

Littérature grise: Typologie et description

Joachim Schöpfel

► **To cite this version:**

Joachim Schöpfel. Littérature grise: Typologie et description. Documentaliste: Sciences de l'Information, 2015, 52 (1), pp.42-44. <hal-01290337>

HAL Id: hal-01290337

<http://hal.univ-lille3.fr/hal-01290337>

Submitted on 18 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Littérature grise : typologie et description

Joachim Schöpfel

Au risque de commencer par une banalité, nous parlons ici de littérature, soit de documents textuels, avec un contenu écrit, ce qui écarte *a priori* d'emblée les vidéos, photos, enregistrements sonores, mais aussi les données, statistiques, tableaux, cartes etc. Quant à la couleur grise, elle renvoie à une zone d'ombre, un champ chromatique sans couleur avec une intensité lumineuse entre le blanc et le noir¹. Autrement dit, nous nous trouvons dans un espace documentaire se situant entre les publications commerciales (blanc) et les documents classés, car confidentiels (noir).

Quels documents ?

Ce préalable s'imposait pour établir une typologie et dresser des listes². Mais toutes ces listes auront deux limites : elles ne seront jamais exhaustives car on verra toujours apparaître de nouveaux types de documents, avec de nouveaux contenus, fonctionnalités et supports ; il restera toujours un point d'interrogation car le même type de document peut être aussi bien gris que blanc ou noir. Ainsi, par exemple, certaines séries de rapports sont commercialisées par des grands groupes d'édition. Il en est de même pour les conférences, notamment quand un éditeur sponsorise ou organise lui-même ces événements, salons, etc. D'autres documents, comme des rapports dans le domaine du renseignement ou de la R&D stratégique sont classés confidentiels et ne circulent pas. Une *lecture note* publiée par un grand éditeur scientifique n'est pas grise. Une revue *samizdat*³ produite en autoédition et diffusée en tout petit nombre l'est. Il faut distinguer diffusion et typologie.

Traditionnellement, le terme « littérature grise » englobe surtout les rapports de tout genre et de toute taille, des travaux universitaires (thèses, mémoires, habilitations etc.), des travaux non publiés (*working papers*, *preprints*) et les conférences. Mais on y trouve aussi des cours d'enseignement et de formation, de la documentation technique, des documents politiques (livres verts), des carnets de laboratoires etc. Par ailleurs, il y a débat sur les nouvelles formes d'édition et d'information, comme les blogs et miniblogs, wikis et réseaux sociaux, un débat ouvert qui défie la documentation, sachant toutefois que les praticiens de la veille ont tranché depuis longtemps et que pour l'intelligence économique et scientifique, il s'agit de sources d'information unique et de valeur. Comme le disait de Pascal Magnier, fondateur et PDG d'Expernova, « nous indexons aujourd'hui près de 30 types de documents : articles passés en comité de relecture, rapports, communications à des conférences, conference proceedings, short communications, posters, thèses, livres, études de validation, essais, etc. mais aussi brevets, projets européens ainsi que les pages web des

¹ Wikipédia <http://fr.wikipedia.org/wiki/Gris>

² La liste de GreyNet <http://www.greynet.org/greysourceindex/documenttypes.html>

³ Cf. la collection unique de la littérature clandestine *drugi obieg* de la Biblioteka Jagiellońska à Cracovie.

organisation que nous cartographions ». Seules comptent la qualité, la pertinence et la fraîcheur de l'information.

Catalogage et indexation

Terminologie et description restent un problème. L'ancienne base européenne SIGLE⁴ et, plus récemment, l'Institute of Computational Linguistics (CNR)⁵ en Italie et, dans le cadre d'un projet de *linked data*, la Bibliothèque Nationale de Technologie à Prague⁶ ont essayé d'harmoniser voire de normaliser certains termes. La Description bibliographique internationale normalisée (ISBD) et UniMARC ignorent ce concept générique mais permettent tout de même une description précise et homogène des conférences, rapports, thèses etc. Mais ce n'est pas le cas des métadonnées du Dublin Core, trop génériques et « permissives » pour ce type de documents.

