (ou classé avant), à des heures et des jours différents, suivant un protocole identique pour tous. Chacun des numéros de téléphone devra être exploité jusqu'au classement définitif de la fiche individuelle (enquête réalisée, refus, etc).

Les règles de rappel seront définies par le Céreq en collaboration avec le prestataire avant le lancement de la collecte. Le paramétrage des appels sera identique sur toute la période d'appel et pour chaque individu (variant jours et heures d'appels) sauf indication contraire donnée par le Céreq en cours de plateau d'enquête.

Le prestataire devra prévoir d'alterner entre des appels effectués avec un numéro de téléphone visible par l'appelé et des appels effectués avec un numéro masqué. Si un individu, d'office, n'accepte pas les numéros masqués, un démasquage devra être effectué sur la totalité des tentatives restantes (afin que les 20 tentatives soient effectives sur ce type de numéro).

Les différents numéros de téléphone disponibles pour chaque individu (au maximum six) devront être essayés selon l'ordre établi par le Céreq. En effet, les numéros de téléphone transmis pour chaque individu seront classés selon un ordre respectant une probabilité décroissante d'être le téléphone de l'individu ou de membres de sa famille. Toutefois, le nombre de tentatives d'appels pourra être augmenté et l'ordre de traitement des numéros de téléphone modifié en cours d'opération à la demande du Céreq.

Le prestataire s'engage à contacter l'ensemble des individus de l'échantillon selon le protocole décrit (20 tentatives par numéro). L'objectif quantitatif de l'interrogation est d'exploiter de façon homogène et poussée le fichier des individus transmis par le Céreq. Ces exigences ne pourront être satisfaites que si le nombre d'enquêteurs dédiés à l'opération est adéquat. Le plateau sera dimensionné pour permettre de réaliser l'intégralité de l'enquête avec le protocole défini avant la date de fin de plateau.

Le candidat précisera quel plan d'action pourra être mis en place en cas de production jugée insuffisante en cours de plateau au regard des objectifs fixés et à quelle vitesse celui-ci pourra être mis en œuvre.

7.4 Le protocole de relances multimode

Les relances seront réalisées par internet, par téléphone, par message laissé sur répondeur et/ou par SMS. En parallèle des relances téléphoniques propres à chaque phase, des relances mails hebdomadaires seront programmées ; les mails sont soit génériques dans le cas de questionnaire non commencé, soit personnalisés dans le cas contraire. La relance peut également se faire ponctuellement par SMS.

En complément de la lettre-avis, un système de relances est organisé autour du statut de l'enquêté une semaine après la réception de celle-ci. Pour les individus n'ayant pas encore répondu à l'enquête et disposant d'un mail valide, une relance mail hebdomadaire est envoyée variant le jour et l'heure d'envoi. Sinon, pour ceux dont l'adresse mail est invalide ou indisponible, des tentatives d'appels sont réalisées.

7.4.1 Les contacts téléphoniques et message sur le répondeur

Pour chaque individu du fichier de l'échantillon, il est prévu 20 tentatives d'appel par numéro de téléphone disponible. De même que les relances par internet, les plans de dialogue des relances par téléphone seront différents si le questionnaire a été commencé et la selon phase de collecte durant laquelle la relance téléphonique est faite.

Une procédure pour laisser des messages sur le répondeur par l'enquêteur sera demandée en milieu d'enquête (mi-mai). Ce message sera élaboré par le Céreq. Il présentera l'enquête Génération en quelques mots et indiquera le numéro gratuit afin que l'individu puisse rappeler et prendre rendez-vous ou encore réaliser l'enquête.

La gestion de ces dépôts de message sur répondeur devra être précisée par le contractant. En effet, il est important qu'un seul message sur le répondeur soit laissé sur une courte période.

7.4.2 Les relances mails

- Deux types de relances par mail seront effectués :

Les relances mail **planifiées** seront effectuées auprès des individus qui n'ont pas commencé le questionnaire (qu'ils aient été contactés par mail ou par téléphone). Un calendrier sera fourni par le Céreq indiquant les jours d'envoi de ces mails. Ces relances pourront s'étaler sur l'ensemble de la période de collecte à raison d'un mail par semaine.

Des relances mails **pour terminer l'enquête** seront envoyées aux individus qui ont commencé et n'ont pas terminé le questionnaire. Dans ce dernier cas, des relances par mail seront faites à J+4, J+8., « J » correspondant au jour où l'individu a commencé le questionnaire.

Le prestataire devra pouvoir montrer sa capacité à programmer des relances mails.

- Les modèles de mails à envoyer :

Le prestataire devra envoyer plusieurs types de mail : le mail de relance pour ceux qui n'ont pas encore commencé à répondre et le mail de relance pour ceux qui ont commencé le questionnaire sans le terminer. Les modèles de ces différents messages seront fournis par le Céreq. Pour chacun de ces mails, il est demandé au

prestataire de créer des liens de connexion personnalisés (intégrant l'identifiant et le mot de passe de connexion), qui permettront d'accéder directement au questionnaire en ligne. Les paramètres d'authentification habituels sont le mois, l'année de naissance et un mot de passe. Les objet et nom d'expéditeur devront faire explicitement référence au Céreq et à l'enquête Génération. Les liens web présents dans les mails devront être des liens personnalisés intégrant l'extension « cereq.fr », ils seront redirigés vers les adresses des serveurs du contractant.

Le candidat démontrera sa capacité à envoyer des mails en masse et précisera les possibilités d'automatisation (totale ou partielle) des requêtes pour l'envoi des mails de relance ainsi que le nombre de champs personnalisables possible dans un mail.

Le mail de relance pour ceux qui n'ont pas encore répondu

Ce courriel sera envoyé aux individus n'ayant pas commencé le questionnaire. Il est important que le système proposé soit automatisé. En effet, les envois de mail se feront à des jours et des horaires différents (parfois le week-end ou le soir). Il est important que la requête de sélection des individus concernés par une relance se fasse juste avant l'envoi du message. En effet, si un mail est programmé le lundi matin mais que la requête a été effectuée le vendredi en fin de journée, tous les individus ayant répondu le week-end se feront relancer à tort. Le contractant précisera la solution retenue.

Le mail de relance pour ceux qui ont commencé sans terminer

Ce courriel sera envoyé uniquement aux individus ayant commencé l'enquête par internet sans l'avoir terminé. Les temporalités d'envoi de ces messages seront précisées lors de l'exécution du marché.

7.4.3 Les relances par SMS

Une relance par SMS sera faite aux individus ayant un numéro de téléphone portable et qui n'ont pas encore répondu à l'enquête au moment du passage dans la deuxième phase de collecte. Le contenu du message sera fourni par le Céreq. Ce message contiendra un lien pour se connecter directement au questionnaire en ligne. Le Céreq demandera aussi l'envoi de sms en cours d'opération pour inciter les individus à répondre à l'enquête.

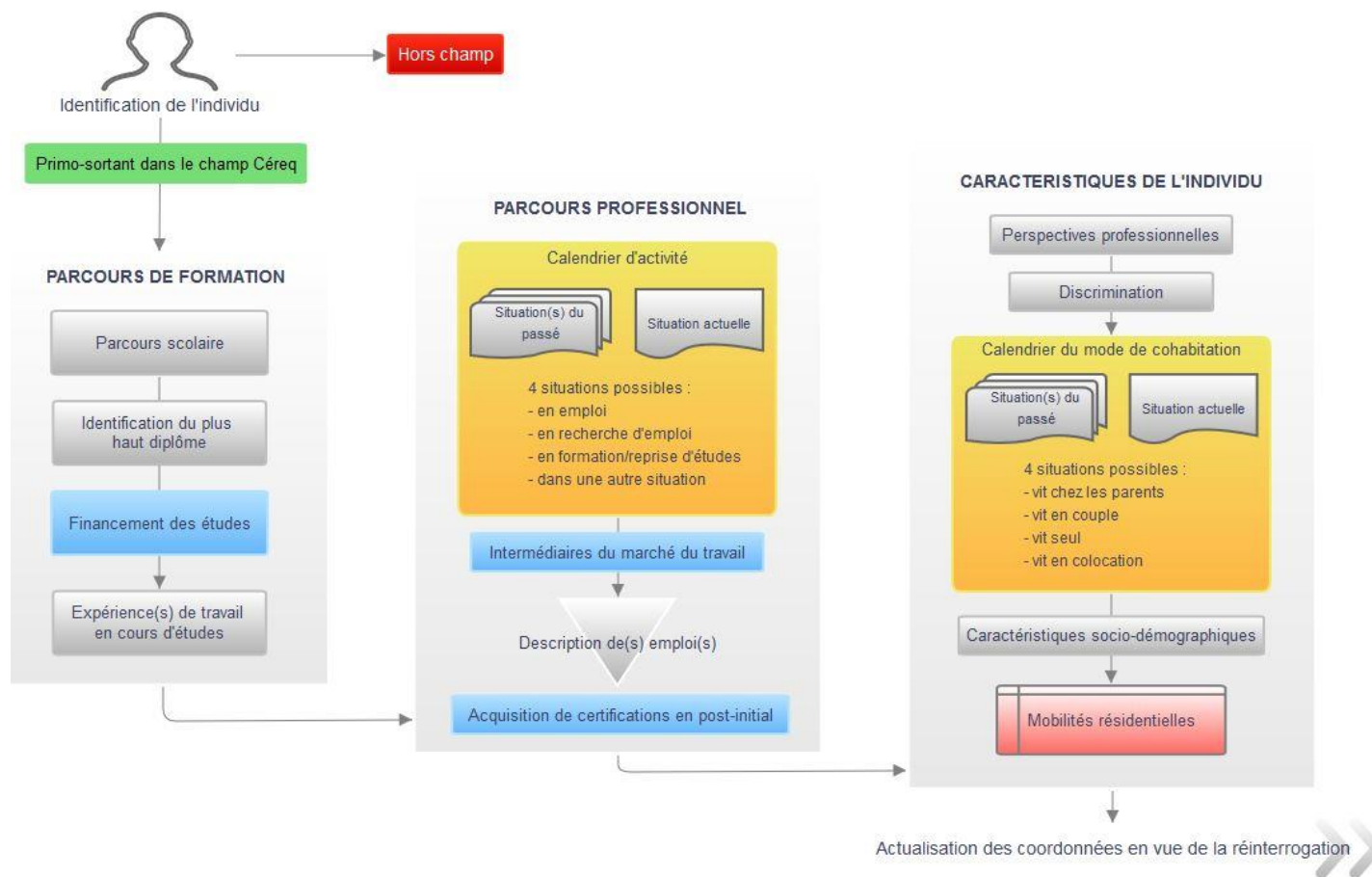
Plusieurs envois seront prévus par numéro de mobile sur tout ou partie de la population encore en production. Ce SMS informera les jeunes de l'enquête et leur notifiera le numéro gratuit pour réaliser l'enquête.

Le contractant devra préciser dans sa réponse les solutions d'envoi associées à cette opération.

8. Spécificités de développement informatique du questionnaire

8.1 Contenu du questionnaire de l'enquête multimode

Figure 3 - Logigramme (provisoire) du questionnaire de l'enquête



Marché d'appel d'offre européen

« 1^{ère} interrogation auprès d'un échantillon de sortants du système éducatif français en 2017 »
Enquête multimode (téléphone, internet)

8.2 Fourniture du questionnaire multimode par le Céreq

Le questionnaire de l'enquête de 2020 à administrer sera transmis par le Céreq en version électronique. Il a été rédigé au format Word (extrait en Annexe 1). Ce questionnaire unique intègre les spécificités de chacun des modes de collecte internet et téléphone. Certaines questions présenteront des variantes de formulation selon le mode de passation. Il est essentiel de respecter ce travail de rédaction lors de la programmation pour collecter une information de qualité. Le Céreq sera vigilant lors du développement de l'outil.

Le questionnaire intègre une partie importante de filtres qui sont de nature diverse :

- des filtres simples qui définissent la population d'intérêt pour une question ou module donné ;
- des filtres d'affichage ou d'adaptation de l'intitulé de la question (en fonction d'informations importées) ;
- des filtres de type renvoi (pour sauter une ou partie du questionnaire) ;
- des filtres de recode (imputer une information du fichier d'import corrigée de l'enquêté) ;
- des filtres intégrant une fonction (de calcul) ;
- des filtres qui permettent de tester la cohérence des réponses.

La numérotation des questions a également son importance (détail dans le questionnaire).

Le questionnaire d'enquête devra être développé par le prestataire sous un système de collecte multimode internet-téléphone (partie 1.1). Le candidat précisera les possibilités de tests automatiques du questionnaire (simulation, robots,...).

8.3 Personnalisation du questionnaire et charte graphique

Pour attester du caractère officiel de l'enquête, il est nécessaire de créer un outil de collecte personnalisé. Les individus contactés seront amenés à participer sans jamais avoir recours à un enquêteur. Il est donc essentiel que l'individu soit rassuré.

Il doit tout de suite avoir un visuel avec des éléments qui démontrent la crédibilité de l'enquête Génération (logo, construction des pages,...) notamment au travers de la personnalisation. Devront également apparaître le numéro gratuit et l'adresse mail de contact, accessible sur clic, pour toute question. Une liste non exhaustive des éléments d'identité du Céreq se trouve en annexe 2.

La présentation et les messages doivent être soignés notamment sur la page d'accueil (pour une première connexion ou une reprise d'enquête), sur la dernière page du questionnaire mais également lorsque l'individu clique sur le bouton « quitter et revenir plus tard ». Le Céreq a des contraintes éditoriales, c'est pourquoi sont prévus des éléments de texte de présentation du Céreq avec des liens sur le site mais aussi de redirection vers les réseaux sociaux. Également, seront inclus des éléments visant à rassurer et à inciter à la poursuite de l'enquête.

Chaque questionnaire visualisé à l'écran devra intégrer les informations nécessaires à l'identification de l'individu. Ces informations figureront dans le fichier d'import fourni par le Céreq. De même, certaines informations saisies en cours d'enquête seront appelées (ou feront l'objet d'un calcul) à s'afficher à un autre moment dans le questionnement, permettant ainsi de vérifier la cohérence et/ou de collecter une information additionnelle lors de la suite de l'entretien (diplôme, nom de l'entreprise, commune, profession etc.)

Les critères de personnalisation demandés peuvent être discutés sur la base de propositions adaptées fournies par le prestataire.

Tableau 2 – Descriptif des critères de charte graphique demandés

Critère	Description
Police, couleur et taille	<p>La charte du questionnaire devra concorder avec la charte graphique du site internet du Céreq. Une liste de police (non exhaustive) pourra être utilisée :</p> <ul style="list-style-type: none"> - IBM Plex sans condensed - Cavar Dream <p>Concernant les couleurs, hormis le noir, le blanc et le gris, cinq couleurs pourront être appliquées. Les codes RVB sont les suivants :</p> <ul style="list-style-type: none"> - 0/139/153 - 239/83/80 - 37/98/153 - 248/172/0 - 123/154/98 <p>Le Céreq fournira davantage d'éléments de charte au moment du développement du questionnaire.</p>
Disposition des objets sur une page	Homogénéisation des pages dans la disposition des objets. Le remplissage du questionnaire doit être simple. L'effet d'apprentissage doit permettre à l'individu de ne pas avoir à chercher les éléments sur la page. Le scrolling doit être évité.
Boutons de fonctionnalité	Lorsque la souris survole un bouton, il doit être mis en surbrillance. Les boutons « suivant » et « précédent » devront être situés en bas de chaque page et devront être fixes. Le bouton « précédent » devra être plus petit que le bouton « suivant ». Un bouton « quitter et revenir plus tard » devra aussi être présent. Sur le dernier écran du questionnaire devra apparaître un bouton « valider le questionnaire ». Ce bouton devra être suffisamment visible afin que l'individu n'oublie pas de cliquer dessus.
Titre de partie et Fil d'Ariane	Un positionnement fixe des objets est indispensable. Le fil d'Ariane dynamique doit intégrer un changement de couleur ou une surbrillance du nom de la partie, fonction de l'avancée dans l'enquête. Prendre en compte le cas des reprises de questionnaire.
QCM représenté par un tableau	<p>Pour collecter des informations comparables dans le cadre d'une collecte multimode, le format des questions à choix multiples est un tableau croisé oui/non.</p> <p>Présentation dynamique du QCM : pour faciliter le remplissage, un changement de couleur devra survenir à chaque survol de souris sur une modalité.</p>
Menu pour une recherche en autocomplétion	Tous les menus de l'enquête seront accessibles par une recherche par complétion automatique. L'affichage du champ de recherche doit être simple et permettre une recherche intuitive avec une proposition d'échos à partir de 3 lettres saisies. Lorsque l'individu ne trouve pas dans la liste, la modalité « autre » est sélectionnée et doit permettre l'affichage instantané d'un champ de saisie.
Calendrier, Tableau de type récapitulatif de saisie	Adapter les calendriers et les tableaux (mobilités résidentielles et langues) en fonction de la charte graphique du questionnaire, sans oublier les <u>pop-ups</u> . Pour le calendrier d'activité, prévoir des visuels pour le besoin d'aide ainsi que pour la vidéo de démonstration. À noter que pour les messages d'erreur, des pop-ups devront s'afficher pour alerter d'une saisie erronée.
Infobulle	Lorsque la souris survole l'infobulle, son contenu doit apparaître sans cacher le texte. Pour ne pas influencer le choix de réponse, la couleur et le positionnement dans la page doivent inviter à la discrétion.
Factorisation de questions	L'affichage général des questions est une question par page. Toutefois, des questions qui portent sur la même thématique ont été regroupées pour alléger le questionnement (indication prévue dans la rédaction du questionnaire).

