

Insertion des jeunes : un projet innovant DEPP/DARES d'appariement de bases de l'éducation et du travail

CNIS – commission services publics et services aux publics – 12 mars 2020

Loïc Midy - DEPP

SOMMAIRE

- La genèse du projet
- La situation en 2018
- **INSERJEUNES** : ses finalités et la description du projet
- Appariements entre bases : les obstacles rencontrés
- Le cœur : l'outil d'appariement
- Les étapes clés du projet

LA GENÈSE DU PROJET

- Article 24 de la loi pour la liberté de choisir son avenir professionnel N° 2018-771 du 5 septembre 2018 :

Art. L. 6111-8 (code du travail) - Chaque année, pour chaque centre de formation d'apprentis et pour chaque lycée professionnel, sont rendus publics quand les effectifs concernés sont suffisants :

- 1° Le taux d'obtention des diplômes ou titres professionnels ;
- 2° Le taux de poursuite d'études ;
- 3° Le taux d'interruption en cours de formation ;
- 4° Le taux d'insertion professionnelle des sortants de l'établissement concerné, à la suite des formations dispensées ;
- 5° La valeur ajoutée de l'établissement.

Pour chaque centre de formation d'apprentis, est également rendu public chaque année le taux de rupture des contrats d'apprentissage conclus.

LA SITUATION EN 2018

- Des enquêtes d'insertion (enquêtes IVA/IPA) menées par la DEPP auprès de tous les jeunes sortants qui sont :
 - lourdes à gérer pour les chefs d'établissements et les différents acteurs intervenant dans la collecte (équipes statistiques DEPP et rectorat, ...)
 - coûteuses (envoi questionnaires, relances,)
 - un seul point : insertion à 7 mois après la sortie
 - Taux de réponse de l'ordre de 50% : ne permettent pas de fournir des taux d'insertion par établissement ni même au niveau national pour toutes les spécialités

INSERJEUNES : SES FINALITÉS

- Créer un nouveau SI par fusion de bases administratives « scolarité » et bases « emploi » afin de calculer chaque année au niveau établissement les indicateurs suivants :
 - taux d'insertion professionnelle des sortants de l'établissement + valeur ajoutée de l'établissement
 - taux de poursuite d'études
 - taux d'interruption en cours de formation
- Réaliser des études (ex : étudier les conditions d'emploi des jeunes sortant de formation, adéquation formation/emploi, ...) : base très riche
 - pour les chargés d'études DEPP/ DARES
 - pour les chercheurs

INSERJEUNES : DESCRIPTION DU PROJET (1)

- Le projet mobilise une vingtaine de bases « élèves » de la sphère éducation et de la sphère travail dont les principales sont :
 - Les remontées administratives des inscriptions des élèves (y compris agriculture) et des apprentis
 - la base MMO (mouvements de main-d'œuvre) de la Dares qui repose sur la DSN (Déclaration Sociale Nominative) couvrant l'emploi salarié, puis les 3 fonctions publiques à compter de 2022.
- Un projet innovant
 - Se rapprocher de l'exhaustivité
 - Construction d'indicateurs d'insertion à des niveaux très fins
 - Avoir ces indicateurs à différents moments après la sortie du système éducatif (6 mois, 12 mois, 18 mois et 24 mois)
 - Délai de mise à disposition de ces taux plus rapide

INSERJEUNES : DESCRIPTION DU PROJET (2)

- un projet en 4 grandes étapes :
 - Étape 1 : calculer le champ des sortants d'apprentissage et de voie professionnelle
 - Comparaison sur 2 années
 - 6 sources de « scolarité » : Lycées, apprentis, supérieur, lycées agricoles, vœux validés formations paramédicales et sociales dans parcours sup, privé hors contrat
 - Appariement sur l'INE (sauf privé hors contrat)
 - Étape 2 : enrichir la base des sortants avec les résultats des examens
 - Étape 3 : appariement par identifiants indirects avec la MMO/DSN (Nom, prénom, date et commune de naissance, sexe)
pas d'identifiant commun entre sources scolarité (INE) et sources emploi (NIR) !
 - Étape 4 : agrégation au niveau établissement

APPARIEMENTS ENTRE BASES : LES OBSTACLES RENCONTRES

■ Sur les méthodes

- Littérature abondante : quelle méthode choisir ?

■ Sur les outils

- de nombreux outils existent : lequel choisir ?

