

Fiche de poste

Contractuel chargé.e de suivi pour la transcription, l'anonymisation manuelles et le contrôle qualité du corpus « Cahiers citoyens »

VOTRE ROLE

Le 11 mars 2025, l'Assemblée Nationale a adopté une résolution relative à la publicisation des 20 000 « cahiers de doléances » ou « cahiers citoyens » collectés dans le cadre du Grand Débat National de 2019, à la suite du mouvement des Gilets jaunes. Ces cahiers, produits dans 17 000 mairies, regroupent plus de 225 000 contributions ; il s'agit d'un des plus vastes corpus d'expression citoyenne de notre époque. Sa mise en ligne répond à une exigence démocratique mais vise aussi à permettre aux chercheurs – politistes, sociologues, linguistes, historiens... – d'exploiter cet ensemble de données unique en son genre. Ces cahiers ont fait l'objet d'une numérisation et d'une transcription (parfois fautive) ; mais pas d'une mise en ligne, faute d'avoir été dûment anonymisés.

Les conditions de la mise en ligne ont été étudiées par [une mission Flash pilotée par le Campus Condorcet](#). À la suite de cette étude, le Campus s'est vu confier, par un comité de pilotage parlementaire, la responsabilité de réaliser une preuve de concept (POC) d'une chaîne de traitement automatisée de ces cahiers. Cette chaîne, développée par les équipes de l'Humathèque Condorcet et de l'IR* Huma-Num, au sein de l'Huma-Num Lab, vise à :

- Transcrire automatiquement le contenu des cahiers citoyens, à partir des numérisations déjà existantes,
- Anonymiser automatiquement les cahiers,
- Afin de permettre leur mise en ligne.

Le POC sera réalisé sur une sous-partie du corpus (environ 5% des cahiers), afin de tester la capacité d'industrialiser le processus.

La personne recrutée assiste la cheffe de produit dans l'encadrement et le suivi de l'équipe des vacataires chargés d'une part de produire les « vérités de terrain » qui permettront d'entraîner et d'améliorer les algorithmes de transcription et d'anonymisation, et d'autre part d'assurer le contrôle qualité des transcriptions et anonymisations réalisées par les processus automatisés. La personne recrutée participe également à la production des vérités de terrain et au contrôle qualité.



VOS MISSIONS

- **Assister la cheffe de produit dans l'encadrement et le suivi de l'équipe des vacataires étudiants**
 - Gérer l'organisation des plages de présence, suivre les états de vacation
 - Suivre l'avancée des travaux, alimenter et contrôler l'alimentation des outils de suivi
 - S'assurer du bon respect des procédures et des règles de production des vérités de terrain et du contrôle qualité, centraliser et faire remonter les difficultés et les cas particuliers à la cheffe de produit
- **Produire des « vérités de terrain » à partir des fichiers image sources**
 - Transcrire des échantillons du corpus manuscrit des « corpus citoyens » selon une méthodologie donnée (règles de transcription)
- **Contrôler la qualité des données produites**
 - Relire le texte transcrit en regard des originaux, repérer les erreurs de transcription et les corriger
 - Vérifier l'anonymisation des données personnelles (nom, signature, profession...) conformément aux règles élaborées par un groupe d'experts
 - Participer à l'analyse des erreurs de transcription et d'anonymisation, repérer les erreurs récurrentes
 - Repérer les textes qui pourraient contenir des données sensibles
 - Participer à l'évolution des procédures et méthodologies de traitement
- **Participer à la chaîne de traitement et à son amélioration**
 - Assurer l'export des transcriptions et participer à la préparation des données lors des différentes étapes de la chaîne de traitement
 - Participer au test des différentes interfaces
 - Participer à l'évaluation de l'ergonomie de la chaîne de traitement et à son amélioration continue

QUI ETES VOUS ?

- **Formation**
 - En sciences humaines et sociales, en archivistique ou en humanités numériques et intéressé par le travail sur des corpus de textes et les problématiques de transcription
- **Compétences**
 - Bonne maîtrise de la langue française
 - Bonne culture générale





- Aisance avec les outils informatiques
 - Capacité à suivre des procédures et à identifier des cas particuliers
 - Connaissances en langage de programmation (Python), en base de données (SQL) et en interface (Jupyter Notebook)
 - Un goût pour la paléographie, et/ou des notions sur la problématique des données personnelles et sensibles seraient appréciées
 - Des notions dans l'entraînement de modèle IA seraient un plus
- **Qualités professionnelles**
 - Sens de l'organisation
 - Sens de la rigueur
 - Facilité à rendre compte
 - Discrétion professionnelle

POURQUOI NOUS REJOINDRE ?

Rejoindre le Campus Condorcet, c'est intégrer un **établissement public national de coopération en plein essor**, au cœur des **transformations de l'enseignement supérieur et de la recherche et du territoire de Seine-Saint-Denis**.

Le Campus Condorcet est la plus grande communauté de chercheurs en sciences humaines et sociales de France, réunissant **11** institutions d'excellence (EHESS, EPHE, etc.), **12 000** chercheurs, doctorants et étudiants et **116** unités, structures, réseaux et projets de recherche. Le Campus offre un équipement unique en France (Humathèque, Hôtel à projets, un centre de colloques...) et accueille chaque année plus **1200** manifestations scientifiques (en 2025). **Notre ambition ?** Faire des sciences humaines et sociales un fer de lance des réponses à apporter aux grands défis sociétaux du XXI^e siècle (climat, démographie, inégalités, numérique, etc.).

Vous évoluerez dans un environnement **intellectuellement stimulant**, au contact direct de projets scientifiques d'envergure nationale et internationale, **d'infrastructures de recherche d'envergure européenne** et aux côtés de partenaires publics et privés de premier plan.

L'Humathèque Condorcet est une bibliothèque de recherche et un centre d'archives de niveau international. Elle a été conçue sur le principe de mutualisation des ressources des établissements fondateurs (environ 1 million de volumes, 5 kilomètres linéaires d'archives de la recherche, une centaine de professionnels de la documentation et des archives). Elle place l'utilisateur au cœur de sa stratégie de services.

CONDITION D'EMPLOI

- **CDD de niveau catégorie A pour une durée de 6 mois**
- **Rémunération selon expérience dans la grille ASI**
- **Les activités liées à ce poste ne sont pas télétravaillables**
- **Date souhaitée de prise de poste : 1er septembre 2026**





- **Localisation :** Campus Condorcet – Aubervilliers, à proximité immédiate de Paris (ligne 12 station Front Populaire ou RER B station La Plaine - Stade de France)

VOUS ETES INTERESSE.E?

Adressez-nous votre CV et lettre de motivation à : recrutement@campus-condorcet.fr ; julie.lenoir@campus-condorcet.fr et à cynthia.pedroja@campus-condorcet.fr

Date limite de candidature : 15 août 2026

Contact pour toutes questions : cynthia.pedroja@campus-condorcet.fr



L'ÉCOLE
DES HAUTES
ÉTUDES EN
SCIENCES
SOCIALES



École
nationale
des
chartes

| PSL ★



ÉCOLE PRATIQUE
des HAUTES ÉTUDES

| PSL ★



fondation
maison des
sciences
de l'homme



Sorbonne
Nouvelle
université des cultures

UNIVERSITÉ
PARIS8
DES CRÉATIONS

Université
Paris Nanterre

UNIVERSITÉ
SORBONNE
PARIS NORD