Une charte graphique plus détaillée sera transmise au prestataire lors du démarrage du développement du questionnaire. Toutefois, le Céreq examinera les propositions faites par le prestataire, le cas échéant.

8.4 Développement du questionnaire et contraintes techniques imposées

Le Céreq souhaite le développement d'un outil unique permettant, à la fois, de collecter de façon simplifiée des données par l'intermédiaire d'un enquêteur et optimal pour une participation à l'enquête par internet. Le produit devra être multimode (téléphone et internet) et connecté à une plateforme de gestion d'appels téléphoniques.

Le candidat devra présenter de façon détaillée le matériel informatique utilisé sur le plateau téléphonique et la gestion des appels. Il devra également décrire les logiciels et langages de développement choisis pour le développement du questionnaire.

L'outil de collecte attendu doit permettre de pouvoir naviguer dans un seul et unique environnement pour limiter les interventions de l'enquêteur et le risque d'abandon de l'internaute. Le déroulé du questionnaire repose en partie sur les variables issues de la base de sondage (partie 3) mais aussi à partir de l'information collectée en cours d'enquête.

Ce questionnaire comprend deux parties :

- Une première partie dont l'objectif est de déterminer si l'individu contacté est celui qui est recherché. L'identification est réalisée à partir de ses nom/prénom/mois et année de naissance. En cas d'invalidation, cet entretien sera terminé et le numéro sera classé « hors cible ». Si l'identification de l'individu est confirmée, il faudra vérifier son éligibilité à l'enquête. Si l'individu ne remplit pas les critères, il sera classé « hors champ ».
- Une seconde partie qui correspond au questionnaire sur l'insertion professionnelle de l'individu. Cette partie est composée du calendrier d'activité mensuel et de modules de questions.

Selon le type de population à enquêter, des questions ou des filtres spécifiques seront à développer.

Afin de faciliter le travail de l'enquêteur et la saisie en auto-administré, le questionnaire développé devra respecter les contraintes techniques suivantes :

Tableau 3 – Descriptif des techniques de développement demandées

Technique	Description
Menu pour une recherche en autocomplétion	Recherche en « complètement automatique » (autocomplétion) Ce mode de recherche permet à l'utilisateur de limiter la quantité d'informations à saisir et ainsi améliorer la fluidité du menu. Au fur et à mesure de la saisie, une ou plusieurs modalités seront proposées à la volée (sans avoir besoin de valider leur saisie par « Entrée »). Il s'agira des modalités du menu qui incluent la chaîne de caractère saisi. Ce mode de recherche doit détecter rapidement la chaîne de caractère saisi dans l'ensemble du menu et sur chaque mot de chaque modalité (<i>i.e.</i> pas de recherche uniquement sur le premier mot, mais sur l'ensemble de la chaîne de caractère de chacune des modalités). De plus, si la recherche inclut plusieurs mots, il faudra rechercher chacun des mots. Par exemple : si recherche du couple de mots « bac électricité », le menu devra proposer l'item « bac professionnel électricité ». Une fois le bon libellé sélectionné, le questionnaire doit-être en mesure de récupérer un ou plusieurs codes ou variables associées à ce libellé dans le menu.

Infobulle	Certains termes présents dans les intitulés de question ou de modalité doivent être définis pour une meilleure compréhension de l'enquêteur ou de l'enquêté. Ces définitions devront être intégrées à l'aide d'infobulles à proximité des termes en question. Ce système d'infobulles ne doit pas alourdir le questionnement.
Contrôle de cohérence	Les masques de saisie devront impérativement intégrer des contrôles de plages (définis dans le questionnaire) et assurer la cohérence entre les modules de questions. Les messages d'erreurs peuvent être du type « pop-up » ou s'afficher au-dessus du libellé de la question. Certains contrôles de cohérence, non prévus dans le questionnaire initial, pourront être demandés suite aux tests du questionnaire.
Question avec un champ de saisie libre ou « autre précisez »	Les champs de saisie libre « texte » devront prévoir un masque de saisie en majuscule, sans accent et pour ceux « numérique » un masque de saisie permettant de bloquer la saisie de texte. La taille du champ de saisie devra être adaptée à l'information collectée. Exemples pour la question ouverte : champ contenant plusieurs lignes de saisie (retour à la ligne compris) et pour l'âge : un champ numérique avec 3 caractères maximum.
Minutage en cours de questionnement	Le questionnaire devra intégrer des « tops » pour estimer au mieux le minutage par module de questionnement (idéalement un minutage par question). Il faudra prévoir la gestion des retours arrière dans le calcul des tops (temps passé sur la question + le temps passé pour le changement de réponse). L'opportunité d'afficher le minutage en cours de questionnement pour informer l'enquêteur du temps passé en ligne avec son interlocuteur sera discutée.
Fil d'Ariane, arborescence du questionnement	Le fil d'Ariane (ou sommaire) représente les différentes parties/thématiques abordées dans l'enquête. Il permet à l'individu de se situer rapidement dans le questionnaire. La barre de progression n'étant pas adaptée à l'enquête Génération, ce fil permet également de fournir une approximation du temps restant avant la fin du questionnement. Le candidat précisera comment créer un fil d'Ariane <u>dynamique</u> ayant la capacité de se mettre à jour au moment de la reprise de questionnaire. Le questionnement étant adapté aux informations déclarées dans le calendrier, il faudra adapter en conséquence l'arborescence (à mettre à jour au moment de la validation du calendrier).
Titres dynamiques	Afin de faciliter le repérage dans le questionnaire, celui-ci devra intégrer des titres de partie dynamiques sur chaque masque de saisie. Certains titres seront calculés à partir de réponses à certaines questions préalables du questionnaire, en particulier du calendrier d'activité (exemple : période d'emploi du <datedebut> à <datefin> dans l'entreprise <entreprise>).
Reprise de l'enquête à la dernière question renseignée	Le candidat devra préciser comment se déroule la reprise d'un questionnaire entamé notamment par le mode internet. L'internaute (ou l'enquêteur) devra pouvoir atteindre la dernière question renseignée (ou celle située immédiatement après) suite à sa nouvelle connexion.
Passage automatique à la question suivante	Pour alléger la passation du questionnaire, une alternative au clic sur le bouton suivant (en dehors des questions spécifiques telles que le QCM, le tableau ou le menu de recherche) est le passage automatique à la question suivante lors de la sélection d'une modalité de la question en cours. La sélection d'une modalité doit pouvoir se faire manuellement pour l'enquêteur (saisie du numéro de la modalité).
Affichage d'éléments de contacts	Tout le long du questionnaire le numéro gratuit et l'adresse de contact devront être visualisables. L'individu pourra sur un simple clic pouvoir transmettre un mail au Céreq.

Tout changement de logiciel en cours de développement doit être dûment justifié et validé au préalable par le Céreq.

Contraintes spécifiques de développement :

Le questionnaire Génération prévoit la collecte de communes à différents moments du parcours scolaire et professionnel et, parfois, des adresses complètes. Une application sous forme de menu de recherche par complétion automatique, dans les fichiers de la Poste intégrant la normalisation des adresses, fait l'objet d'un marché indépendant. Est incluse également dans ce marché la collecte des établissements/entreprises, via le pop-up emploi du calendrier d'activité, où l'individu a travaillé sur les trois dernières années de vie active. Cette recherche sera réalisée en direct dans la base entreprise SIRENE.

À noter, le candidat retenu pour la prestation du présent marché peut être différent de celui choisi pour la prestation de la collecte des communes et des données entreprises. Le candidat du présent marché devra réaliser un travail collaboratif avec ce dernier pour l'intégration de cet outil nécessaire au bon déroulement de la collecte et indispensable à la qualité des informations recueillies.

La rénovation de l'enquête Génération inclut également la rénovation de la nomenclature et du protocole de collecte de la profession (PCS) initiée par le Cnis et mise en œuvre par L'Insee. Le nouvel outil de collecte revêt, une nouvelle fois, la forme d'un menu par complétion automatique. Cette application sera fournie par l'Insee. Les modalités des échanges ne sont pas encore définies ; ils se feront soit directement avec l'Insee, soit par l'intermédiaire du Céreq.

Ces applications devront être intégrées dès le premier test en juin 2019 pour s'assurer de leur bon fonctionnement, quel que soit le mode de collecte, en conditions réelles.

Le candidat devra démontrer sa capacité technique à développer un questionnaire multimode en détaillant les éléments suivants :

- Pertinence de l'outil de développement choisi pour le développement du questionnaire général (pertinence au vu des contraintes Céreq et expérience de développement sur ce logiciel) ;
- Capacité technique sur le logiciel choisi (import du questionnaire Céreq à partir d'un fichier Word ou Excel existant, expérience en ergonomie de questionnaire, pop-up, arborescence ou fil d'Ariane, etc.) ;
- Gestion de la personnalisation des questionnaires par des variables fournies en import, et/ou saisies au cours du remplissage du questionnaire ;
- Intégration de masques de saisie avancés ;
- Contrôles automatiques de plages de saisies et messages d'erreurs ;
- Intégration de définitions par infobulles pour aider le télénquêteur en cas de question de l'enquête ;
- Retour-arrière et capacité à modifier une information déjà saisie (avec mise à blanc automatique), y compris dans le calendrier ;
- Menus déroulants par saisie « complètement automatique » (autocomplétion) et récupération d'une ou plusieurs variables associées ;
- Capacité à gérer un remplissage automatique (ex : saisie du code postal et chargement automatique du nom de la commune) ;

8.5 Le calendrier interactif

8.5.1 Le calendrier d'activité

Description de l'outil

Le calendrier d'activité est le cœur du questionnaire de l'enquête Génération. Il doit permettre de reconstituer mois par mois la chronologie des situations passées de la date de sortie du système éducatif en 2016-2017 jusqu'à la date de l'enquête au printemps 2020, soit environ 3 ans sur le marché du travail.

Il est indispensable que le calendrier soit parfaitement intégré dans le questionnaire multimode. Le calendrier d'activité permet de piloter l'ouverture de pop-up, détaillant *a minima* les situations, et de modules de questions spécifiques associés aux situations professionnelles décrites.

Le calendrier d'activité est un tableau à double entrée qui croise les mois et les situations connues suivi d'un récapitulatif de saisie modifiable. La plage de saisie est fluctuante entre la date de fin des études, déclarée dans le questionnaire filtre entre octobre 2016 et décembre 2017, et la date de l'enquête en 2020, d'avril à juillet.

Ci-après, un calendrier décrivant le parcours professionnel d'un individu ayant quitté le système éducatif pour la première fois en mars 2017 et interrogé en mai 2020. Un tableau récapitulatif permet une ultime validation du parcours notamment avec l'enquêteur (*le cas échéant*).

Figure 4 – Visuel du calendrier d'activité de l'enquête 2020 (exemple)

	2017												2018												2019												2020						
	jan	fev	mars	avr	mai	juin	juil	août	sept	oct	nov	déc	jan	fev	mars	avr	mai	juin	juil	août	sept	oct	nov	déc	jan	fev	mars	avr	mai	juin	juil	août	sept	oct	nov	déc	jan	fev	mars	avr	mai	juin	juil
Emploi																																											
Recherche d'emploi																																											
Formation ou Reprise d'études																																											
Autre situation																																											

Récapitulatif de saisie :

	Début	Fin	Situation	Emploi	
				Entreprise	Commune
+ ✎ 🗑	avr-17	avr-18	Emploi	A	13014 Marseille
+ ✎ 🗑	mai-18	août-18	Emploi	B	75012 Paris
+ ✎ 🗑	sept-18	déc-18	Recherche d'emploi		
+ ✎ 🗑	janv-19	mai-20	Emploi	A	13014 Marseille

Son développement devra favoriser la fluidité de remplissage et, en particulier, devra en faciliter la correction/modification.

Pour être intuitif, le calendrier permettra :

- D'entrer les informations de façon chronologique ou non afin de faciliter la restitution du parcours par l'enquêté (la numérotation des situations ne doit pas dépendre de l'ordre de saisie mais uniquement de la chronologie des événements) ;
- De saisir de l'information de manière optimisée : un clic de souris pour sélectionner la date du début de période, un clic pour la fin de période, ou alors sélection de la période avec un cliquer-glisser de souris ;

- D'effacer des informations saisies pour en saisir de nouvelles de la manière la plus optimale possible (par exemple clic droit sur la période puis supprimer) ;
- De consulter sur le calendrier une période saisie (par exemple, clic droit puis consulter, ou passage de la souris qui génère l'affichage d'une info-bulle) ;
- De faciliter l'intégration d'une période entre deux périodes déjà existantes ;
- De proposer la possibilité d'allonger ou de raccourcir une période déjà saisie sans être obligé de ressaisir l'intégralité de l'information déjà en mémoire ;
- D'effectuer des contrôles de cohérence en direct (par exemple, une seule situation décrite pour un mois donné, pas de mois non renseigné, ni de chevauchement) ;
- De visualiser un récapitulatif modifiable au fur et à mesure de la saisie du calendrier. Ce récapitulatif pourra être modifié sur son contenu, de la même manière que le calendrier : modification du nom de l'entreprise, de la commune de l'entreprise, des dates de début et de fin de période ;
- D'utiliser un système de couleurs pour repérer les différentes situations.

Ce calendrier est complété par des pop-up dynamiques de description des situations. Un pop-up s'ouvre chaque fois qu'une situation est saisie. Le détail sera fourni dans le questionnaire. Ci-dessous, toutefois, quelques exemples d'informations collectées au fur et à mesure du remplissage du calendrier :

- Type de situation mois par mois ;
- Date de début et date de fin de chaque situation ;
- Pour une situation d'emploi : nom et commune de l'établissement employeur. Il s'agira, pour collecter cette information, d'intégrer une application externe de recherche de l'entreprise (application non demandée dans ce marché). Un travail collaboratif est prévu avec le prestataire de collecte des communes (partie 8.4).
- Pour une situation d'emploi en intérim. En fonction de critères de sélection qui dépendent de la position de la situation dans le calendrier ou de sa durée, il sera demandé à l'enquêté de caractériser soit la mission d'intérim la plus longue, soit la première ou la dernière ;
- Pour une situation de recherche d'emploi, une série de questions à choix unique.
- Pour une situation de formation-reprise d'études, une question permettra de déterminer le type de formation. Dans certains cas, il faudra demander le nom et la commune de l'entreprise (basculent vers l'emploi) ;
- Pour une séquence « autre situation », des basculements vers l'emploi seront possibles en fonction de la nature de l'inactivité ;

Chacun des pop-ups devra proposer deux boutons de fonctionnalité pour valider la saisie des informations et pour supprimer la situation. L'affichage devra être dynamique et occuper l'espace à l'écran (éviter le *scrolling*). Un affichage conditionnel facilitera la navigation.

Pour aider au remplissage du calendrier, un bouton spécifique devra apparaître sur la page. Ce bouton ou image, intitulé « Besoin d'aide » ou « Aide au remplissage », devra permettre d'accéder à un lexique qui s'affichera dans un pop-up. Ce lexique est un menu en autocomplétion. Un bouton supplémentaire, nommé « Comment remplir le calendrier ? » permettra de lancer une vidéo de quelques secondes, toujours dans le cadre d'un pop-up, explicitant comment remplir le calendrier.

