

ASSEMBLÉE NATIONALE

8 janvier 2016

RÉPUBLIQUE NUMÉRIQUE - (N° 3318)

Non soutenu

AMENDEMENT

N° CE29

présenté par

M. Chassaigne, M. Asensi, M. Bocquet, Mme Buffet, M. Candelier, M. Carvalho, M. Charroux,
M. Dolez, Mme Fraysse et M. Sansu

ARTICLE ADDITIONNEL**APRÈS L'ARTICLE 17, insérer l'article suivant:**

I. – Après le vingtième alinéa de l'article L. 122-5 du code de la propriété intellectuelle, il est inséré un 10° ainsi rédigé :

« 10° Les copies ou reproductions numériques réalisées à partir d'une source licite, en vue de l'exploration de textes et de données pour les besoins de la recherche publique, à l'exclusion de toute finalité commerciale. Un décret fixe les conditions dans lesquelles l'exploration des textes et des données est mise en œuvre, ainsi que les modalités de conservation et communication des fichiers produits au terme des activités de recherche pour lesquelles elles ont été produites. »

II. – Après le cinquième alinéa de l'article L. 342-3 du même code, il est inséré un 5° ainsi rédigé :

« 5° Les copies ou reproductions numériques de la base réalisées par une personne qui y a licitement accès, en vue de fouilles de textes et de données dans un cadre de recherche, à l'exclusion de toute finalité commerciale. La conservation et la communication des copies techniques issues des traitements, au terme des activités de recherche pour lesquelles elles ont été produites, sont assurées par des organismes désignés par décret. Les autres copies ou reproductions sont détruites. »

EXPOSÉ SOMMAIRE

L'exception pour la fouille automatique de données de texte (text et data mining) consiste à autoriser la recherche automatisée parmi un volume très important de textes ou de données : il est possible d'accéder à des résultats qui n'auraient pas pu être découverts par une autre méthode. Cela donnerait une force nouvelle à l'entrée de la recherche française à l'heure des mégadonnées (big data) et de réaliser des gains de productivité très importants, alors même que d'autres pays, comme le Royaume-Uni, le Japon et les Etats-Unis, ont pris une avance considérable dans ce domaine. La fouille automatisée de textes et de données, en tant qu'activité de lecture et d'extraction d'informations, est une pratique qui ne se distingue pas fondamentalement du relevé manuel des

informations qui a toujours été effectué par la recherche. Pourtant, les grands éditeurs qui détiennent la majeure partie des publications scientifiques, peuvent aujourd'hui proscrire, par des solutions contractuelles, la fouille de textes et de données aux chercheurs, même lorsque ces derniers disposent d'un accès légal à l'ensemble des publications scientifiques comprises dans les bases de données fouillées.