Littérature grise : des collections et des portails

Joachim Schöpfel

Pour trouver et sélectionner la littérature grise, il faut connaître les sources et voies de distribution, savoir évaluer la fiabilité et la qualité des documents, avoir de bonnes relations et de la patience. L'article dresse un panorama des collections disponibles.

Les collections de thèses, rapports, *working papers*, conférences et autres documents non publiés sont parfois le fruit d'un investissement de long terme de réseaux professionnels. Dans les années 80, les pays européens ont mis en place un réseau de signalement avec des collections décentralisées dont la pièce centrale – la base Sigle – existe toujours en libre accès et avec de nouvelles fonctionnalités sous le nom d'OpenGrey¹. D'autres réseaux sont moins visibles, aux périmètres et portées plus limités. Ligria, la coordination au niveau du Premier Ministre de la production grise des grandes administrations centrales en faisait partie, dans les années 90. Certains sites seront utiles pour identifier ces réseaux, comme celui de Dean Giustini de Vancouver².

Obligation de dépôt

Certaines collections résultent d'une obligation de dépôt³, basée sur une réglementation nationale ou régionale. Ainsi, la bibliothèque nationale allemande d'IST (Technische Informationbibliothek Universitätsbibliothek (TIB) de Hanovre) reçoit tous les rapports de recherche financée par le gouvernement fédéral. Sa collection contient plus de 236 000 rapports imprimés et 45 000 rapports électroniques, et croît de près de 900 rapports chaque mois. Jusqu'en 2002, la bibliothèque royale du Danemark a constitué une collection de documents et rapports du Conseil d'Europe, de l'Onu, de l'Otan, de l'OCDE, de l'Unesco et d'autres institutions internationales. Le centre russe d'IST à Moscou (VNTIC) a reçu plus de 500 000 thèses depuis 1982. Le dépôt des thèses de doctorat en France suit une autre logique, celle de collection distribuée autour d'un catalogue collectif⁴. Mais ces collections « obligatoires » présentent des risques : que faire si la subvention est insuffisante ? Et que faire, surtout, pour assurer la continuité si le mandat est suspendu ?

Stratégie d'acquisition

Certaines bibliothèques ont mis en place une politique d'acquisition sans mandat officiel de dépôt obligatoire. C'est le cas de la British Library avec sa collection de plus de 450 000 conférences en anglais et de plus de 10 millions de rapports scientifiques et techniques des secteurs privé et public, y compris de la Nasa, de

¹ Cf. l'article sur OpenGrey

² www.slideshare.net/giustinid/finding-the-hard-to-finds-searching-for-grey-gray-literature-2010. Voir aussi les sources dans l'annexe

³ Cf. aussi l'article sur le dépôt légal

⁴ Cf. l'article sur les thèses

l'ESA et de la FAO. La TIB possède 2 millions de rapports d'institutions européennes et américaines dont ceux du National Technical Information Service (NTIS) de l'US Department of Commerce. En France, la BIU Santé recueille les actes de congrès biomédicaux français et étrangers, soit 70 000 documents à partir de 1845. Elle détient aussi la plupart des thèses de médecine et de chirurgie dentaire soutenues en France et, depuis 1998, des thèses de biologie ainsi que des mémoires spécialisés (180 000 documents). Le nouveau dispositif CollEx serait l'occasion de valoriser d'autres fonds gris moins importants.

Numérisation

Les bibliothèques et centres d'IST ont entrepris de numériser d'anciennes collections, soit surtout des thèses du 18^e au 20^e siècle. Par exemple, dans le cadre du projet EThOS, avec le soutien du JISC, la British Library numérise les thèses des universités britanniques ; la TU de Delft et l'université d'Uppsala ont numérisé leurs collections de thèses anciennes tandis que le réseau catalan TDX et l'équipe PSNC de Poznan ont pris en charge la rétronumérisation de thèses de plusieurs établissements. En France, l'ANRT de Lille propose ce service aux bibliothèques universitaires.

Les numérisations de conférences, rapports ou travaux non publiés sont plus rares. Le JISC a financé plusieurs projets de *special collections* dont 600 volumes de rapports historiques de recensement de l'université d'Essex. En France, le BRGM a numérisé ses rapports pour le portail InfoTerre⁵, l'INIST son fonds historique de rapports scientifiques pour le portail Lara⁶, et le réseau des bibliothèques de mathématiques une trentaine de séminaires de 1948 à 2007 sur la plateforme Numdam⁷.

Chaque fois il s'agit d'un fonds de littérature grise clairement identifié, avec un caractère de patrimoine scientifique mais sans coordination avec d'autres initiatives. Il manque un projet d'envergure comme celui des rapports américains mis en œuvre par l'OSTI (DOE) en partenariat avec d'autres organismes (Fermi, LANL, etc.)⁸. Par ailleurs, une partie des projets soumis au financement dans le cadre de la BSN5 concerne de la littérature grise (conférences, mémoires, etc.). Aujourd'hui, ces collections numérisées sont souvent diffusées avec des documents numériques « natifs »⁹. Ici, la jonction entre patrimoine et production courante se fait d'une manière organique¹⁰.

Portails

On ne dispose d'aucune cartographie des collections de littérature grise ni d'outil - hors Google Scholar dont l'indexation est défailante et l'algorithme opaque - qui, sur le plan de la documentation non commerciale, pourrait concurrencer le Web of Sciences ou Scopus. Plusieurs portails facilitent toutefois la recherche et l'accès à ces documents, notamment les thèses, tels que Thèses.fr en France¹¹,

⁵ <http://infoterre.brgm.fr>

⁶ <http://lara.inist.fr>

⁷ www.numdam.org

⁸ www.osti.gov

⁹ C'est le cas des 45 000 thèses diffusées sur TEL dont 40 000 sont numériques « natives ».

¹⁰ Cf. l'article sur les archives ouvertes.

¹¹ Cf. l'article sur les thèses

EThOS au Royaume Uni, Hellenic Dissertations en Grèce ou DART-Europe pour plus de 550 universités de l'UE. Le portail de l'ETH Zurich référence plus de 2 millions de rapports d'autres bibliothèques, bases de données et outils de recherche tandis que le portail DiVA¹² donne accès aux travaux des chercheurs et étudiants de 27 universités suédoises et norvégiennes dont près de la moitié des 300 000 documents relève de la littérature grise (thèses et mémoires, conférences, rapports, etc.).

Mais toutes ces réalisations ne concernent que les documents publics, au détriment des rapports et autres travaux produits par le secteur privé. Aussi, de nouvelles formes échappent en partie à ces stratégies. Signe que l'idée d'une collection est révolue ? Ou problème de transition, en attendant des outils de veille et d'accès plus puissants ? ■

¹² www.diva-portal.org