Description des variables synthétiques

Le calendrier d'activité, entièrement renseigné, doit fournir une liste de codes, appelés « CAL », dans l'ordre chronologique du parcours. Pour chacune des situations déclarées, un code doit être affecté en fonction de sa nature, sa durée et le moment où elle apparaît dans la description du parcours (son calcul précis sera détaillé dans le questionnaire). Le tableau ci-après présente l'ensemble des 11 valeurs de la variable « CAL » possibles avec nom des pop-up et modules à ouvrir.

Tableau 4 – Description de la variable de synthèse CAL du calendrier d'activité

SITUATION	CAL	Descriptif de CAL	Temporalité	Durée	Pop up à ouvrir	Module à ouvrir
1	01	Emploi court	passé	≤ 12 mois	EP	
	02	Emploi long	passé	> 12 mois		
	03	Emploi court	actuel	≤ 6 mois	EA	
	04	Emploi long	actuel	> 6 mois		
2	05	Recherche d'emploi	passé	Pas de notion de durée	REP	Pas de module
	06		actuel		REA	
3	07	Formation / Reprise d'études	passé		FREP	
	08		actuel		FREA	
4	09	Autre situation (Inactivité)	passé		AUP	
	10		actuel		AUA	
5	11	Job de vacances	passé	≤ 5 mois (05 à 09/2017)	Pas de module	

D'autres variables de synthèse utilisées dans certains filtres devront être calculées au même moment que la variable CAL (liste non exhaustive) :

SITUATION = 1 emploi/2 recherche d'emploi/3 formation reprise d'études/4 autre situation/5 job de vacances

CALORIG = première situation saisie dans le calendrier avant basculement en emploi (valeur du CAL soit de la situation formation-reprise d'études ou autre situation)

CAL1EMP = premier emploi autre qu'un job de vacances (numéro de la séquence)

CALEMPACTU = en emploi à la date de l'enquête (1=oui, 0=non)

CALEMP = présence d'au moins une situation d'emploi (1= oui, 0=non)

CALRE = présence d'au moins une situation de recherche d'emploi (1= oui, 0=non)

CALFRE = présence d'au moins une situation de formation-reprise d'études (1=oui, 0=non), tenir compte des repêchages (calcul à réaliser à partir du CAL final)

CALAUT = présence d'au moins une autre situation 1=oui, 0=non), tenir compte des repêchages (calcul à réaliser à partir du CAL final)

Chacune des situations à identifier est indépendante de sa temporalité, elle peut avoir été déclarée au passé ou à la date de l'enquête.

Le candidat précisera si le développement du calendrier est intégré au questionnaire multimode ou si un développement spécifique est nécessaire. Dans ce dernier cas, le candidat précisera la nature du développement et le raccordement au questionnaire. Il devra expliciter dans quelle mesure la fluidité de remplissage pourrait être impactée ainsi que le temps de passation du questionnaire.

Par expérience, il est nécessaire de prévoir une procédure de rappel de l'enquête en cas de problème informatique. Pour maximiser les chances de terminer l'enquête, le délai de réinitialisation d'un tel outil doit être

limité à 48 heures. Cette procédure devra également être mise en place pour tout autre problème informatique survenant en cours d'enquête.

Le candidat devra préciser et démontrer dans sa réponse le mode de fonctionnement précis du calendrier, la totalité de ses possibilités, ainsi que les interactions entre le calendrier et le questionnaire. Le Céreq souhaite une démonstration des capacités du candidat à développer ce type de produit (et non un descriptif sans justification technique). Un exemple sera fortement apprécié dans la réponse du candidat. Cet exemple pourra être intégralement détaillé sur papier (mode de fonctionnement, impressions d'écran etc.). Il pourra éventuellement être présenté « en fonctionnement réel » par l'intermédiaire d'une vidéo d'écran ou d'un programme (exécutable sur format libre). Une attention toute particulière sera portée par le Céreq sur cette partie.

8.5.2 Le calendrier du mode de cohabitation

Le calendrier du mode de cohabitation interactif est un calendrier qui retrace mois par mois le mode de cohabitation des individus (chez les parents, en couple, seul ou en colocation).

Il doit être d'apparence identique à celui du calendrier d'activité, mentionné dans la partie 8.5.1, ainsi que son mode de fonctionnement. À noter que la période d'observation est également équivalente. Pour chaque individu interrogé, le démarrage du remplissage du calendrier coïncide avec la date de fin des études déclarée + 1 mois et se poursuit jusqu'à la date d'interrogation en 2020.

Figure 5 – Visuel du calendrier du mode de cohabitation de l'enquête 2020

	2017												2018												2019												2020											
	jan	fev	mars	avril	mai	juin	juil	août	sept	oct	nov	déc	jan	fev	mars	avril	mai	juin	juil	août	sept	oct	nov	déc	jan	fev	mars	avril	mai	juin	juil	août	sept	oct	nov	déc	jan	fev	mars	avril	mai	juin	juil					
Parents																																																
En couple																																																
Seul																																																
En colocation																																																

Pour ce calendrier, seule la description de la situation du mode de cohabitation mois par mois est demandée (aucun pop-up n'est à prévoir). Les situations sont exclusives.

Le candidat devra démontrer sa capacité à développer un calendrier interactif à double entrée :

- Détailler et démontrer le choix technique autour du développement du calendrier ;
- Démontrer l'interactivité du calendrier en détaillant la méthode de remplissage en multimode ainsi que les possibilités de correction (modifications de périodes, suppressions, ajouts, contrôles et messages d'erreur) ;
- Indiquer par des impressions d'écran et par une démonstration vidéo ou par un programme exécutable, le fonctionnement de l'outil proposé ;
- Justifier techniquement et systématiquement les possibilités annoncées (y compris l'ouverture de pop-up dynamiques et le pilotage des modules en fonction du parcours décrit dans le calendrier) ;
- Décrire le développement du tableau récapitulatif de la saisie (avec les possibilités de correction des situations décrites dans les pop-up).

8.6 Le tableau dynamique

8.6.1 Le tableau des mobilités résidentielles

Pour faciliter la collecte de ces informations, il est préconisé de reproduire le tableau récapitulatif du calendrier d'activité (partie 8.5.1).

Pour celui des *mobilités résidentielles*, il s'agira d'importer les informations concernant les communes « connues », celles du fichier d'import si validées, celles corrigées ou celles collectées en cours d'enquête. L'individu devra ensuite compléter le tableau en mentionnant tous les changements de communes survenus depuis la fin de ses études en renseignant la date, le code postal et la commune. Le recueil du code INSEE est aussi indispensable mais n'est pas nécessaire à afficher.

Le même système que celui du calendrier d'activité devra être mis en place, c'est à dire offrir la possibilité d'ajouter, de modifier ou de supprimer une commune. Toutefois, les informations importées ne devront pas être modifiables.

Une recherche par complétion automatique devra être proposée pour chacun des champs. La double recherche code postal ou commune devra être identique à celle proposée dans le pop-up d'emploi. Sera étudiée la possibilité d'intégrer un module de recherche RNVP avec le prestataire du marché associé (module de recherche externe à ce présent marché).

Figure 6 – Visuel du tableau des mobilités résidentielles

			Date de changement		Code postal	Commune	
			Mois	Année			
+					13002	Marseille	<i>Classe de sixième</i>
+					13002	Marseille	
+					13002	Marseille	
+			septembre	2018	69001	Lyon	<i>Année du Bac</i>
+			juin	2019	13002	Marseille	

8.6.2 Le tableau des langues parlées

Le tableau des *langues parlées à la maison* sera du même format que précédemment. Il sera présenté vierge d'informations. Un menu « langue » et un menu « parents » seront fournis par le Céreq. Une recherche par complétion automatique devra être intégrée dans chacun des champs.

Figure 7 – Visuel du tableau des langues parlées

			Langue parlée	Parents
+			Français	Les deux
+			Portugais	Père
+			Anglais	Mère

9. Tests du questionnaire de l'enquête

9.1 Tests manuels et tests automatiques du questionnaire

Une fois le questionnaire développé, le prestataire et le Céreq effectueront des contrôles sur le respect de l'intitulé des questions, des modalités, des filtres et de tout ce qui est relatif à la programmation du questionnaire mais également de l'ergonomie de l'outil (indispensable à une collecte multimode).

Le prestataire mettra également en place un système permettant au Céreq d'effectuer des tests à distance. Il mettra par ailleurs à disposition du Céreq plusieurs terminaux sur son plateau et dans ses locaux dans l'éventualité de tests effectués directement chez le prestataire.

Du fait de la complexité du questionnaire, cette phase de développement et de tests de conformité nécessite de nombreux allers-retours entre le Céreq et prestataire avant d'aboutir à une version stabilisée. Une grande disponibilité du prestataire dans cette phase est donc souhaitée. Le candidat précisera ainsi le nombre, la quotité de travail, l'expérience ainsi que les compétences des personnes dédiées au développement du questionnaire. Il indiquera également l'organisation du travail pendant cette phase. Un exemple d'organisation possible pourrait prendre la forme suivante :

- Les demandes de modifications du Céreq sont adressées par mail de manière groupée deux fois par semaine dans un fichier (Excel) de modifications évolutives qui permet d'assurer la traçabilité des demandes Céreq et des modifications effectuées par le développeur.
- Deux conférences téléphoniques ont lieu par semaine, permettant de clarifier certaines demandes et de trouver d'éventuelles solutions aux problèmes rencontrés. À l'issue de chaque conférence, le prestataire annoncera par mail les délais pour réaliser les modifications décidées.
- Entre 2 et 4 séances de travail auront lieu chez le prestataire entre l'équipe Céreq et le (ou les) personne(s) en charge du développement du questionnaire et du calendrier. Ces séances permettront de réaliser les correctifs et tester les résultats immédiatement. Sur certaines parties délicates du questionnaire, ce fonctionnement permet de gagner en efficacité.

Pour faciliter le travail de développement du questionnaire, le candidat devra préciser dans sa réponse s'il peut mettre à disposition du Céreq des licences flottantes de l'outil de développement. L'accès aux filtres et à la visualisation en direct du questionnaire permettra de limiter les échanges sur les modifications de filtres.

Par ailleurs, le candidat précisera s'il a la possibilité d'effectuer des tests de manière automatique, permettant de contrôler les différents filtres ainsi que les plages de saisie. Autrement dit, s'il est en mesure de mettre en place une solution de test (sous forme de « robot » ou de « simulation ») permettant d'injecter des valeurs aléatoires et de créer des jeux de données exploitables en sortie. Le prestataire et le Céreq analyseront alors les jeux de données générés afin de détecter d'éventuels problèmes et de lister les corrections à apporter. Autant de jeux de données que nécessaires aux vérifications devront être créés afin de valider les différents cas engendrés par le questionnaire.

9.2 Tests réels en plateau d'enquête

L'organisation des tests, en conditions réelles, devra permettre de tester des enquêtes par téléphone et par internet. Deux échantillons distincts seront fournis pour chacun des tests. Le premier échantillon (E1) concernera une collecte multimode et chaque individu présent dans le fichier d'import disposera d'une adresse mail. Le second échantillon (E2) sera spécifiquement dédié à une collecte par téléphone (taille réduite).

Pour chaque test, la période de collecte s'étalera sur 15 jours. Le premier jour du test, un mail avis, dédié au test, avec un lien de connexion au questionnaire sera transmis aux individus de E1 (sans relance, ni réception d'appel pendant toute la semaine). La deuxième semaine sera consacrée aux appels et relances téléphoniques sur les individus de E1 n'ayant pas encore participé et ceux ayant commencé à répondre sans avoir terminé l'enquête. Les individus de E2 seront contactés également cette deuxième semaine pour une participation uniquement par téléphone.

Trois tests seront effectués par le prestataire :

Test 1 = Un premier test sera effectué en juin 2019 visant à collecter 300 questionnaires. L'objectif de ce test est d'alimenter le dossier pour le Comité du label de la statistique publique (septembre 2019), qui attribue le label de qualité statistique et d'intérêt général. Ce premier test sera réalisé sur l'intégralité du questionnaire, y compris les modules d'extensions thématiques. Ce test sera également adossé d'un test qualitatif dans le cadre d'un « Youthproofing » (partie 9.3).

Test 2 = Le deuxième test aura lieu en décembre 2019 et visera 300 questionnaires complets.

Test 3 = Le troisième test, s'il s'avère nécessaire, aura lieu en février 2020 et visera également 300 questionnaires complets.

La formation des enquêteurs et superviseurs pour les tests sera conforme à la procédure généralisée. Il sera apprécié que les superviseurs de l'opération participent en tant qu'enquêteurs à ces phases de test.

L'objectif général de ces tests est :

- De contrôler en conditions réelles le pilotage du questionnaire par le calendrier ainsi que les contrôles de plages et de cohérences ;
- De vérifier le bon déroulement du questionnaire ;
- De faire une estimation de la durée du questionnaire à travers un minutage précis des différents modules de questions ;
- De tester les nouveaux modules de questions et les changements de formulation (*Youth-proofing*) ;
- D'estimer la part des enquêtes réalisées par mode de collecte ;
- De valider le fichier d'enquêtes.

Durant les phases de tests, le prestataire communiquera les fichiers intermédiaires (au 50^{ème} questionnaire réalisé et au 100^{ème}) permettant de corriger, dans la mesure du possible, les problèmes majeurs. Des aménagements/corrections du questionnaire pourront être effectués pendant et à l'issue des phases de tests. Le candidat précisera si les modifications réalisées en cours de plateau sont déployées instantanément (*i.e.* dès le début de la première enquête qui suit la modification) sans interrompre la production en cours. L'ensemble des questionnaires sera communiqué en fin de test au Céreq pour analyse.

Le troisième test pourra être annulé si, à l'issue du deuxième test, le questionnaire est considéré stabilisé par le Céreq. Les questionnaires des tests seront facturés comme ceux de l'enquête réelle (avec la DMC finale de l'enquête).

Lors de ces tests, l'équipe Céreq Marseille et la personne qualité du Céreq seront présentes pour la formation et le suivi du plateau. Des postes d'écoutes sur le plateau du prestataire et à distance devront être mis à disposition.

9.3 Test « Youthproofing » du questionnaire

Le principe est de faire tester par un échantillon d'enquêtés, et en présence du Céreq, le questionnaire de l'enquête Génération sur le mode internet et téléphone.

L'objectif général de ce test est de :

- Tester la compréhension des questions en auto-administré et par téléphone ;
- Tester les nouveaux modules de questions et les changements de formulation en auto-administré et au téléphone ;
- Discuter des incompréhensions et de l'ergonomie du site avec les enquêtés en groupe (2 ou 3 groupes en fonction de l'effectif).

Calendrier :

En parallèle du test 1 au mois de juin 2019.

Modalités :

Tester le questionnaire avec des jeunes en présentiel dans les locaux du Céreq ou chez le prestataire. Le test sera réalisé sous sa forme multimode.

Un premier groupe test réalisera le questionnaire sur internet et un autre par téléphone. Les enquêtes par téléphone seront effectuées par les chargés d'études du Céreq ou par les enquêteurs du prestataire. Des chargés d'études seront présents lors du remplissage du questionnaire, que cela soit par téléphone ou internet. Ils réaliseront ensuite un débriefing avec les participants au test.

Recrutement : 20 participants (cible souhaitée)

Les jeunes à contacter seront issus de la base de sondage et résideront à proximité du lieu du test.

Il sera nécessaire que le prestataire ait développé une partie du questionnaire, sur le contact et l'éligibilité, pour recruter les jeunes pour la phase de test. L'objectif est de proposer aux individus dans le champ de l'enquête de participer aux groupes de test du questionnaire (remplissage du questionnaire et échanges avec les concepteurs). Si le principe est accepté, alors un rendez-vous sera proposé dans les semaines qui suivent.

Synthèse des opérations demandées au prestataire :

- Développement d'un questionnaire de « recrutement » pour le test ;
- Passation du questionnaire de « recrutement » par des enquêteurs du prestataire courant mai 2019 sur un échantillon fourni par le Céreq ;
- Prise de rendez-vous et transmission de la liste des participants au Céreq ;
- Organisation des tests internet et téléphone dans les locaux du prestataire (*le cas échéant*) ;
- Mise à disposition dans les locaux du prestataire d'une salle et d'ordinateurs pour le test (*le cas échéant*).