Pour les thèses, plusieurs recommandations proposent des formats de métadonnées adaptées et détaillées, notamment TEF⁷ en France et ETD-MS⁸ au niveau international. En ce qui concerne les rapports, un groupe de travail international a élaboré un guide pour la mise en forme, la diffusion et la description, y compris les métadonnées⁹. D'une manière plus générale, la norme METS de la Library of Congress¹⁰ contient plusieurs profils de littérature grise, dont les thèses électroniques et les conférences. À l'avenir, il faudra être attentif aux travaux menés autour des fonctionnalités des notices bibliographiques¹¹, notamment en ce qui concerne les attributs des entités Œuvre et Manifestation, y compris pour les données d'autorité (forme, support...), et du nouveau code de catalogage RDA¹².

En matière d'identifiants, certains documents gris ont un ISBN, d'autres un ISSN (série de rapports ou de conférences), voire les deux. Dans les années 90, l'ISO avait lancé le projet d'un identifiant unique pour les rapports (ISRN, ISO 10444), mais la norme a été abandonnée depuis, en raison de son manque de succès et de soutien institutionnel, ce qui veut dire qu'au mieux les rapports ont un identifiant institutionnel, comme auparavant. Quant aux thèses et dissertations, selon les pays, on leur attribue un identifiant unique (numéro, code...) au niveau national ou par établissement. À l'heure actuelle, le DOI est probablement devenu l'identifiant le plus répandu et représentatif de la documentation grise. Tout cela n'est pas très utile pour l'identification, la recherche et la gestion. Être sur le Web ne suffit pas. Cette existence « en *marge des dispositifs de contrôle*

⁴ On en trouve des traces dans OpenGrey <http://www.opengrey.eu/>

⁵ MARZI C., PARDELLI, G., et SASSI, M., A terminology-based re-definition of grey literature. In *Twelfth International Conference on Grey Literature. Prague, 6-7 December 2010* http://www.greynet.org/images/GL12_S1P_Marzi_et_al.pdf

⁶ Cf. <http://invenio.nusl.cz/record/81435>

⁷ Thèses électroniques françaises (TEF) <http://www.abes.fr/abes/documents/tef/>

⁸ Interoperability Metadata Standard for Electronic Theses and Dissertations <http://www.ndltd.org/standards/metadata>

⁹ Cf. article sur GLISC

¹⁰ Metadata Encoding & Transmission Standard <http://www.loc.gov/standards/mets/>

¹¹ Functional requirements for bibliographic records (FRBR) cf.

http://www.bnf.fr/documents/frbr_rapport_final.pdf

¹² Resource Description and Access <http://rda-en-france.enssib.fr/>

bibliographique », pour reprendre les termes de l'ADBS¹³, est sans doute le prix à payer pour une production et diffusion en dehors des circuits commerciaux de l'édition, sans enjeux économiques. De ce point de vue, on n'a pas beaucoup progressé depuis vingt ans¹⁴. L'accès à ces documents n'en devient pas plus simple pour autant.

Et après ?

Il est toujours plus facile de trouver des articles ou livres que des rapports, mémoires ou présentations. Et seule une partie de ces documents sont en libre accès. Pourtant, c'est dans ces documents que l'on trouve souvent davantage d'informations et d'autres types d'informations que dans les articles et les livres. Or, seule une partie de cette information sera un jour publiée sous forme d'article ou de livre¹⁵.

A force d'étudier les différents types et catégories de la littérature grise, avec leurs spécificités et leur potentiel dans l'environnement numérique actuel, la question se pose : pouvons-nous vraiment parler d'une littérature grise, ou est-ce que les particularités des différents types de documents ne sont-elles pas finalement plus importantes que leurs points communs ? Dans certains cas, est-ce que la littérature (soit le document) n'est-elle pas en train de se transformer en service numérique autour d'une information en format texte ? Question ouverte.

¹³ Sources: http://www.adbs.fr/litterature-grise-17647.htm?RH=OUTILS_VOC

¹⁴ Cf. COMBEROUSSE M. (1993). La littérature grise. *Bulletin des Bibliothèques de France*, (5).

¹⁵ Cf. SCHÖPFEL J., Grey literature on bilingualism in Belgium. *Cahiers de la Documentation / Bladen voor Dokumentatie*, 2008, 62(2):22-30.