

LA DIFFUSION DES THÈSES DE DOCTORAT ÉLECTRONIQUES FRANÇAISES : petit tour d'horizon statistique



Grâce à l'application nationale STAR¹, épicerie du dépôt national des thèses de doctorat, l'Abes dispose du réservoir de données le plus complet concernant la diffusion des thèses de doctorat électroniques françaises. À l'heure tout à la fois de la science ouverte, de l'essor du travail à distance pour cause de pandémie, mais aussi des indicateurs comptables et des évaluations, il a semblé pertinent d'exploiter ces données et de les exposer au public. Le rapport intitulé *La diffusion des thèses de doctorat électroniques françaises – Bilan statistique 2020* a été publié à cette fin².

Un peu de méthodologie

Le rapport a été établi à partir du corpus des thèses électroniques traitées dans STAR, dans le cadre du dépôt national des thèses de doctorat. Les thèses diffusées en dehors de ce cadre – les dépôts effectués par les auteurs eux-mêmes sur TEL³ par exemple – ne sont pas comptabilisées. Le dépôt national des thèses de doctorat sous forme électronique a été rendu possible en 2006⁴, mais est d'abord facultatif. Il prend véritablement son essor à partir de 2011, puis se généralise en septembre 2016⁵, date à laquelle il est finalement rendu obligatoire par les textes réglementaires. Entre 2006 et 2016, le corpus des thèses électroniques traitées dans STAR ne représente donc pas l'ensemble des thèses de doctorat soutenues en France.

De plus, on constate qu'en dépit des obligations réglementaires, entre 500 et 1000 thèses de doctorat échappent chaque année au dépôt national. Enfin, il faut compter en moyenne un an pour qu'une thèse soit traitée. De ce fait, le traitement des thèses soutenues au cours de l'année 2019, voire de l'année 2018, est loin d'être achevé : le corpus analysé n'est donc pas exhaustif.

Les données exploitées sont : le choix de diffusion opéré par l'auteur (en accès libre sur internet ou en accès restreint sur intranet), les restrictions temporelles appliquées à la diffusion (embargo ou confidentialité), la durée de ces restrictions, la date de soutenance et le domaine disciplinaire, tous éléments qui permettent d'avoir une vision large de l'amplitude de diffusion des thèses concernées.

Les résultats

Le bilan statistique montre que 74 % des docteurs ont fait le choix de diffuser leur thèse en libre accès sur internet : 60 % immédiatement, 14 % après une période d'embargo et/ou de confidentialité. Les 26 % restants ont choisi de ne diffuser leur thèse qu'en accès restreint : 24 % sans restrictions temporelles, 2 % après une période de confidentialité.

Le pourcentage annuel de thèses disponibles en libre accès est en augmentation continue : il est passé de 67 % en 2010 à 77 % en 2018. En parallèle, on constate que le pourcentage annuel de thèses dont la diffusion est soumise à une restriction temporelle est passé de 11 % en 2010 à 21 % en 2018. En d'autres termes, les docteurs sont plus enclins à diffuser leur thèse, mais la diffusion est de plus en plus différée dans le temps. La durée moyenne des embargos, un an environ, reste stable dans le temps, de même que la durée moyenne des confidentialités, qui est égale à 3 ans.

Si on compare les pratiques des docteurs en Lettres, sciences humaines et sociales (LSHS) et celles des docteurs en Science, technologie, médecine (STM), on constate des différences importantes. Les docteurs en STM diffusent librement leur thèse dans 85 % des cas, contre 51 % des docteurs en LSHS. En revanche, l'usage des restrictions temporelles est plus répandu en STM (20 % des cas) qu'en LSHS (10 %

des cas). On observe par ailleurs des disparités au sein même de ces deux grands ensembles. La libre diffusion est plus répandue en mathématiques, informatique, physique et géologie qu'en sciences de la vie et chimie, pour lesquelles on remarque également que les embargos sont plus fréquents. Dans le cas des LSHS, ce sont les docteurs en géographie, économie, gestion, sciences de l'éducation et psychologie qui diffusent le plus, et les docteurs en histoire, littérature, droit et arts qui diffusent le moins.

Le rapport mis à disposition présente les statistiques propres à chaque discipline sans proposer d'interprétation. Ainsi, chaque communauté sera à même d'expliquer les données au regard de ses pratiques en matière de publication scientifique.

Maïté Roux

Abes – Service des thèses
m.roux@abes.fr

[1] www.abes.fr/Theses/Applications-pour-le-signallement-des-theses/Star-Signallement-des-Theses-ARchivage

[2] Disponible ici : <http://www.abes.fr/Media/Fichiers/Theses-Fichiers/Diffusion-des-theses-electroniques-francaises-bilan-statistique-2020>
Les statistiques portent sur les thèses soutenues entre le 1er septembre 2006 et le 31 décembre 2019, dont le traitement était achevé dans STAR au moment de l'extraction des données, fin mars 2020.

[3] <https://tel.archives-ouvertes.fr>

[4] Arrêté du 7 août 2006.
www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000635069

[5] Arrêté du 25 mai 2016.
www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000032587086