Le candidat devra indiquer les possibilités de test « Youthproofing » dans ses locaux ainsi qu'une proposition d'organisation pour réaliser ce test dans les conditions satisfaisantes pour sa réussite (quel que soit le lieu final de réalisation). Le candidat pourra également être force de proposition à partir de son expérience de ce type de test.

Le candidat devra démontrer sa capacité technique et organisationnelle pour cette phase de développement et de tests :

- Procédure utilisée pour réaliser des tests à distance ;
- Possibilité de faire des tests automatisés ;
- Organisation interne du développement et relations avec le Céreq lors de cette phase ;
- Dimensionnement du plateau pour les tests (nombre d'enquêteurs, de superviseurs et nombre de jours de plateau) ;
- Capacité et réactivité à faire des modifications sur le questionnaire en cours de plateau (pendant les tests) ;
- Organisation du test « Youthproofing ».

10. Composition des équipes, présences plateaux et formations

10.1 Responsable du projet et organisation

Le candidat devra détailler l'organisation générale du projet (à l'aide d'un retro planning précis) et le positionnement de ses équipes dédiées à l'opération Céreq. Il devra mentionner les possibilités de contacts directs, indirects, les quotités de travail dédié à l'opération, les plages horaires et les jours de contact de référence ainsi que le rôle de chacun des membres de l'équipe dans l'opération « Enquête Génération du Céreq ». Il devra préciser leur qualification professionnelle et leur expérience dans les domaines d'activité des enquêtes.

Il devra détailler les équipes dédiées aux opérations suivantes :

- Encadrement global de l'opération, contact principal ;
- Équipe de développement des outils informatiques (développement du calendrier, envoi de mail en masse,...) ;
- Équipe de développement autour du questionnaire d'enquête ;
- Maintenance informatique, maintenance des outils, responsable sécurité ;
- Sécurité des données individuelles dans le cadre strict du règlement général sur la protection des données (prévoir un contact direct entre les DPO des deux parties) ;
- Responsable plateau et suivi de la production.

10.2 Équipes et plateau

Le candidat devra préciser dans son offre le nombre d'enquêteurs et de superviseurs ainsi que leur qualification professionnelle pour le plateau d'enquête téléphonique. Il devra également préciser les jours et les plages horaires sur lesquels les enquêteurs seront positionnés.

Les appels devront être passés du lundi au samedi. Afin d'avoir un maximum de chances de contacter les individus, les équipes devront se répartir sur des créneaux matin, après-midi et soirée. Le nombre d'enquêteurs sur chaque créneau horaire, sur chaque jour de la semaine, devra être précisé dans la réponse. Par expérience, les soirées et le samedi sont des créneaux à privilégier.

Le nombre d'enquêteurs dédiés à l'opération Céreq peut évoluer en cours de plateau (sur la durée des 3,5 mois). Dans ce cas, le candidat précisera exactement l'évolution de la dimension du plateau sur toute la durée prévue. Le nombre total d'heures de production devra être précisé. En cours de plateau, le prestataire indiquera au Céreq de manière quotidienne, dans les statistiques de suivi, le nombre d'enquêteurs présents et le nombre total d'heures de production.

Il est important que le taux de renouvellement des équipes soit limité afin d'assurer une qualité de collecte de l'information constante sur l'ensemble de l'opération. Le candidat précisera le nombre d'enquêteurs travaillant de manière exclusive sur l'opération du Céreq, en distinguant le nombre d'enquêteurs permanents travaillant pour le candidat. Il précisera également le taux de renouvellement moyen observé sur des plateaux de cette taille réalisés sur cette même période. En cours de plateau, le prestataire indiquera au Céreq de manière hebdomadaire, le taux de renouvellement de ses enquêteurs. Il mettra en place une politique de ressources humaines (RH) permettant de garantir un taux de renouvellement inférieur à 25 %. Dans sa réponse, le candidat pourra préciser sa politique RH ainsi que les éventuelles démarches qualité mises en place.

Le plateau d'enquête pourra être découpé en trois terrains en fonction du niveau de qualification des jeunes interrogés : non diplômés et bas niveaux, sortants du secondaire, sortants du supérieur. Les enquêteurs devront être spécialisés par type de terrain pour optimiser la passation du questionnaire en fonction des différents publics. Ce point pourra être rediscuté avec le prestataire.

10.3 Formation des superviseurs et des enquêteurs

Le prestataire assurera la formation des enquêteurs à la manipulation de son outil informatique, c'est-à-dire au déroulement du questionnaire Céreq et particulièrement :

- La manipulation du téléphone : appel, mise en attente, bis, appels masqués/démasqué ;
- Le classement d'appels ;
- La maîtrise des outils informatiques mis à disposition : phase de contact, questionnaire, calendrier.

Il organisera ensuite, dans ses locaux, des réunions de formation à l'attention de l'ensemble des enquêteurs et superviseurs dédiés à l'opération, en deux étapes :

- Le Céreq présentera de manière précise l'opération aux superviseurs ;
- En étroite collaboration avec les chefs d'équipe responsables de la mise en place du terrain, le Céreq présentera ensuite aux enquêteurs par groupe de 10 à 15 personnes sur environ ½ journée, les objectifs de l'enquête, les questions cruciales et celles pouvant poser problème (calendrier, nomenclatures de professions, d'activité et de diplômes) au travers d'exercices et de simulations de questionnaires.

Afin de réaliser au mieux ces formations, il est indispensable :

- D'avoir une salle à disposition équipée de postes de travail connectés directement au questionnaire de test ;
- D'avoir autant de postes de travail que d'enquêteurs à former ;
- D'avoir au préalable chargé des fiches d'individus fictifs identiques pour tous les postes de travail.

Le Céreq assistera au début de l'opération. Le terrain pourra alors être écourté pour faire un briefing ou un débriefing des enquêtes qui permettra d'insister sur certaines consignes de passation des enquêtes.

10.4 Suivi du plateau, responsable et personne « qualité »

Le plateau sera suivi quotidiennement par le responsable Céreq de la production des enquêtes Génération (présent sur Marseille) et par une personne « qualité » (présente le long de l'opération chez le prestataire).

Le responsable Céreq s'entretiendra hebdomadairement par téléphone avec le prestataire pour faire un point sur l'avancement de la collecte et gérer l'organisation générale de la production. Le prestataire établira un relevé de décision de l'entretien adressé par mail à l'ensemble de l'équipe en charge des enquêtes Génération du Céreq.

Une procédure d'écoutes à distance devra permettre au personnel Céreq (chargés d'études, chargés d'enquêtes et direction) de faire des écoutes depuis Marseille ou tout autre site. Cette écoute à distance devra permettre d'écouter la conversation entre l'enquêté et l'enquêteur ainsi que de suivre la saisie à l'écran de l'enquêteur (applications externes comprises).

La personne « qualité » sera spécialement dédiée au suivi de l'opération et sera présente de manière permanente sur le lieu de réalisation des enquêtes (présence possible de deux personnes sur le plateau d'enquêtes). Elle sera l'interlocutrice du prestataire pour tous les aspects concernant la collecte de l'information par les enquêteurs ou par internet.

Le prestataire mettra à sa disposition un poste de travail à proximité des enquêteurs. Ce poste sera équipé du matériel nécessaire au suivi en direct et à la supervision de la passation des enquêtes (écoute des enquêteurs et visionnage de l'enquête en direct, suivi statistique en temps réel, etc.) ainsi que d'un accès à l'interface d'enquête.

Un accès internet sera également mis à disposition de la personne qualité, via son poste de travail ou via une connexion sans fil. Cette connexion devra s'effectuer sur le plateau d'enquête pour assurer la continuité du travail dans les différentes tâches à opérer. La personne qualité sera associée aux équipes du prestataire lors des formations et briefings dispensés aux enquêteurs et aux équipes encadrantes. Elle sera en contact continu avec les équipes encadrantes ainsi qu'avec les enquêteurs et participera à l'élaboration et à la diffusion des consignes et recommandations relatives à la passation du questionnaire.

Quotidiennement, elle recevra l'ensemble des documents nécessaires au suivi de la production : statistiques de production, analyse des résultats des appels, reporting des difficultés rencontrées, gestion des effectifs des enquêteurs, statistiques de remplissage des enquêtes par internet, etc... Elle sera informée des décisions affectant la réalisation des enquêtes comme par exemple l'augmentation ou la diminution du nombre d'enquêteurs dédiés à la réalisation d'enquêtes. Le cas échéant, elle pourra être amenée à émettre des recommandations visant à améliorer la qualité ou la quantité des enquêtes réalisées.

Le candidat devra démontrer sa capacité organisationnelle pour la réalisation générale de l'opération en détaillant les éléments suivants :

- Plages d'appels et jours dédiés à l'opération ;
- Dimensionnement du plateau selon les plages et les jours (nombre d'enquêteurs et superviseurs, taux de renouvellement, nombre total d'heures de production) ;
- Moyens mis à disposition pour former les équipes (postes informatiques, salles de réunion) ;
- Moyens mis à disposition des personnes qualité et du personnel Céreq de Marseille pour le suivi de l'opération ;
- Processus d'ajustement des effectifs d'enquêteurs en cas de sous-production (notamment en précisant les délais opérationnels).

11. Transmission, suivi et bilans de la phase enquête téléphonique

Afin de détecter rapidement d'éventuelles incohérences au cours des premiers entretiens, le prestataire fournira au Céreq le fichier complet des 500 premiers questionnaires réalisés, accompagné du programme (script) du questionnaire (sous format Word par exemple) et de l'ensemble des fichiers historiques de suivi (partie 6). Jusqu'à la fin de l'opération, le prestataire fournira d'autres fichiers de ce type sur demande du Céreq.

Chaque jour, le prestataire fera parvenir au Céreq des statistiques détaillées sur l'avancée de la production :

- Suivi des enquêtes réalisées avec un détail par sous-population, selon une liste de variables issues du fichier d'import définie par le Céreq (exemple : terrain, destinataire des données, plus haut diplôme, etc.) mais également de variables issues des questionnaires complétés ;
- Suivi des adresses (adresses appelées, non appelées) ;
- Classement et état des fiches par rapport à la grille de qualification des appels définie conjointement par le Céreq et le prestataire ;
- Des statistiques permettant de suivre l'avancement du protocole, notamment concernant les 20 tentatives d'appel par numéro disponible (échanges avec le prestataire pour définir d'un ou plusieurs indicateurs) ;
- Durée moyenne de passation par terrain et au niveau global ;
- Suivi par enquêteur anonymisé (nombre d'enquêtes réalisées et durée moyenne) ;
- Éléments nécessaires à la facturation de l'opération (afin de produire des estimations budgétaires en cours de plateau).

Ces statistiques seront détaillées par jour de production et en cumul.

Un point hebdomadaire avec l'équipe Céreq (sur site ou par téléphone/visioconférence) sera mis en place pendant la période de plateau, avec compte rendu à la charge du prestataire, lequel devra être validé par le Céreq. Ce point pourra être très court si tout se passe bien ; dans le cas contraire, il permettra de noter les difficultés et les décisions prises au fur et à mesure.

En fin de collecte, des statistiques complémentaires pourront être demandées au prestataire notamment pour établir le bilan de la collecte.

En milieu d'opération, le prestataire fera parvenir au Céreq le fichier complet des enquêtes réalisées et les fichiers historiques afin que le Céreq puisse mettre en place ses programmes de traitement de l'opération.

Les bases de données transmises devront être livrées dans un format défini à l'avance (idéalement format SAS). Afin d'exploiter rapidement ces fichiers :

- L'identifiant Céreq de l'individu devra être présent dans tous les fichiers ;
- Les variables « vides » qui ont servi au calcul ou à la gestion de filtres devront être supprimées ;
- Les données liées aux calendriers (activité et mode de cohabitation) seront dans des fichiers séparés du fichier correspondant au reste du questionnaire. Ces fichiers liés aux calendriers devront être sous la forme suivante : un individu et une situation décrite par ligne. Par exemple, un individu qui décrit 3 situations dans le calendrier d'activité apparaîtra dans le fichier correspondant 3 fois (une ligne pour chaque situation décrite).

Dans tous les cas, le fichier de données sera accompagné d'un fichier (ou deux) contenant le dessin précis du fichier (noms des variables, caractère ou numérique, longueur de la chaîne - fixe ou variable -, libellé de la question, code et labels des modalités de la variable, etc.), ainsi que de la documentation (dictionnaire des

variables) nécessaire à leur exploitation (le programme papier du questionnaire sous format Word par exemple). Le type de séparateur retenu devra être précisé (et totalement interdit pendant la saisie).

En fin d'opération, le fichier final sera transmis au Céreq selon les mêmes caractéristiques que le fichier intermédiaire, sous un délai d'environ une semaine après l'arrêt complet de la phase terrain. Ce fichier comprendra, pour l'ensemble des individus initialement transmis par le Céreq (échantillon), les mêmes informations que le fichier intermédiaire.

Il sera assorti :

- Des statistiques, par sous-population et au niveau global, sur les différentes phases de l'opération, c'est-à-dire les résultats de l'ensemble des contacts téléphoniques effectués (selon la grille de qualification des appels) ;
- Des diverses remarques des enquêteurs.

Il est impératif que la transmission de tout fichier comprenant des données individuelles, directement ou indirectement nominatives, soit préalablement crypté par le prestataire selon une procédure validée par le Céreq et soit accessible par transmission des données sécurisée. Cette procédure concerne la transmission des fichiers intermédiaires et finaux, des fichiers historiques et des fichiers d'enquêtes (complets ou non).

Le prestataire produira un bilan détaillé de l'opération de collecte et le présentera à l'équipe du Céreq lors d'une réunion de bilan qui clôturera l'opération.

À l'issue du plateau, le questionnaire devra être accessible dans son intégralité (en ligne) pour le Céreq pendant un an (soit jusqu'en juillet 2021), l'objectif principal de cette mise à disposition étant de faciliter le travail de création du fichier de diffusion. Également, pour disposer des éléments mettant en évidence les choix réalisés pendant l'enquête, en termes d'ergonomie de questionnaire, à intégrer lors de la rédaction du bilan scientifique et méthodologique.

Le candidat pourra proposer d'autres formes d'accessibilité au questionnaire multimode.

Le prestataire fournira au Céreq :

- Le questionnaire multimode développé (questionnaire + calendrier) sous format texte intégrant la totalité des filtres et des formules (Word par exemple) ;
- Les fichiers des statistiques de suivi quotidien de la production ;
- L'ensemble des bases de données collectées (questionnaires complets ou non, données de la phase de contact, historiques) en versions intermédiaires et en version finale ;
- Un fichier de type tableur (et/ou fichier sas) comprenant le libellé de la question, le nom de la variable et le type de variable correspondante ;
- Un fichier de type tableur (et/ou fichier sas) contenant les modalités pour chaque variable présente dans le questionnaire ;
- Un bilan détaillé de l'opération.

12. Préconisations pour l'application du Règlement général de protection des données (RGPD)

Applicable depuis le 25 mai 2018 à l'ensemble de l'Union européenne, le règlement européen sur la protection des données (RGPD) renforce les droits des résidents européens sur leurs données et responsabilise l'ensemble des acteurs traitant ces données (responsables de traitement et sous-traitants) qu'ils soient ou non établis au sein de l'Union européenne.

Le sous-traitant indiquera dans sa réponse si, au sein de son établissement, un DPO (Data protection officer) a été désigné et, si tel n'était pas le cas, s'il prévoit d'en nommer un. Le règlement impose des obligations spécifiques aux sous-traitants dont la responsabilité est susceptible d'être engagée en cas de manquement.

Lorsque le sous-traitant interviendra dans la mise en œuvre des traitements de données personnelles, il devra offrir au Céreq « **des garanties suffisantes quant à la mise en œuvre de mesures techniques et organisationnelles appropriées de manière à ce que le traitement réponde aux exigences du présent règlement et garantisse la protection des droits de la personne concernée** » (article 28 du règlement européen).

Il devra notamment assister et conseiller le Céreq dans sa conformité à certaines obligations prévues par le règlement européen (analyses d'impact, notification de violation, sécurité, destruction des données).

