



**Conseil national
de l'information statistique**

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

Réunion du 12 mai 2022

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

**Collecte de la base de sondage de l'enquête
Génération 2021**

1. Titre

Intitulé de l'enquête	Collecte de la base de sondage de l'enquête Génération 2021
-----------------------	---

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	Établissement
Champ statistique couvert	<p>L'enquête s'adresse aux établissements de formation délivrant de la formation initiale certifiante, situés en France.</p> <p>Les établissements de formation pour lesquels le Céreq collecte des fichiers d'inscrits directement auprès des ministères centralisateurs ne sont pas concernés par cette enquête.</p>
Champ géographique	L'enquête a pour champ géographique les établissements de formation situés en France métropolitaine ou dans un département ou région d'outre-mer.

4. Présentation de l'enquête

Titre du(es) questionnaire(s)	Collecte de la base de sondage de l'enquête Génération 2021
URL de consultation du questionnaire	
Objectifs	<p>Cette collecte est une étape indispensable pour la réalisation de l'enquête Génération 2021 du Céreq.</p> <p>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 doit donc constituer cette base de sondage en amont de chaque nouvelle enquête Génération, à partir de différentes sources et en opérant divers traitements.</p> <p>Deux principales opérations de collecte ont lieu pour récupérer des listes nominatives d'élèves ou d'étudiants inscrits dans un</p>

établissement de formation initiale français en 2020-2021 et 2021-2022.

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

- le fichier des élèves inscrits dans le second degré, 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 par le Céreq : c'est précisément de cette enquête qu'il s'agit ici. La liste des établissements à enquêter est 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 d'inscrits en 2020-2021 et 2021-2022 permet de faire des comparaisons de fichiers afin de repérer les poursuites d'études : les inscrits en 2020-2021 et poursuivants en 2021-2022 ne seront pas conservés dans la base de sondage de l'enquête Génération 2021.

Thèmes du (ou des) questionnaires	<p>Les informations demandées aux établissements de formation portent sur les élèves inscrits au cours des deux années scolaires 2020-2021 et 2021-2022. Elles doivent permettre d'identifier les individus (état civil, INE, coordonnées) et de caractériser la formation suivie au cours de l'année scolaire concernée (niveau et spécialité).</p> <p>Plus précisément, les informations demandées sont les suivantes: nom et prénom(s), INE (lorsque celui-ci est disponible), adresse, numéro(s) de téléphone, mail, date et lieu de naissance, sexe, classe et spécialité de formation, obtention ou non du diplôme préparé.</p>
Historique	<p>Le dispositif Génération a été mis en place à la fin des années quatre-vingt-dix. Huit générations de sortants du système éducatif ont été interrogées jusqu'à présent : les Générations 1992, 1998, 2001, 2004, 2007, 2010, 2013 et 2017.</p> <p>Pour chacune de ces générations, une collecte de fichiers d'élèves a été réalisée pour construire la base de sondage : pour partie auprès de ministères, pour partie directement auprès des établissements de formation. Jusqu'à la Génération 2017, l'enquête auprès des établissements de formation n'avait pas fait l'objet d'une demande de labellisation par le comité du Label. En vue de l'enquête auprès de la Génération 2021 (nom complet : "Enquête de cheminement tous niveaux de sortie du système éducatif. Génération 2021."), le Céreq a décidé de faire de cette collecte une enquête de la statistique publique à part entière.</p>
Faits marquants précédentes éditions	sans objet
Concertation	sans objet
Origine de la demande	Le comité du label lui-même a encouragé le Céreq à faire labelliser cette enquête auprès des établissements de formation (lors de la labellisation de l'enquête Génération 2017 "Première interrogation auprès d'un échantillon de sortants du système éducatif français

	<p>en 2017”, en décembre 2019), face aux difficultés dont le Céreq a fait part au sujet de cette phase préliminaire à l’enquête Génération 2017. L’entrée en vigueur de la RGPD a rendu les établissements de formations plus réticents à transmettre des données nominatives. L’objectif de la labellisation de cette collecte, avec caractère obligatoire, est de fournir aux établissements de formation une base légale (l’obligation légale en l’occurrence) pour fournir les données nominatives demandées par le Céreq, indispensables à la réalisation de l’enquête Génération 2021.</p>
Place dans un dispositif statistique	<p>Cette collecte est une phase préliminaire de l’enquête Génération 2021, enquête panélysée qui donnera lieu à deux interrogations (“Première interrogation auprès d’un échantillon de sortants du système éducatif français en 2021” et “Deuxième interrogation auprès d’un échantillon de sortants du système éducatif français en 2021”)</p>
Extensions géographiques	<p>sans objet</p>
Utilisateurs	<p>Le Céreq sera le seul utilisateur des données collectées. Un plan de gestion des données décrira l’ensemble des opérations réalisées ainsi que les durées de conservation des données.</p>

5. Caractéristiques techniques

Périodicité de l'enquête	<p>Ponctuelle ou pluri-annuelle</p>
Période de collecte	<p>La collecte se déroulera de janvier à mai 2023.</p>
Mode de collecte	<p>Il s’agit d’une enquête monomode téléphonique.</p>
Comitologie	<p>sans objet</p>
Contraintes pour l'enquête	<p>La contrainte pour les établissements contactés ne réside pas dans le temps passé au téléphone mais dans le temps de préparation des fichiers demandés par le Céreq. Le Céreq cherche autant que possible à minimiser ces efforts : par exemple, concernant les</p>

	établissements dépendant d'Apogée (universités notamment), un patch est développé en collaboration avec l'Amue pour faciliter l'extraction des données et leur mise en forme.
Coût de l'enquête	<p>Six personnes et demi (6,5 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 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.</p> <p>Concernant la collecte de la base de sondage, une personne de l'équipe est mobilisée pour le suivi du plateau téléphonique. Deux à trois vacataires sont recrutés pour une période d'environ 5 mois, afin de gérer la réception des fichiers et la vérification des informations contenues.</p> <p>La gestion des appels pour collecter les fichiers est confiée à une société externe choisie selon la procédure officielle des marchés publics. Pour cette enquête, le coût externe est estimé à environ 70 000 euros HT</p>
Organisme collecteur	<p>La collecte des fichiers sera confiée à un prestataire sous-traitant choisi selon la procédure officielle des marchés publics.</p> <p>Conformément au cahier des charges et des clauses techniques rédigé par le Céreq, il sera chargé de concevoir le développement informatique d'effectuer les contacts par téléphone auprès des établissements de formation (premier contact, relances).</p>
Plan de sondage	<p>Il n'y a pas de tirage d'échantillon, le but étant de couvrir l'intégralité des inscrits dans des établissements de formation en 2020-2021 et 2021-2022.</p> <p>Tous les établissements délivrant de la formation initiale non couverts par les bases centralisées (voir rubrique "Objectifs") sont concernés par cette collecte.</p>
Taille de l'échantillon	Environ 3500 établissements sont sollicités

6. Service(s) producteur(s)

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

Partenariat	Des partenariats seront noués pour l'enquête Génération 2021, que cette enquête préliminaire permettra
Date prévisionnelle de la première publication	Les publications suivront l'enquête Génération 2021 : cette enquête aura lieu en 2024 et les premières publications paraîtront en 2025.