

---

# L'éducation, l'enseignement, la formation aux temps du contrôle social fondé sur les classifications de grandes masses de données : quels enjeux éthiques en éducation et en formation ?

Khansa Ghabara\*<sup>1</sup>

<sup>1</sup>Education Discours Apprentissages – Université Paris Descartes - Paris 5 : EA4071 – 45, rue des Saints Pères 75270 Paris Cedex 06, France

## Résumé

La production massive et quasi-permanente de traces liées aux actions des utilisateurs en ligne, combinée au déploiement croissant des environnements numériques pour une visée éducative offre, aujourd'hui, de vastes terrains pour l'exploitation des données d'apprentissage. Notre recherche vise à explorer les données circulant entre les différentes plateformes (ENT, ressources, solutions de vie scolaire) et à combler un manque de recherche dans le domaine.

On souhaite ainsi, faire des analyses conjointes de jeux de données issues de tâches d'apprentissage disciplinaire et de données issues de l'ENT et des outils généraux de vie scolaire, dans l'objectif de comprendre ce que l'on peut en tirer, de quelle manière on peut les exploiter et de quelle manière elles peuvent s'articuler et s'enrichir mutuellement.

Nous proposons, dans cette communication, une classification générale des données éducatives issues des plateformes numériques, avec un focus sur trois catégories identifiées dans le cadre de notre projet : les données relatives à la vie scolaire : comportant des informations statiques liées à l'organisation pédagogique (les cycles pédagogiques, les enseignements, ...), des informations liées au parcours des élèves à fréquence de production variable (bilan de cycle, bilan périodique, moyenne), des informations en forme d'observations des enseignants et enfin des informations avec une haute fréquence de production, comme par exemple les absences, les incidents, les devoirs, les notes des détaillées, etc.), les données d'utilisation des services traçant les accès aux différents services de l'ENT (emploi du temps, cahier de texte, ...) et enfin les données d'usages des ressources pédagogiques se déclinant en deux catégories : la description des ressources pédagogiques (titres, résumés, relations avec les enseignements concernés, caractéristiques d'usage comme les notes minimales et maximales pour une activité notée, ect.) et les traces d'usage engendrées par l'utilisation individuelle de ces ressources et décrivant l'interaction des élèves avec ces éléments.

Ces données peuvent inclure des informations très sensibles simplement par leur nature; c'est notamment le cas des données démographiques d'une personne mais également celles relatives à son suivi pédagogique et disciplinaire dans un contexte scolaire.

Par ailleurs, les analyses ayant souvent une finalité de prise automatisée de décisions ou

---

\*Intervenant

d'intervention sont susceptibles d'avoir un impact à court terme sur le processus d'apprentissage mais également des conséquences sur le long terme sur le parcours des élèves.

C'est la raison pour laquelle, nous souhaitons, à travers cette communication, porter une réflexion sur les problématiques d'éthique et d'équité liées à l'analyse de ces données dans un contexte éducatif.

Pour ce faire, nous nous appuyerons sur la cartographie ci-dessus et sur l'état de l'art portant sur les techniques d'analyse pratiquées dans les communautés EDM et LAK.

**Mots-Clés:** données, environnements numériques, analyse, éthique, scolaire