Concrètement, cela impliquera :

Une obligation de transparence et de traçabilité

Le contractant devra :

- Mettre à la disposition du Céreq toutes les informations nécessaires pour démontrer le respect des obligations ;
- Tenir un registre qui recense les traitements effectués comprenant :
 - ✓ Le nom et les coordonnées du ou des sous-traitants et de chaque responsable du traitement pour le compte duquel le sous-traitant agit ainsi que, le cas échéant, les noms et les coordonnées du représentant du responsable du traitement ou du sous-traitant et celles du délégué à la protection des données ;
 - ✓ Les catégories de traitements effectués pour le compte de chaque responsable du traitement ;
 - ✓ Le cas échéant, les transferts de données à caractère personnel vers un pays tiers ou à une organisation internationale, y compris l'identification de ce pays tiers ou de cette organisation internationale et, dans le cas des transferts visés à l'article 49, paragraphe 1, deuxième alinéa, les documents attestant de l'existence de garanties appropriées ;
 - ✓ Dans la mesure du possible, une description générale des mesures de sécurité techniques et organisationnelles visées à l'article 32, paragraphe 1.

La prise en compte des principes de protection des données dès la conception et de protection des données par défaut

Le contractant devra offrir les garanties nécessaires afin que le traitement mis œuvre pour le compte du Céreq réponde aux exigences du règlement européen et protège les droits des personnes concernées.

Cela signifie notamment que :

- dès leur conception, les outils, produits, applications ou services intègrent de façon effective les principes relatifs à la protection des données ;
- par défaut, les outils, produits, applications ou services garantissent que seules sont traitées les données nécessaires à la finalité du traitement au regard de la quantité de données collectées, de l'étendue de leur traitement, de la durée de conservation et du nombre de personnes qui y a accès ;
- de purger automatiquement et sélectivement les données d'une base active à l'issue d'une certaine durée ou de gérer des habilitations et droits d'accès informatiques « donnée par donnée » ou sur demande des personnes concernées.

Une obligation de garantir la sécurité des données traitées

- Les employés du contractant qui traitent les données du Céreq doivent être soumis à une obligation de confidentialité ;
- Le contractant devra notifier au Céreq toute violation de ses données ;
- Le contractant devra prendre toute mesure pour garantir un niveau de sécurité adapté aux risques ;
- Au terme de la prestation et selon les instructions du Céreq, le contractant devra :
 - ✓ renvoyer toutes les données au Céreq puis les supprimer ;
 - ✓ détruire les copies existantes, sauf obligation légale de les conserver.

Une obligation d'assistance, d'alerte et de conseil

Sur chaque opération, le Céreq attend des conseils (techniques, mise en conformité) des différentes opérations. Le service informatique du Céreq sera sollicité pour analyser, si besoin techniquement, la sécurité de certaines opérations.

- Si, selon le contractant, une instruction du Céreq constitue une violation des règles en matière de protection des données, le contractant devra l'en informer immédiatement ;
- Lorsqu'une personne exerce ses droits (accès, rectification, effacement, portabilité, opposition, ne pas faire l'objet d'une décision individuelle automatisée, y compris le profilage) le contractant devra, dans toute la mesure du possible, aider le Céreq à donner suite à cette demande ;
- Compte tenu des informations mises à disposition, le contractant aidera le Céreq à garantir le respect des obligations en matière de sécurité du traitement, de notification de violation de données et d'analyse d'impact relative à la protection des données.

Le candidat devra fournir des éléments de réponse aux questions suivantes :

- Présence d'un DPO dans l'établissement/entreprise ?
- Le profil du DPO est-il technique/juridique/les deux ?
- Pouvez-vous détailler votre procédure interne (et votre position en tant que sous-traitant) à suivre en cas de violation de données ?
- Les employés de l'établissement/entreprise sont-ils soumis à une charte informatique ?
- Quelles sont les mesures techniques mises en œuvre pour garantir la protection des données personnelles ?
 - Chiffrement des données
 - Gestion des droits d'accès
 - Lutte contre les intrusions
 - Politique de mot de passe postes, sécurisation des postes
 - Flux et liens sécurisés
 - Sécurisation des pc portables, téléphone mobiles (chiffrement en cas de vol, etc.)
 - Sécurisation des serveurs
 - Sécurisation des sites web dédiés à l'opération
 - Sécurisation des échanges sur la transmission de fichiers
- Quelles sont les mesures organisationnelles mises en œuvre pour garantir la protection des données personnelles ?
 - Sensibilisation/formation des équipes métier et IT
 - Privacy by design
 - Analyse de risque
 - Gestion des droits des agents sur l'accès aux serveurs de fichiers, etc.
- Quelles sont les mesures physiques (protection des locaux) mises en œuvre pour garantir la sécurité des données ? (Salles sécurisées, traçage des accès etc.)
- À quelle fréquence sont sauvegardées les données ?
- Quel est le recul de sauvegarde des données ?
- Quel est le lieu d'hébergement des données ? (Chez-vous, chez un sous-traitant ? En France, en Europe, à l'étranger ?)
- Les outils informatiques intègrent-ils de façon effective les principes relatifs à la protection des données ? (Précisez, par outil, les règles appliquées)
- Avez-vous mis en œuvre un traçage des accès aux données personnelles ?

13. Calendrier prévisionnel des opérations

Date	Opération
16 janvier 2019	Lancement du marché
13 février 2019	Date limite de remise des offres
25-27 février 2019	Audition des candidats sélectionnés
27 février 2019	Notification au candidat retenu
15 mars 2019	Début d'exécution du marché et fourniture du questionnaire. Le Céreq fournira au prestataire le questionnaire de l'enquête afin qu'il soit développé au format multimode. Certains modules thématiques pourront être adjoints après cette date.
À partir d'avril 2019	Démarrage du travail collaboratif avec l'Insee pour le volet collecte de la profession et avec le prestataire du marché dédié à la collecte de l'activité de l'entreprise et des adresses postales des individus interrogés.
1er avril 2019	Fourniture d'une identité visuelle du questionnaire. Basée sur la charte graphique du Céreq et des indications en termes d'ergonomie.
15 avril 2019	Fourniture d'une version de questionnaire multimode. Le prestataire devra mettre à disposition du Céreq une première version du questionnaire développé, un mois après la remise du questionnaire.
À partir du 15 avril 2019	Tests et échanges sur le développement du questionnaire. Le questionnaire fera l'objet de contrôles manuels par le Céreq entre le 15 avril 2019 et 31 mai 2019 - soit 1,5 mois - afin qu'à ce terme, le questionnaire soit dans une version stabilisée pour démarrer les tests en réel. Test, également, de l'intégration des applications externes profession, activité et adresse postale. D'autres types de tests devront être effectués par le prestataire en parallèle au fur et à mesure des modifications réalisées avec une mise à disposition des jeux de données aléatoires. Durant et à l'issue des phases de test, des modifications pourront être apportées par le Céreq.
10-22 Juin 2019	Test 1 en réel (pour la constitution du dossier du label). Test multimode du questionnaire pour obtenir 300 répondants par téléphone et par internet. Ce test intègre également un test qualitatif avec une cellule d'écoutes (Youthproofing). Le Céreq analysera les résultats et demandera éventuellement à modifier la programmation.
2-14 décembre 2019	Test 2 en réel. Test multimode du questionnaire pour obtenir 300 répondants par téléphone et par internet (minimum 250). Le Céreq analysera les résultats et demandera éventuellement à modifier la programmation.
3-15 février 2020	Test 3 en réel. Test multimode du questionnaire pour obtenir 300 répondants par téléphone et par internet (minimum 250). Le Céreq analysera les résultats et demandera éventuellement à modifier la programmation.
2 mars 2020	Validation définitive du questionnaire. Développement informatique du questionnaire et visuel de l'outil de collecte.
Courant mars 2020	Le prestataire réalisera la commande du numéro gratuit.
Fin mars 2020	Formations des enquêteurs à l'outil (interne au prestataire)
30-31 mars 2020	Envoi des lettres avis par voie postale (Céreq)
31 mars 2020	Ouverture du numéro gratuit (prestataire)
1-3 avril 2020	Envoi des mails avis (prestataire)
1er avril 2020	Formation des enquêteurs (par le Céreq)
2 avril 2020	Démarrage de la collecte. Ouverture du plateau d'enquêtes pour une durée de 3,5 mois.
9 avril 2020	Phase 1 – Collecte internet (mode de collecte principal)
4 mai 2020	Phase 2 – Collecte multimode (proposition des deux modes de collecte)
8 juin 2020	Phase 3 – Collecte téléphone (mode de collecte principal)
15 juillet 2020	Fin de la collecte*
22 juillet 2020	Réception des fichiers finaux
7 septembre 2020	Réunion de bilan

* Ce calendrier est susceptible d'être modifié. En particulier, si le taux de réponse minimum acceptable n'est pas atteint. Une prolongation du plateau d'enquête d'une semaine au maximum serait envisagée. Cette situation n'est évidemment pas souhaitée.

14. Annexes

Annexe 1 : Extrait du questionnaire

MODULE : ACQUISITION DE CERTIFICATIONS EN POST-INITIAL– CALLED « DIP »
ASK ALL

FIDIP – title (safeguard)

Acquisition de certifications en post-initial

TDDIP010 - safeguard

DIP010A

ASK IF FREP030=2 AND FREA020≠1 AND...ENSEMBLE DES VARIABLES DE DIPLOMES DU CALENDRIER (EMPLOI)

Depuis l'obtention du <DIPSOR> <spécialité> en <mois d'obtention> <année d'obtention>, avez-vous obtenu un nouveau diplôme ?

Consigne : Décrire le dernier diplôme obtenu si 2 diplômes obtenus de niveau équivalent

Oui un seul..... 1
Oui, plusieurs..... 2
Non 3

FIDIP010B - recode

DETERMINATION DU DIPLÔME EN POST-INITIAL LE PLUS ELEVE

DO i=1 TO 7

IF FREP050 NOT EMPTY AND NIV=i THEN FIDIP010B=FIFREP051

ELSE IF FREA050 NOT EMPTY AND NIV=i THEN FIDIP010B=FIFREA051

ELSE IF (DIPLÔME EMPLOI) NOT EMPTY AND NIV=i THEN FIDIP010B= DIPLÔME EMPLOI

END

DIP010B

ASK IF DIP010A=EMPTY

Vous nous avez déclaré avoir obtenu le diplôme suivant : <FIDIP010B>. Est-ce le diplôme le plus élevé que vous ayez obtenu depuis l'obtention du <DIPSOR> <spécialité> en <mois d'obtention> <année d'obtention> ?

Consigne : Décrire le dernier diplôme obtenu si 2 diplômes obtenus de niveau équivalent

Oui 1
Non 2

DIP010C

ASK IF DIP010A=EMPTY AND DIP010B=EMPTY

Après avoir arrêté vos études en <CLASSNEW> en <ANSOR>, avez-vous obtenu un diplôme ?

Consigne : Décrire le dernier diplôme obtenu si 2 diplômes obtenus de niveau équivalent

Oui un seul.....	1
Oui, plusieurs.....	2
Non	3

FIDIP020 – display

IF DIP010A=1 OR DIP010C=1, CODE 1

IF DIP010A=2 OR DIP010C=2, CODE 2

IF DIP010B=2, CODE 3

Quel est ce diplôme.....1

Parmi ces diplômes, quel est le plus élevé.....2

Quel était alors le diplôme le plus élevé que vous ayez obtenu depuis <mois d'obtention diplôme de sortie> <ANSOR>3

DIP020

ASK IF DIP010A=1,2 OR DIP010B=2 OR DIP010C=1,2

<FIDIP020> ?

Consigne : Sélectionnez « autre » si diplôme non trouvé

Menu diplôme

CODE DU DIPLÔME

DIP021 - safeguard

LIBELLE DU DIPLÔME

DIP030

ASK IF DIP010A=1,2 OR DIP010B=2 OR DIP010C=1,2

Quelle est la spécialité de ce diplôme ?

Consigne : Sélectionnez « autre » si spécialité non trouvée

Menu spécialité

CODE DU SPECIALITE

DIP031 - safeguard

LIBELLE DE LA SPECIALITE

Affichage sur une page (au même niveau)
DIP040

ASK IF DIP010A=1,2 OR DIP010B=2 OR DIP010C=1,2

En quelle année l'avez-vous obtenu ?

Menu année

DIP041

ASK IF DIP010A=1,2 OR DIP010B=2 OR DIP010C=1,2

En quel mois de <DIP040> ?

Menu mois

DIP050

ASK IF DIP010A=1,2 OR DIP010B=2 OR DIP010C=1,2

Est-ce dans le cadre d'une validation des acquis de l'expérience (VAE) ?

Oui 1

Non 2

Une infobulle contenant la définition suivante :

VAE = Toute personne, quels que soient son âge, sa nationalité, son statut et son niveau de formation, qui justifie d'au moins 1 an d'expérience en rapport direct avec la certification visée, peut prétendre à la VAE. Cette certification qui peut être un diplôme, un titre ou un certificat de qualification professionnelle doit être inscrite au Répertoire national des certifications professionnelles (RNCP).

DIP060

ASK IF FREP010≠3 AND FREA010≠3

Nous allons parler des autres certifications que vous avez éventuellement obtenues. Après avoir arrêté vos études en <CLASSNEW> en <ANSOR>, avez-vous obtenu un titre professionnel ?

Consigne : Décrire le titre professionnel le plus élevé obtenu si plusieurs

Oui un seul..... 1

Oui, plusieurs..... 2

Non 3

FIDIP070 – display

IF DIP060=1, CODE 1

IF DIP060=2, CODE 2

IF FREP010=3 AND FREP040=1, CODE 3

Quel est ce titre professionnel ? 1

Quel était alors le titre professionnel le plus élevé que vous ayez obtenu depuis <ANSOR> ? 2

Nous allons parler des autres certifications que vous avez obtenu. Vous nous avez déclaré avoir obtenu au moins un titre professionnel. Quel est l'intitulé du titre professionnel le plus élevé que vous ayez obtenu depuis <ANSOR> ? 3

Affichage sur une page (conditionnel)

DIP070

ASK DIP060=1,2 OR (FREP010=3 AND FREP040=1)

<FIDIP070>

Consigne : Sélectionnez « autre » si titre professionnel non trouvé

Menu titre professionnel

CODE DU TITRE PROFESSIONNEL

DIP070MC – text field
ASK IF DIP070=AUTRE (DÉFINIR CODE)

Veillez préciser :

DIP071 - safeguard

LIBELLE DU TITRE PROFESSIONNEL

Affichage sur une page (au même niveau)

DIP080

ASK DIP060=1,2 OR (FREPO10=3 AND FREPO40=1)

En quelle année l'avez-vous obtenu ?

Menu année

DIP081

ASK DIP060=1,2 OR (FREPO10=3 AND FREPO40=1)

En quel mois de <DIP080> ?

Menu mois

DIP090

ASK IF FREPO10≠3 AND FREA010≠3

Après avoir arrêté vos études en <CLASSNEW> en <ANSOR>, avez-vous obtenu un CQP ?

Oui 1

Non 2

Une infobulle contenant la définition :

CQP = certification de qualification professionnelle créée et délivrée par une branche professionnelle

DIP100

ASK IF FREPO10≠4 AND FREA010≠4

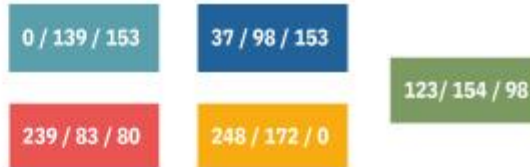
Après avoir arrêté vos études en <CLASSNEW> en <ANSOR>, avez-vous obtenu une habilitation ?

Oui 1

Non 2

TFDIP100 - safeguard

Annexe 2 : Éléments de charte graphique



Police 10 p / IBM Plex sans condensed / Light
 Police 10 p / IBM Plex sans condensed / Light Ital

Police 10 p / IBM Plex sans condensed / regular
 Police 10 p / IBM Plex sans condensed / regular

Police IBM 11 p / Plex sans condensed / medium
 Police IBM 11 p / Plex sans condensed / medium ital

Police IBM 12 p / Plex sans condensed / medium
 Police IBM 12 p / Plex sans condensed / medium ital

Police 13 p / IBM Plex sans condensed / bold
Police 13 p / IBM Plex sans condensed / bold

Police 13 p / IBM Plex sans condensed / semi bold
Police 13 p / IBM Plex sans condensed / semi bold

Police CAVAR DREAM
 Police CAVAR DREAM
 Police CAVAR DREAM
 Police CAVAR DREAM



Génération



Annexe 3 : Quelques données de cadrage sur les appels téléphoniques

Ces quelques statistiques ont été établies à partir des précédentes enquêtes générations. Ces dernières étant monomode (téléphone), le comportement des enquêtés lors de cette future vague pourrait évoluer et le volume d'appels diminuer en faveur des réponses internet.