■ Les sources

- Amélioration du champ couvert (DOM, fonction publique, etc.)
- Amélioration des variables (ex: ajout code COG commune de naissance en plus du libellé de la commune)

■ Cadre juridique/RGPD

- Analyse d'impact en cours et 2 DPD concernés
- Procédure d'effacement des variables identifiantes

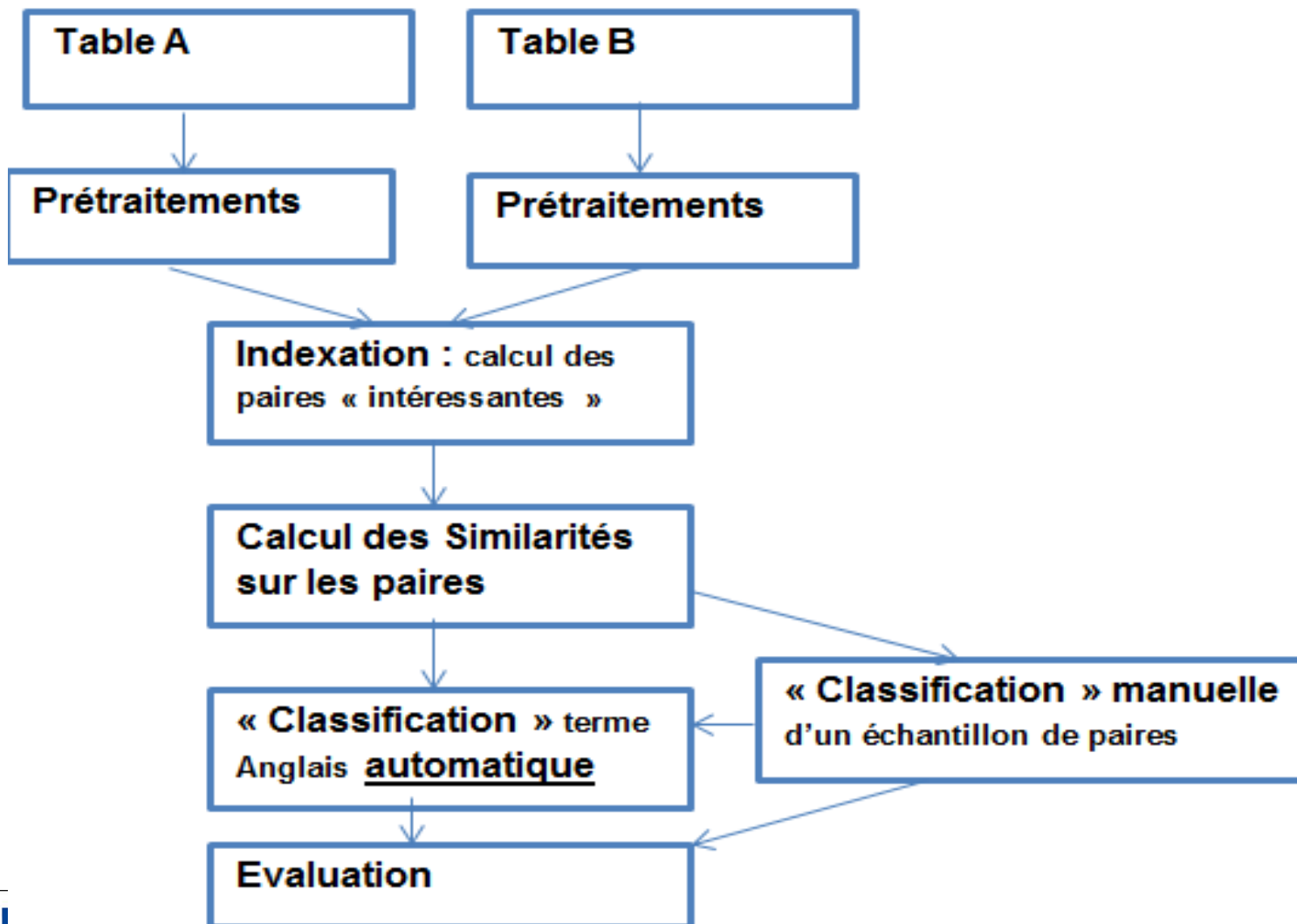
■ Sécurité

- Échanges de données DEPP/DARES : RIE + canal chiffré + fichiers chiffrés.

LE CŒUR : L'OUTIL D'APPARIEMENT (1)

- Etude de plusieurs outils et méthodes existantes : outil DEPP, « matchID », programmes SAS DARES, méthode « tours de clé » Insee
- Décision de faire un outil en utilisant Python et des librairies existantes
 - À un **coût raisonnable** (environ 60 jours/hommes)
 - **Facile à installer**
 - **Adaptable** à d'autres appariements
 - **Rapide** : l'appariement de 400 000 * 15 millions de lignes en 45 mn
 - Avec un module **qualité** sur l'appariement : taux d'appariement de 97% entre source apprentis et source contrats d'apprentissage.
- Cet outil sera mis en open source d'ici fin 2020 avec un guide d'utilisation et accompagné d'un document méthodologique

LE CŒUR : L'OUTIL D'APPARIEMENT (2)



ETAPES CLES DU PROJET

■ T4 2018 et S1 2019

- Travaux sur l'outil d'appariement et niveaux de diffusion
- Premiers calculs des indicateurs (taux d'emploi et de leur valeur ajoutée) sur les sortants à l'été 2017 et comparaison avec IVA/IPA

■ S2 2019

- Démarrage des développements informatiques
- Calcul sur les sortants à l'été 2018 et comparaison avec IVA/IPA
- Échange avec les établissements fin 2019 sur les résultats

■ 2020

- Fin des développements informatiques
- Calcul sur les sortants à l'été 2019 et échange avec les établissements
- 1^{ère} diffusion des indicateurs du nouveau SI mi-décembre 2020



Merci de votre attention