Néanmoins ces chiffres permettront au contractant.e de remplir au mieux l'annexe financière qui suit. En effet il est rappelé que seules les enquêtes complètes sont facturées au Céreq. Les appels infructueux, le travail de recherche du bon individu (exemple : contact avec un tiers), les questionnaires incomplets, ... ne seront pas facturés. Le candidat en tiendra compte (en répercutant ces coûts) dans la grille de prix proposée.

Taille de l'échantillon attendu : entre 200 000 et 300 000

Volumes d'appels lors de l'enquête 2013 auprès de la Génération 2010 :

Taille de l'échantillon	279 484
Échantillon effectivement appelé	254 144
Injoignables	105 736
Nombre d'appels aux injoignables	236 532

15. Annexe financière

Le la contractant.e indique précisément, dans l'annexe financière, le prix qu'il.elle applique pour les prestations décrites ci-après. Le prix est unitaire et exprimé en **euros H.T.**
Toutes les prestations proposées doivent être inscrites dans l'annexe et comprises dans le montant total de l'offre.

		Situation minimum : questionnement de <u>25 minutes</u> avec un nombre total d'enquêtes complètes de <u>21 000</u>			Situation cible : questionnement de <u>30 minutes</u> avec un nombre total d'enquêtes complètes de <u>31 000</u>			Situation maximum : questionnement de <u>35 minutes</u> avec un nombre total d'enquêtes complètes de <u>46 000</u>			
		Unité de compte	Prix unitaire H.T. minimum	Quantité minimum	Total H.T. minimum	Prix unitaire H.T. cible	Quantité cible	Total H.T. cible	Prix unitaire H.T. maximum	Quantité maximum	Total H.T. Maximum
Prise en charge des fichiers		Forfait		1			1			1	
Ouverture du numéro gratuit		Forfait		1			1			1	
Ouverture de numéros régionalisés et redirection sur le numéro gratuit		Forfait		1			1			1	
Envoi de mails (avis, relances)		Millier		75 ¹			300 ¹			600 ¹	
Envoi de sms		Millier		10 ¹			30 ¹			50 ¹	
Développement	Questionnaire d'enquête format multimode	Forfait		1			1			1	
	Calendriers multimodes (activité, habitat)	Forfait		1			1			1	
Facturation des enquêtes complètes réalisées selon la durée moyenne cumulée finale (DMC) par téléphone	DMC = 25 min	Enquêtes complètes réalisées		10 500			15 500			23 000	
	DMC = 26 min			10 500			15 500			23 000	
	DMC = 27 min			10 500			15 500			23 000	
	DMC = 28 min			10 500			15 500			23 000	
	DMC = 29 min			10 500			15 500			23 000	
	DMC = 30 min			10 500			15 500			23 000	
	DMC = 31 min			10 500			15 500			23 000	
	DMC = 32 min			10 500			15 500			23 000	
	DMC = 33 min			10 500			15 500			23 000	
	DMC = 34 min			10 500			15 500			23 000	
DMC = 35 min		10 500			15 500			23 000			

Réalisation intégrale de l'enquête	Par internet	Enquêtes complètes réalisées		10 500			15 500			23 000	
Hébergement de l'enquête CAWI (à remplir selon la facturation logicielle)	(ou) Facturation sur connexions simultanées	À préciser par le contractant									
	(ou) Facturation à l'enquête entamée	Unité									
Mise à disposition du CATI et du CAWI pour contrôle par le Céreq		À préciser par le contractant									
Relance téléphonique avec passage du questionnaire filtre, ayant pour objectif une réponse à terme par internet (qualification)		Individu qualifié									
Relance téléphonique pour terminer l'enquête commencée par internet (coût unitaire à la minute)		(Enquête complète réalisée) x (temps moyen passé pour terminer l'enquête)		1 500			2 500			3 000	
Formation des enquêteurs.trices		Enquêteurs.trices formé.e.s		60 ²			100 ²			140 ²	
Formation des superviseurs.res		Superviseur.e.s formé.e.s		3 ²			6 ²			9 ²	
Transmission des fichiers et du bilan		Forfait		1			1			1	
TOTAL											

¹ Ces volumes sont fictifs, la base de sondage étant toujours en phase de constitution ; Ils serviront de base à la comparaison des offres

² Ces volumes sont donnés à titre d'exemple pour comparer les offres. Le contractant précisera dans son offre le volume d'enquêteurs et de superviseurs à former.

Le candidat doit remplir toutes les cellules vides bleues du tableau.

Tous les éléments présents dans cette annexe financière ne font pas forcément l'objet d'une facturation. Si c'est le cas, le/la contractant.e précise sur la ligne correspondante la mention « offert » ou « sans objet ».

Cas de minute supplémentaire ou de déduction :

	Coût en euros (H.T.)
Minute supplémentaire d'enquête (au-delà de 35 minutes)	
Déduction d'une minute d'enquête (en-deçà de 25 minutes)	

Le/la contractant.e indique précisément le prix qu'il applique pour les prestations décrites ci-dessus. Le prix est unitaire et exprimé en **euros H.T.**
Les prix s'entendent **fermes** et **définitifs** pendant la durée du marché.

Date :

Cachet et signature du/de la candidat.e :



Notice de l'annexe financière

Toutes les cases vides bleues doivent être renseignées par le candidat.

Calcul des prix pour les situations intermédiaires (en termes de quantités)

L'annexe financière permet de proposer des prix différents selon que l'on se trouve dans la situation minimum (PUmin), cible (PUcible) ou maximum (PUmax). À la fin de l'enquête, une fois les quantités finales connues, les prix finaux seront calculés par interpolation linéaire comme indiqué ci-dessous.

Pour une ligne de l'annexe financière donnée, les quantités minimum (Qmin), cible (Qcible) et maximum (Qmax) sont connues.

1^{er} cas : si $Q_{min} < Q_{cible} < Q_{max}$

Si la quantité finale (Qfin) est inférieure à Qcible, alors le Prix unitaire finale (PUfin) est déterminé par la formule :

$$PU_{fin} = \frac{PU_{cible} - PU_{min}}{Q_{cible} - Q_{min}} \cdot (Q_{fin} - Q_{min}) + PU_{min}$$

Si la quantité finale (Qfin) est supérieur à Qcible, alors le prix unitaire finale (PUfin) est déterminé par la formule :

$$PU_{fin} = \frac{PU_{max} - PU_{cible}}{Q_{max} - Q_{cible}} \cdot (Q_{fin} - Q_{cible}) + PU_{cible}$$

2^e cas : si $Q_{min} = Q_{cible} = Q_{max}$ (et que les prix unitaires sont différenciés selon la situation min, cible, max), on utilise le nombre d'enquêtes pour calculer l'interpolation.

Si le nombre d'enquêtes finales (NBEfin) est inférieur à 10000, alors le Prix unitaire finale (PUfin) est déterminé par :

$$PU_{fin} = \frac{PU_{cible} - PU_{min}}{15500 - 10500} \cdot (NBE_{fin} - 10500) + PU_{min}$$

Si le nombre d'enquêtes finales (NBEfin) est supérieur à 10000, alors le prix unitaire finale (PUfin) est déterminé par :

$$PU_{fin} = \frac{PU_{max} - PU_{cible}}{23000 - 15500} \cdot (NBE_{fin} - 15500) + PU_{cible}$$

La durée de l'enquête

Les enquêtes sont facturées en fonction de la durée moyenne cumulée en fin de plateau (DMC) des enquêtes complètes réalisées.

Il n'y a pas de distinction dans la tarification entre les appels sur numéros fixes, mobiles ou spéciaux. Le candidat en tiendra compte dans la manière de calculer ses prix unitaires (se reporter à l'Annexe 3).

Seules les enquêtes complètes sont facturées au Céreq. Les appels infructueux, le travail de recherche du bon individu (exemple : contact avec un tiers), les questionnaires incomplets, l'animation de la hotline (téléphone et mail), ... ne seront pas facturés. Le candidat en tiendra compte (en répercutant ces coûts) dans la grille de prix proposée.

La durée du questionnaire étant inconnue au moment de la rédaction du marché. L'annexe financière prévoit toutes les situations possibles. Le candidat devra donc remplir les prix unitaires d'un questionnaire, minute par minute, dans la fourchette [14 minutes ; 24 minutes] dans les trois situations (minimum, cible, maximum). En dehors de cette fourchette, le candidat indiquera dans le deuxième tableau, le prix d'une minute supplémentaire (au-delà de 24 minutes) ou la déduction d'une minute de questionnement (en deçà de 14 minutes).

Concernant les enquêtes multimodes réalisées en partie sur internet et en partie sur téléphone, elles seront comptabilisées comme des enquêtes téléphoniques de durée égale au temps effectif passé au téléphone. Ces durées seront prises en compte dans le calcul de la DMC.

Afin de pouvoir comparer les différentes offres, la colonne totale ne sera renseignée qu'une seule fois par situation :

- Pour la situation minimum, le total ne sera calculé que pour un questionnaire de 14 minutes
- Pour la situation cible, le total ne sera calculé que pour un questionnaire de 17 minutes
- Pour la situation maximum, le total ne sera calculé que pour un questionnaire de 24 minutes

Précision sur le calcul du prix du questionnaire en fonction de la durée.

La fonction du prix d'un questionnaire en fonction du minutage ($PU(m)$) sera ainsi connue (via le renseignement des deux tableaux par le candidat) pour chaque minutage m entier.

Or, à la fin du plateau, la DMC finale (DMC_{fin}) ne sera pas forcément un nombre entier. Il existe M_{fin} un minutage tel que DMC_{fin} appartient à l'intervalle $[M_{fin} ; M_{fin}+1[$. (M_{fin} est la partie entière de DMC_{fin})

Trois cas :

- si M_{fin} est dans l'intervalle [14 ; 24], alors le prix du questionnaire sera donné par la formule suivante

$$PU(DMC_{fin}) = (PU(M_{fin} + 1) - PU(M_{fin})) \cdot (DMC_{fin} - M_{fin}) + PU(M_{fin})$$

- si DMC_{fin} est inférieur à 14, on note pxd la déduction du prix par minute,

$$PU(DMC_{fin}) = PU(25) - pxd \cdot (25 - DMC_{fin})$$

- si DMC_{fin} est supérieur à 24, on note pxs le prix de la minute supplémentaire

$$PU(DMC_{fin}) = PU(35) + pxs \cdot (DMC_{fin} - 35)$$

Ces calculs du prix en fonction de la durée sont effectués pour les trois situations (minimum, cible et maximum), puis on utilise les formules intégrant les quantités (présentées ci-dessus) pour calculer le prix unitaire du questionnaire en tenant compte du nombre d'enquêtes finales.

Les formations

L'annexe financière prévoit deux lignes correspondant à la formation des enquêteurs et des superviseurs. Les quantités sont fixées afin de permettre une comparaison de prix équitables entre les différents candidats.

En effet, prenons l'exemple d'un candidat A qui propose d'affecter 100 enquêteurs sur l'opération et un autre candidat qui propose d'en affecter 50. Faisons l'hypothèse que le coût de la formation par enquêteur est identique pour les deux candidats. En laissant libre les quantités dans l'annexe, le candidat qui propose le plus d'enquêteurs se retrouverait, toutes choses égales par ailleurs, pénalisé sur le critère du prix alors même que le dimensionnement du plateau est plus intéressant. C'est la raison pour laquelle les quantités sont strictement imposées dans cette annexe financière.

Le candidat précisera donc clairement dans sa proposition le dimensionnement prévu de son plateau dans les différentes situations (minimum, cible, maximum). Le dimensionnement du plateau (et plus précisément le nombre d'heures de production prévus) est un critère essentiel pour permettre de réaliser l'opération dans les délais impartis. La proposition du candidat pourra être inférieure à 30 ou supérieure à 90. En revanche, dans l'annexe financière, il renseignera le prix pour respectivement 30, 60, 90 enquêteurs formés (indépendamment du dimensionnement qu'il propose).

Un enquêteur formé à l'opération mais qui finalement ne participe pas ou quasiment pas à l'enquête ne sera pas facturé.

Les envois

Pour la même raison que précédemment (comparaison équitable des offres), les quantités d'envois de mails-avis et de SMS ont été arbitrairement fixées. Si les envois de mails-avis seront parfaitement connus dès le démarrage (nombre d'individus présents dans l'échantillon possédant un e-mail), les envois concernant les relances mails ou sms dépendront du nombre d'individus encore « actifs » (non définitivement classés) le jour de la relance. Des statistiques de suivi quotidiennes permettront d'arbitrer conjointement sur la date et le volume de relances à effectuer.

La transmission des fichiers et du bilan

Cette ligne de l'annexe financière correspond à la transmission des éléments suivants :

- Tests du questionnaire (réels et simulés) ;
- Fichiers de suivi statistique ;
- Fichiers intermédiaires des individus ;
- Fichier final de la totalité des individus transmis enrichis par le résultat des enquêtes ;
- Fichier historique d'appels contenant, pour chaque appel, le classement de l'appel, le numéro de téléphone, le mode d'appel (masqués/démasqués), l'ensemble des variables de la phase de contact correspondant à l'appel ;
- Bilan quantitatif et qualitatif de l'opération.

Comparaison des offres

Les offres financières seront comparées à partir du prix moyen (PM).

On considère les prix totaux correspondant aux trois situations minimum (PT_{min}), cible (PT_{cible}), maximum (PT_{max}). Le prix moyen PM est obtenu en faisant une moyenne pondérée des trois prix totaux calculés comme suit :

$$PM = 0,5 \cdot PT_{cible} + 0,3 \cdot PT_{max} + 0,2 \cdot PT_{min}$$

A.6 – Inscription au registre des traitements de données à caractère personnel du Céreq

Registre DPO

Le déclarant

Nom prénom du déclarant *	OLARIA
Département / Service / Mission *	DEEVA/EIGE
Adresse email du déclarant*	olaria@cereq.fr
Email du responsable hiérarchique du projet * (Ex : Chef de l'unité)	vignale@cereq.fr

*Les adresses email saisies recevront une copie modifiable de la déclaration.
Il est néanmoins nécessaire d'aller jusqu'à la validation définitive du formulaire pour que les mails soit générés.

Description du traitement

Quel traitement de données à caractère personnel avez-vous mis en place ?

Le traitement concerne toute opération ou tout ensemble d'opérations effectuées ou non à l'aide de procédés automatisés et appliquées à des données ou des ensembles de données à caractère personnel, telles que la collecte, l'enregistrement, l'organisation, la structuration, la conservation, l'adaptation ou la modification, l'extraction, la consultation, l'utilisation, la communication par transmission, la diffusion ou toute autre forme de mise à disposition, le rapprochement ou l'interconnexion, la limitation, l'effacement ou la destruction;

Nom du traitement	Enquête Génération 2017
N° / Ref (réservé DPO)	2019DEEVA02
Date de la première déclaration	08/04/2019
Date de mise à Jour	
Numéro de la fiche programme associée	

Les acteurs (Champs réservés au DPO)

	Nom	Adresse	CP	Ville	Pays	Tel
Responsable du traitement*	Florence Lefresne -	10 place de la Joliette	13002	Marseille	France	04911328 00
DPO**	Florence Ryk	10 place de la joliette	13002	Marseille	France	04911328 57
Représentant***					France	

Responsable(s) conjoint(s)****

France

***Responsable du traitement** : la personne physique ou morale, l'autorité publique, le service ou un autre organisme qui, seul ou conjointement avec d'autres, détermine les finalités et les moyens du traitement;

****DPO** : Data protection officer

*****Représentant** : une personne physique ou morale établie dans l'Union, désignée par le responsable du traitement ou le sous-traitant par écrit qui les représente en ce qui concerne leurs obligations respectives en vertu du présent règlement;

******Responsable conjoint** : A compléter Lorsque deux responsables du traitement ou plus déterminent conjointement les finalités et les moyens du traitement

Les finalités du traitement

Finalité : Objectif principal d'une application informatique de données personnelles.

Exemples de finalité : gestion des recrutements, gestion des clients, enquête de satisfaction, surveillance des locaux, etc.

Finalité	Description
Finalité principale Enquête multimode (internet/téléphone) Génération 2017 auprès des sortants du système éducatif français en 2016-2017.	Le Centre d'études et de recherches sur les qualifications (Céreq) produit ces enquêtes statistiques nationales qui portent sur l'insertion et le cheminement des sortants du système éducatif français lors de leurs premières années de vie active. Ces enquêtes ont pour objectif principal de produire des indicateurs statistiques sur l'insertion professionnelle (taux d'emploi, taux de chômage, taux d'emploi à durée indéterminée, etc.), selon les niveaux de formation, les filières et spécialités de formation, à destination des acteurs publics. Elles produisent ainsi des informations qui contribuent à la compréhension des processus d'insertion et des différenciations des parcours en début de carrière. 3 tests sont prévus : en juin, décembre 2019 et février 2020 puis l'enquête aura lieu d'avril à juillet 2020
Sous-finalité 1	
Sous-finalité 2	
Sous-finalité 3	
Sous-finalité 4	
Sous-finalité 5	

Fondement légal du traitement

Pour être valable, un traitement de données à caractère personnel doit être licite, ce qui signifie qu'il doit reposer sur l'un des six fondements énoncés par l'article 6 du RGPD

Sur quel fondement légal se situe la licéité du traitement ?

- e) le traitement est nécessaire à l'exécution d'une mission d'intérêt public ou relevant de l'exercice de l'autorité publique dont est investi le responsable du traitement ;
- f) le traitement est nécessaire aux fins des intérêts légitimes poursuivis par le responsable du traitement ou par un tiers (...) ».

Argumentaire licéité du traitement :

Le dispositif Génération est commandé par les tutelles du Céreq soit le ministre en charge de l'éducation et le ministre en charge du travail. Ses objectifs sont inscrits dans le COP.
De plus, depuis sa création il y a 20 ans, ce dispositif a toujours obtenu le label d'intérêt général et de qualité statistique.

Contexte et données personnelles

Les données personnelles récupérées ont-elles été transmises par un tiers ?

- Oui

Oui, par quel(s) organisme(s) ?

Pour réaliser l'échantillon, une collecte auprès des établissements dispensant de la formation initiale a été réalisée. Des données centralisées ont également été fournies par la DEPP, la DGESEP, la DARES, le ministère des sports, le ministère de la culture. Cette opération a elle-même été référencée.

Les données personnelles ont-elles été récupérées via un formulaire papier ou en ligne ? (ex: inscription à un événement Céreq)

- Non

Ces données sont-elles amenées à être accessibles en ligne ? (ex : questionnaire en ligne internet, formulaire)

- Oui

Ces données sont-elles stockées sur le serveur Céreq dans un dossier restreint, accessible uniquement aux personnes qui en ont besoin ?

- Oui

Mesures de sécurité

Quelles sont les mesures de sécurité mises en place autour de la donnée ?

Mesures de sécurité : Compte tenu de l'état des connaissances, des coûts de mise en œuvre et de la nature, de la portée, du contexte et des finalités du traitement ainsi que des risques, dont le degré de probabilité et de gravité varie, pour les droits et libertés des personnes physiques, le responsable du traitement et le sous-traitant mettent en œuvre les mesures techniques et organisationnelles appropriées afin de garantir un niveau de sécurité adapté au risque, y compris entre autres, selon les besoins
Voir article 32 du règlement

<https://www.cnil.fr/fr/reglement-europeen-protection-donnees/chapitre4#Article32>

Mesure principale		Description
Mesure de sécurité techniques (identification, pseudonymisation, chiffrement, traçabilité)	Données sécurisées par mots de passe et placées dans un dossier accessible uniquement par les personnes habilités.	
Mesures de sécurité organisationnelles (PSSI, PRA, PCA, Audit sécurité, minimisation, habilitation, rebut, sensibilisation)	Données sécurisées par mots de passe et placées dans un dossier accessible uniquement par les personnes habilités.	

Catégorie de données personnelles concernées

Catégorie des données personnelles créées

Les données personnelles collectées ont une durée de conservation, jusqu'à ce que la finalité du traitement ait été atteinte. Il est interdit de conserver un fichier contenant des données à caractère personnel après ce délai. Vos données anonymisées ne sont plus soumises à la loi informatique et liberté.

	Traitement ? O/N	Description	Délai d'effacement
Etat civil, identité, données d'identification, images...	<input type="radio"/> Oui	Coordonnées des individus (Nom, prénom, adresse postale, téléphones, mail)	10 ans
Vie personnelle (habitudes de vie, situation familiale, etc.)	<input type="radio"/> Oui	Des questions sont posées en cours d'enquêtes sur la situation familiale	Sur demande de l'individu
Vie professionnelle	<input type="radio"/> Oui	Des questions sont posées en cours d'enquêtes sur la situation professionnelle des individus au cours des 3 dernières années	Sur demande de l'individu
Informations d'ordre économique et financier (revenus, situation financière, situation fiscale, etc.)	<input type="radio"/> Oui	Les revenus sont demandés	Sur demande de l'individu
Données de connexion (adress IP, logs, etc.)	<input type="radio"/> Oui	Seuls les dates et heure de logs sont récupérées	Sur demande de l'individu
Enregistrement de la voix	<input type="radio"/> Oui	Aléatoirement, certaines enquêtes sont enregistrées afin de procéder au contrôle qualité de l'enquête	
Données de localisation (déplacements, données GPS, GSM, etc.)	<input type="radio"/> Non		
Vidéosurveillance	<input type="radio"/> Non		
Vidéos diverses pour site Internet	<input type="radio"/> Non		
D'autres informations	<input type="radio"/> Non		

Le traitement repose t'il sur des données de personnes de moins de 16 ans ?

Jamais

Données sensibles

Attention, les données sensibles énumérées ici ne peuvent pas être collectée sans des démarches spécifiques auprès de la Cnil.

Traitement ? O/N	Description	Délai d'effacement
Données révélant l'origine raciale ou ethnique <input type="checkbox"/> Oui	Des questions sur la discrimination en emploi sont posées. Nous pouvons savoir s'il s'agit d'une discrimination en lien avec la couleur de peau l'origine, la religion etc mais ne demandons pas le détail. La seule donnée que nous ne contrôlons pas c'est lorsque l'individu répond "autre" et saisie sa réponse en clair.	Sur demande de l'individu
Données révélant les opinions politiques <input type="checkbox"/> Non		
Données révélant les convictions religieuses ou philosophiques <input type="checkbox"/> Non		
Données révélant l'appartenance syndicale <input type="checkbox"/> Non		
Données génétiques <input type="checkbox"/> Non		
Données biométriques aux fins d'identifier une personne physique de manière unique <input type="checkbox"/> Non		
Données concernant la santé <input type="checkbox"/> Oui	Nous demandons seulement si l'individu est en bonne santé et s'il présente un handicap avec en réponse oui / non / ne veut pas dire. En aucun cas nous demandons des détails	
Données concernant la vie sexuelle ou l'orientation sexuelle <input type="checkbox"/> Non		
Données relatives à des condamnations pénales ou infractions <input type="checkbox"/> Non		
Numéro d'identification nationale unique (NIR pour la France) <input type="checkbox"/> Non		

Catégorie de personnes concernées par le traitement

Lister tous les types de personnes faisant l'objet du traitement de données.
Exemple : salariés, clients, patients, prospects ...

Type de personnes	Description
Catégorie de personnes 1 <input type="text" value="Étudiants"/>	Individus sortant du système éducatif en 2016-2017

Catégorie de personnes 2	<input type="text" value="Élèves"/>	<input type="text" value="Individus sortant du système éducatif en 2016-2017"/>
Catégorie de personnes 3	<input type="text"/>	<input type="text"/>
Catégorie de personnes 4	<input type="text"/>	<input type="text"/>
Catégorie de personnes 5	<input type="text"/>	<input type="text"/>

Les destinataires

Si vous travaillez avec des prestataires, des partenaires etc. et que vous devez transférer des fichiers de données personnelles, vous devez opter pour un moyen sécurisé : ex. filsender + cryptage des données avec mot de passe + fourniture du mot de passe sur un autre canal que l'envoi du fichier.

Destinataire		Type de destinataire	Description
Destinataires 1	<input type="text" value="Médiamétrie"/>	<input type="text" value="Sous-traitants"/>	<input type="text" value="Prestataire en charge de la collecte"/>
Destinataires 2	<input type="text" value="Non connu"/>	<input type="text" value="Sous-traitants"/>	<input type="text" value="Prestataire en charge de l'enrichissement des coordonnées téléphonique"/>
Destinataires 3	<input type="text" value="Non connu"/>	<input type="text" value="Sous-traitants"/>	<input type="text" value="Prestataire en charge de l'impression et du routage des courriers papiers."/>
Destinataires 4	<input type="text" value="Non connu"/>	<input type="text" value="Sous-traitants"/>	<input type="text" value="Prestataire RNVP et rachat des déménagés"/>
Destinataires 5	<input type="text"/>	<input type="text"/>	<input type="text"/>

Transferts hors UE

	Destinataire	Pays	Type de garantie
Organisme destinataire 1	<input type="text"/>	<input type="text"/>	<input type="text"/>
Organisme destinataire 2	<input type="text"/>	<input type="text"/>	<input type="text"/>
Organisme destinataire 3	<input type="text"/>	<input type="text"/>	<input type="text"/>
Organisme destinataire 4	<input type="text"/>	<input type="text"/>	<input type="text"/>

Validation de la déclaration

!!! Information !!!

Vous allez valider votre déclaration de traitement, une copie modifiable va vous être envoyée. Votre DPO recevra aussi une copie de votre déclaration de manière automatique, il prendra contact avec vous pour valider certains points.

Pour rappel, les données personnelles que vous traitez doivent être sécurisées. Elles ne doivent pas être déformées, leur intégrité doit être respectée, et seules les personnes autorisées peuvent y avoir accès.

Pour ce faire, et pour sécuriser au mieux toutes les procédures, vous vous engagez dans le meilleur du possible à :

- Ne pas copier sur des supports amovibles des données à caractère personnel
- Ne pas copier de données à caractère personnel en local sur votre ordinateur mais à privilégier le réseau qui est sécurisé, et sauvegardé
- A restreindre les accès des données à caractère personnel aux seules personnes habilitées
- A sécuriser au maximum les transferts de données à caractère personnel (filesender, zipp + mot de passe etc.)
- A responsabiliser au mieux les sous-traitants qui manipulent ces données personnelles
- A contacter votre DPO et la directrice du Céreq en cas de vol de données et dès que vous en aurez pris connaissance
- A supprimer / anonymiser les données à caractère personnelles une fois que la finalité du traitement est atteinte

Le service des usages numériques et de l'immobilier est là pour assurer la sécurité du système d'information et mettre à disposition des outils pertinents, votre DPO pour vous conseiller et vous aider dans les démarches à mettre en oeuvre pour chaque traitement.

En cas de doute, ou pour toute question, prenez contact avec votre DPO !

Cliquer sur "Enregistrer la déclaration" pour qu'elle soit sauvegardée.

Informations du PIA

Nom du PIA

Enquête Génération

Nom de l'auteur

Arnaud PHILIPPE / Lucie KELLER / Sébastien VALSEMEY (Médiamétrie)

Nom de l'évaluateur

Manon OLARIA (Cereq)

Nom du validateur

Florence RYK (Cereq)

Date de création

10/04/2019

Contexte

Vue d'ensemble

Quel est le traitement qui fait l'objet de l'étude ?

L'objectif de l'enquête Génération est d'interroger l'ensemble des personnes sorties du cursus scolaire (au sens large, c'est à dire incluant collège, lycée, université, grande école, etc...) depuis trois ans pour analyser leur parcours scolaire, leur parcours professionnel et plus généralement leur insertion sur le marché de l'emploi selon leurs caractéristiques individuelles.

Quelles sont les responsabilités liées au traitement ?

Quels sont les référentiels applicables ?

Règlement Général sur la Protection des Données UE 2016/679

Loi dite Informatique & Libertés n°78-17 modifiée

Loi sur l'obligation, la coordination et le secret en matière de statistiques de 1951

Appel d'offres public du CEREQ, soumis au Code de l'Achat Public

Données, processus et supports

Quelles sont les données traitées ?

Données sources

Nom, Prénom, Adresse électronique, Adresse postale, Numéros de Téléphone, Année de sortie du système scolaire, dernier établissement fréquenté, mois et année de naissance, formation et spécialité (attention degré nominatif plus poussé dans le cadre d'une thèse)

Données générées par les systèmes

ID spécifique pour chaque interviewé, les adresses IP (sans géolocalisation)...

Données enrichies auprès de l'individu et les données recueillies auprès d'un tiers (vie professionnelle dont la rémunération, caractéristiques socio-démo, données construites autour de l'individu)

Données sur la condition physique de l'individu (données de santé)

Données sur les origines de l'individu

Données sur les éventuelles discriminations subies par l'individu au cours de la vie professionnelle

Comment le cycle de vie des données se déroule-t-il (description fonctionnelle) ?

1- Acquisition de la donnée initiale par le CEREQ

2- Intégration de la donnée initiale dans le système d'information de Médiamétrie

3- Collecte des informations prévues par le questionnaire et selon les modalités définies

4- Restitution des données au CEREQ

5- Destruction des données

Pour plus de détails, consulter la réponse à l'appel d'offre

Quels sont les supports des données ?

Données en base Cereq - Stockées directement

RENATER - Transfert de données

Système d'Information de Médiamétrie

Les données sont stockées sur les systèmes de l'entreprise et dans les Data Center dédiés à Médiamétrie

Transfert de données chez le routeur et l'imprimeur : A compléter en fonction des choix réalisés (publication des marchés en septembre - à reprendre ensuite une fois la sélection faite)

Système d'information de NP6

Microsoft - Office 365

Principes fondamentaux

Proportionnalité et nécessité

Les finalités du traitement sont-elles déterminées, explicites et légitimes ?

Le traitement a pour finalité le recueil d'informations auprès des individus sortis du système scolaire depuis 3 ans afin de disposer de statistiques sur l'insertion professionnelle des individus.

Cette enquête est réalisée par le Centre d'Etudes et de Recherches sur les Qualifications dont les missions dans le cadre de son expertise formation-emploi sont les suivantes :

- Mener des études et des recherches sur les qualifications
- Evaluer les formations, les dispositifs et politiques publiques mis en œuvre
- Formuler des avis et des propositions

L'enquête réalisée et le traitement d'information qu'elle nécessite sont donc en parfaite adéquation avec l'une des raisons d'être du Cereq.

Enfin, préalablement à tout recueil d'information, chaque individu sera destinataire d'un courrier et/ou d'un e-mail expliquant précisément l'objet de l'enquête. L'information sera faite aussi lors du contact téléphonique uniquement pour ceux n'ayant reçu ni courrier, ni e-mail.

Cette enquête a reçu l'avis d'opportunité favorable et passera en Commission pour obtenir le label d'Interêt Général et de Qualité Statistique du CNIS (Conseil National d'Information Statistique).

Quel(s) est(sont) les fondement(s) qui rend(ent) votre traitement licite ?

La licéité du traitement repose sur plusieurs fondements :

- l'intérêt légitime du Cereq à l'enquête au regard des missions dévolues du Centre de Recherche
- le label d'Interêt Général et de Qualité Statistique du CNIS (Conseil National d'Information Statistique) délivré pour la réalisation de l'étude
- le consentement des individus à répondre aux questions qui leur seront posées

INTEGRER ICI LES CARACTERISTIQUES DEVOLUES A L'ENQUETE AU REGARD DE SON "STATUT" D'ENQUETE STATISTIQUE PUBLIQUE

Les données collectées sont-elles adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées (minimisation des données) ?

-> **Phase de constitution du fichier**

Le fichier est constitué de données personnelles recueillies sur la base de l'intérêt légitime auprès de tiers (établissement scolaire) - La Commission Nationale Informatique & Libertés a émis un avis en 2012 mentionnant qu'il serait manifestement disproportionné de rechercher préalablement à la constitution de cette base le consentement de chaque individu, celui-ci ayant aux étapes suivantes la possibilité de s'opposer au traitement de ses données personnelles. Par ailleurs, la démarche est centrée sur le recueil des informations auprès de l'individu lui-même.

QI - Phase d'identification internet

Les questions posées permettent d'informer l'individu de l'objet de l'étude, de l'identité du Responsable du Traitement et de s'assurer que le répondant est bien l'individu attendu

PDC - Phase de contact

Equivalent téléphonique du module QI

Ce module comporte des questions qui permettent de recueillir des données personnelles sans le consentement de l'individu qui en est le détenteur.

- la légalité du recueil de ces informations selon ce mode opératoire se fonde sur l'intérêt légitime du CEREQ.
- le recueil de ces données selon ce mode opératoire tire sa légalité du label 'Statistique Publique' attribué à l'étude.

Questionnaire filtre

Les questions posées visent à déterminer si l'individu en ligne est bien celui que l'on cherche à interroger et s'il est éligible à l'enquête.

A la fin de ce module, l'individu choisit le mode (téléphone ou Internet) par lequel il va répondre au questionnaire.

Actualisation des coordonnées

L'objectif de ce chapitre est de s'assurer que les données de contacts figurant dans la base sont à jour.

Parcours de formation

Les données recueillies dans le cadre de ce chapitre sont en lien direct avec l'objectif de l'étude menée (cf. finalités du traitement)

Parcours professionnel

Les données recueillies dans le cadre de ce chapitre sont en lien direct avec l'objectif de l'étude menée (cf. finalités du traitement)

Le module "Autre situation du passé" va engendrer pour les questions AUP010T, AUP010W, AUP020, AUA010T, AUA10W et AUA020 la possible collecte d'informations relatives à l'état de santé de l'individu répondant. La collecte de ces informations sert la finalité globale de l'étude. Ces informations, lorsqu'elles seront recueillies par téléphone, ne feront pas l'objet d'un consentement écrit tel que la Commission Nationale Informatique & Libertés le recommande pour ce type de données. La légalité du recueil de ces informations selon ce mode opératoire se fonde sur l'intérêt légitime du CEREQ.

Le recueil de ces données selon ce mode opératoire tire sa légalité du label "Statistique Publique" attribué à l'étude.

Caractéristiques de l'individu

Ce chapitre permet de recueillir des informations sur les caractéristiques socio-démographiques de l'individu répondant et ainsi d'identifier des parcours professionnels selon des caractéristiques sociologiques.

La combinaison des réponses aux questions SD050 SD060 SD070 SD080 SD090 SD100 donne des indications sur les origines ethniques de l'individu répondant. Elles prennent donc ainsi le caractère de données sensibles. Ces informations, lorsqu'elles seront recueillies par téléphone, ne feront pas l'objet d'un consentement écrit tel que la Commission Nationale Informatique & Libertés le recommande pour ce type de données. La légalité du recueil de ces informations selon ce mode opératoire se fonde sur l'intérêt légitime du CEREQ.

Le recueil de ces données selon ce mode opératoire tire sa légalité du label "Statistique Publique" attribué à l'étude.

Les questions SD210 FISD220 ont trait à l'état de santé de l'individu. Le recueil de ces informations permet d'analyser l'impact de la santé des individus sur leur parcours professionnel. Ces informations, lorsqu'elles seront recueillies par téléphone, ne feront pas l'objet d'un consentement écrit tel que la Commission Nationale Informatique & Libertés le recommande pour ce type de données. Le recueil de ces données selon ce mode opératoire tire sa légalité du label "Statistique Publique" attribué à l'étude.

Dans la phase de test (label non encore attribué), la légalité du recueil de ces informations selon ce mode opératoire se fonde sur l'intérêt légitime du CEREQ.

Les questions DI010 DI0202 DI030 DI040 traitent de situation de discrimination à l'embauche ou dans l'emploi de l'individu. Le recueil de ces informations est utile à l'étude dans la mesure où une discrimination peut avoir un impact sur le parcours professionnel de l'individu. Au regard des autres informations déjà enregistrées, ces données peuvent, par combinaison prendre le caractère de données sensibles. Ces informations, lorsqu'elles seront recueillies par téléphone, ne feront pas l'objet d'un consentement écrit tel que la Commission Nationale Informatique & Libertés le recommande pour ce type de données. La légalité du recueil de ces informations selon ce mode opératoire se fonde sur l'intérêt légitime du CEREQ. Le recueil de ces données selon ce mode opératoire tire sa légalité du label "Statistique Publique" attribué à l'étude.

Actualisation des coordonnées

Le recueil de ces informations permet de sécuriser tout nouveau contact qui pourrait être nécessaire, notamment pour les études qualitatives ou quantitatives complémentaires prévues.

Les données sont-elles exactes et tenues à jour ?

Les données directement identifiantes sont fournies par les établissements scolaires français et certains ministères au Cereq. Ces données peuvent être enrichies via des annuaires téléphoniques.

Etablissement

Ministère

Toutes les autres informations recueillies le sont sur une base déclarative des individus.

Les données sont actualisées pendant l'interview par l'individu.

Quelle est la durée de conservation des données ?

Mesures protectrices des droits

Comment les personnes concernées sont-elles informées à propos du traitement ?

Les individus sont informés à plusieurs étapes :

- par le courrier papier et/ou le courriel qui leur est adressé préalablement à l'enquête
- au moment du contact téléphonique avec l'enquêteur qui lui rappelle l'objectif du traitement avant le recueil d'autres informations
- au moment du recueil d'information lors de l'interview via Internet

Point d'information Internet via Site Cereq, Page FB dédiée, Twitter

Évaluation : En attente

Si applicable, comment le consentement des personnes concernées est-il obtenu ?

- Le traitement des données personnelles effectué pour le premier contact est réalisé sur la base de l'intérêt légitime du Cereq à traiter ces informations
- Le traitement des données personnelles ultérieur est réalisé avec le consentement des individus lorsqu'ils acceptent de répondre au questionnaire

Question de fin de questionnaire : ré-interrogation et/ou entretien avec un chercheur

Évaluation : En attente

Comment les personnes concernées peuvent-elles exercer leurs droit d'accès et droit à la portabilité ?

Dans le cadre de cette enquête, le droit à la portabilité sera traité de façon identique au droit d'accès.

Concernant le droit d'accès, une adresse électronique générique sera créée pour permettre aux individus de formuler leurs demandes. Celles-ci seront traitées par Médiamétrie en coordination avec le Cereq.

- Seules seront fournies les informations directement données par l'individu dans le cadre de cette enquête. C'est à dire que l'intitulé des questions ne sera pas transmis, ni les variables calculées

Évaluation : En attente

Comment les personnes concernées peuvent-elles exercer leurs droit de rectification et droit à l'effacement (droit à l'oubli) ?

L'adresse électronique générique dpo-generation@cereq.fr permettra à chacun des individus de formuler ses demandes.

Celles-ci seront traitées par Médiamétrie en coordination avec le Cereq.

Certains membres de l'équipe de Médiamétrie auront un accès à la boîte mail dédiée pour traiter les demandes.

Évaluation : En attente

Comment les personnes concernées peuvent-elles exercer leurs droit de limitation et droit d'opposition ?

L'adresse électronique générique permettra à chacun des individus de formuler ses demandes. Celles-ci seront traitées par Médiamétrie en coordination avec le Cereq.

Évaluation : En attente

Les obligations des sous-traitants sont-elles clairement définies et contractualisées ?

cf. Appel d'offres, réponse à l'appel d'offres et lettre d'engagement

Évaluation : En attente

En cas de transfert de données en dehors de l'Union européenne, les données sont-elles protégées de manière équivalente ?

Sans objet en l'espèce

A voir pour la réponse aux individus situés hors UE exerçant leurs droits

Évaluation : En attente

Risques

Mesures existantes ou prévues

Chiffrement des flux de données

Fonctionnalité fournie nativement par Askia

Cloisonnement des informations

Traçabilité des accès

Protection des locaux

Contrôle d'accès par badge nominatif

Les zones sensibles sont restreintes au personnel habilité

Présence systématique d'un agent de sécurité sur site pendant les périodes de vacation

Vidéo-surveillance, alarme anti-intrusion, alarme incendie

Politique de Sécurité - Politique de Sécurité des Systèmes d'Information - Charte SI

Politique de gestion des données personnelles

Sensibilisation des collaborateurs

Réalisation d'audits

Privacy By Design

Gestion des droits d'accès aux données

Gestion et supervision de l'ensemble du système d'information

Astreinte 24/7 avec des ressources dédiées

Accompagnement par un CERT pour la veille sécurité

Protection Internet

Blocage sur les pages d'identification des injections de type SQL

Gestion des identifications

Détection éventuelle de tentatives répétées d'identification erronées (prévention des attaques par force brute)

Identifiant à 50 caractères

Sécurité du Code Source

Utilisation du framework Symfony 4.2

Protection globale des infrastructures informatiques

Segmentation du réseau

Pare-feux réseau (Checkpoint) et applicatifs (DenyAll)

Maîtrise des flux Internet sortants (Proxy Zscaler)

Hebergement sécurisé des données

Centres de Données Equinix certifiés ISO 27001 / Tier 3 minimum

Liaisons électriques, réseaux et climatisation redondées

Accès nominatifs et biométriques

Suite privative

Traçabilité, suivi et audit des accès aux données

Surveillance via l'utilisation de Varonis DataGuard

Sauvegarde régulière des données

Veille et alerte des informations de traçabilité

Accès illégitime à des données

Quels pourraient être les principaux impacts sur les personnes concernées si le risque se produisait ?

Usurpation d'identité, Sollicitation des personnes concernées pour du marketing direct non désiré, Divulgence des coordonnées pour la réalisation d'opérations frauduleuses de type phishing, Divulgence d'informations confidentielles (ex : salaire, coordonnées, etc...), Chantage ou dénonciation (certains individus travaillant sans être déclarés peuvent l'indiquer en clair dans certaines questions), Discrimination à l'embauche (problèmes de santé déclaré en cours d'études ou actuellement, information sur des arrêts de travail, motif de licenciement), Discrimination à l'accès au logement ou d'un autre service en raison de l'indication de difficultés financières dans le cadre de l'étude, Atteinte à l'image d'une personnalité publique pour des faits qui ne font pas partie du débat d'intérêt général

Quelles sont les principales menaces qui pourraient permettre la réalisation du risque ?

Introduction frauduleuse dans le système d'information par une personne extérieure à l'entreprise, Introduction frauduleuse dans le système d'information contenant les données personnelles des interviewés par un collaborateur non autorisé, Copie et extraction automatisée des données personnelles par un outil, Copie et extraction manuelle des données par un collaborateur malveillant

Quelles sources de risques pourraient-elles en être à l'origine ?

Tiers malintentionné et disposant de moyens et de compétences pour s'introduire dans les systèmes de l'entreprise, Collaborateur malveillant, Organisation illégale visant à dérober des données à grande échelle par des moyens automatisés

Quelles sont les mesures, parmi celles identifiées, qui contribuent à traiter le risque ?

Politique de Sécurité - Politique de Sécurité des Systèmes d'Information - Charte SI, Politique de gestion des données personnelles, Sensibilisation des collaborateurs, Privacy By Design, Chiffrement des flux de données, Gestion des droits d'accès aux données, Traçabilité des accès, Hebergement sécurisé des données, Sécurité du Code Source, Protection Internet, Cloisonnement des informations, Protection des locaux, Gestion des identifications, Traçabilité, suivi et audit des accès aux données, Protection globale des infrastructures informatiques

Comment estimez-vous la gravité du risque, notamment en fonction des impacts potentiels et des mesures prévues ?

Importante, La gravité du risque est ici notée à son maximum mais dépendrait en réalité du moment où l'accès illégitime aux données interviendrait : en tout début de processus les données présentes dans les systèmes sont bien plus limitées que les données présentes à l'issue du recueil de l'information.

Toutefois, certaines mesures devraient permettre de limiter la gravité, notamment par le cloisonnement des informations.

Comment estimez-vous la **vraisemblance du risque**, notamment au regard des menaces, des sources de risques et des mesures prévues ?

Négligeable, Les différentes mesures prévues constituent des barrières efficaces au regard des menaces identifiées. Ainsi, il semble assez négligeable qu'elles surviennent.

Évaluation : Acceptable

Commentaire d'évaluation :

Vu avec M.OLARIA et F.RYK

Modification non désirées de données

Quels pourraient être les principaux **impacts sur les personnes concernées** si le risque se produisait ?

L'étude n'aboutissant à aucune "contre-partie", la modification des données n'aurait aucun impact sur les personnes concernées

Quelles sont les principales **menaces** qui pourraient permettre la réalisation du risque ?

Absence de sécurisation, Erreur de programmation, Bug Logiciel, Malveillance interne, Erreur-Négligence

Quelles **sources** de risques pourraient-elles en être à l'origine ?

Humain, Source logique

Quelles sont les **mesures**, parmi celles identifiées, qui contribuent à traiter le risque ?

Cloisonnement des informations, Sensibilisation des collaborateurs, Gestion et supervision de l'ensemble du système d'information, Traçabilité des accès, Politique de Sécurité - Politique de Sécurité des Systèmes d'Information - Charte SI, Protection des locaux, Hébergement sécurisé des données, Protection globale des infrastructures informatiques, Sauvegarde régulière des données, Veille et alerte des informations de traçabilité

Comment estimez-vous la **gravité du risque**, notamment en fonction des impacts potentiels et des mesures prévues ?

Négligeable, En matière d'atteinte à la vie privée des individus, une modification des données serait de gravité négligeable (cf. Impacts)

Comment estimez-vous la **vraisemblance du risque**, notamment au regard des menaces, des sources de risques et des mesures prévues ?

Limitée, OK

Évaluation : Acceptable

Disparition de données

Quels pourraient être les principaux **impacts sur les personnes concernées** si le risque se produisait ?

Aucun impact

Quelles sont les principales **menaces** qui pourraient permettre la réalisation du risque ?

Malveillance interne, Erreur de manipulation des données, Malveillance externe, Dysfonctionnement de l'ensemble du SI entraînant un effacement des données

Quelles **sources** de risques pourraient-elles en être à l'origine ?

Humain, Source logique

Quelles sont les **mesures**, parmi celles identifiées, qui contribuent à traiter le risque ?

Gestion des identifications, Traçabilité des accès, Sensibilisation des collaborateurs, Protection globale des infrastructures informatiques, Veille et alerte des informations de traçabilité, Sauvegarde régulière des données, Protection des locaux, Cloisonnement des informations

Comment estimez-vous la **gravité du risque**, notamment en fonction des impacts potentiels et des mesures prévues ?

Négligeable, La gravité est estimée ici au regard de l'impact sur la vie privée des individus

Comment estimez-vous la vraisemblance du risque, notamment au regard des menaces, des sources de risques et des mesures prévues ?

Limitée, Les mesures de protection, la sensibilisation des collaborateurs et les sauvegardes régulières des données limitent ce risque

Évaluation : Acceptable

Commentaire d'évaluation :

Vu avec M.OLARIA et F.RYK

Plan d'action

Vue d'ensemble

Principes fondamentaux

- Finalités
- Fondement
- Données adéquates
- Données exactes
- Durée de conservation
- Information des personnes
- Recueil du consentement
- Information des personnes
- Droit de rectification et d'effacement
- Droit de limitation et d'opposition
- Sous-traitance
- Transferts

Mesures existantes ou prévues

- Chiffrement des flux de données
- Cloisonnement des informations
- Traçabilité des accès
- Protection des locaux
- Politique de Sécurité - Politique de Sécurité des Systèmes d'Information - Charte SI
- Politique de gestion des données personnelles
- Sensibilisation des collaborateurs
- Réalisation d'audits
- Privacy By Design
- Gestion des droits d'accès aux données
- Gestion et supervision de l'ensemble du système d'information
- Protection Internet
- Gestion des identifications
- Sécurité du Code Source
- Protection globale des infrastructures informatiques
- Hebergement sécurisé des données
- Traçabilité, suivi et audit des accès aux données
- Sauvegarde régulière des données
- Veille et alerte des informations de traçabilité

Risques

- Accès illégitime à des données
- Modification non désirée de données
- Disparition de données

Mesures Améliorables
Mesures Acceptables

Principes fondamentaux

Aucun plan d'action enregistré.

Mesures existantes ou prévues

Aucun plan d'action enregistré.

Risques

Aucun plan d'action enregistré.