



HAL
open science

Vers une nouvelle définition de la littérature grise

Joachim Schöpfel

► **To cite this version:**

Joachim Schöpfel. Vers une nouvelle définition de la littérature grise. Cahiers de la Documentation, 2012, 66 (3), pp.14-24. sic_00794984

HAL Id: sic_00794984

https://archivesic.ccsd.cnrs.fr/sic_00794984

Submitted on 26 Feb 2013

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.

Vers une nouvelle définition de la littérature grise

Joachim Schöpfel
Maître de conférences
Univ Lille Nord de France
UDL3, GERIICO
BP 60149
59653 Villeneuve d'Ascq Cedex
France

joachim.schopfel@univ-lille3.fr

Date de fin de rédaction de l'article : 26 novembre 2011

L'article contient une synthèse des travaux sur le concept de la littérature grise, les résultats d'une enquête et une analyse de la définition de référence (définition de New York). Les nouvelles technologies et les nouveaux modes de diffusion sur internet posent un problème pour l'interprétation traditionnelle de la littérature grise. L'article discute trois nouveaux attributs (qualité, propriété intellectuelle, médiation) et propose une nouvelle définition (définition de Prague) qui ajoute ces nouveaux attributs à l'approche économique de l'ancienne définition. D'après cette nouvelle définition, la littérature grise correspond à tout type de document produit par le gouvernement, l'administration, l'enseignement et la recherche, le commerce et l'industrie, en format papier ou numérique, protégé par les droits de propriété intellectuelle, de qualité suffisante pour être collecté et conservé par une bibliothèque ou une archive institutionnelle, et qui n'est pas contrôlé par l'édition commerciale.

Article rédigé suite à la conférence donnée par l'auteur dans le cadre du colloque GL12, Twelfth International Conference on Grey Literature: Transparency in Grey Literature. Grey Tech Approaches to High Tech Issues, organisé par le Grey Literature Network Service, les 6-7 décembre 2010, à Prague.

Introduction

La littérature grise est omniprésente. Selon les disciplines, elle représente jusqu'à 30% de la production scientifique. Les analyses scientométriques montrent son impact au niveau des citations, et elle fait partie des collections, portails et archives ouvertes (Farace & Schöpfel, 2010).

Cependant, le terme de "littérature grise" reste parfois mal défini, imprécis, avec des contours flous. Son concept est historique. Il fut un temps où l'on ne connaissait pas de littérature grise. Le concept a été propagé à partir des années 80, notamment par Auger (1989) dans un livre fondateur sur ces documents souvent riches d'information détaillée et fraîche mais difficiles à acquérir et à gérer. Considérés comme non- ou semi-publiés, ces documents sont diffusés en dehors du marché de l'édition, souvent en petit nombre, et leur catalogage n'est pas facile à cause de leur spécificité.

Auger parlait en premier lieu de rapports techniques et scientifiques. D'autres auteurs ont essayé de préciser ce concept et de l'étendre comme terme générique à d'autres documents possédant les mêmes caractéristiques – surtout des thèses mais aussi des communications, des documents de travail, manuscrits (pré-prints) etc.

La plus connue de ces définitions est celle dite "*de Luxembourg*" approuvée lors de la troisième conférence internationale sur la littérature grise en 1997 : "*(La littérature grise est) ce qui est produit par toutes les instances du gouvernement, de l'enseignement et la recherche publique, du commerce et de l'industrie, sous un format papier ou numérique, et qui n'est pas contrôlé par l'édition commerciale*". Quelques années plus tard, en 2004, la conférence de New York a précisé que cette définition incluait (aussi) les éditeurs "*où la publication ne constitue pas l'activité principale*".¹

Cette définition souligne le côté "offre" de la littérature grise, sa production et diffusion sur papier et en format électronique, avec son caractère éphémère, non conventionnel ou "underground". Un groupe de travail américain a précisé qu'il s'agit de documents qui n'entrent pas dans les systèmes habituels d'édition, de distribution, de contrôle bibliographique ou d'acquisition via les libraires ou agences d'abonnements (U.S. Interagency Gray Literature Working Group) tandis que Mackenzie Owen (1997) observe que "*gris n'implique aucune qualification (mais) est simplement une caractérisation du mode de distribution*".

Aujourd'hui, la définition de New York est généralement acceptée et sert de référence. Cependant, les professionnels et chercheurs partagent majoritairement l'avis qu'elle nécessite une révision.

Où se situe le problème ? La "*définition de New York*" ne décrit pas un ou plusieurs types de documents, mais s'appuie sur un modèle de production et de diffusion. Or, internet a profondément modifié le marché de l'édition, et toute la tectonique de la communication scientifique commence à bouger. Le web offre de nouveaux outils et canaux pour produire, diffuser et évaluer la littérature scientifique. Auteur et lecteur, producteur et consommateur changent leur comportement d'information. Nous avons définitivement quitté l'ère de Gutenberg.

Alors que dire de la définition de la littérature grise ? Est-elle encore valable ? Le web fait-il disparaître la littérature grise ? Pour en savoir davantage, nous avons effectué en 2010 une synthèse des publications sur la littérature grise², une enquête auprès de 108 professionnels et chercheurs du réseau international Greynet³, et une étude analytique sur la définition de New York. Les résultats ont été présentés fin 2010 à Prague, accompagnés de la proposition d'une nouvelle "*définition de Prague*" (Schöpfel, 2010). Voici l'essentiel de notre proposition, suivi d'une discussion de cette nouvelle définition sous l'aspect juridique et par rapport au libre accès.

Les éléments d'un concept

"*La littérature grise est difficile à définir*" (Wood & Smith, 1993). C'est la raison pour laquelle beaucoup d'études commencent par une explication ou une interprétation de la définition de référence. Plusieurs travaux ont été entièrement consacrés à la terminologie et au concept de la littérature grise, comme par exemple Di Cesare & Sala (1995), McDermott (1995), Gokhale (1997) et Nahotko (2007). Quels sont les points communs ?

La plupart mettent en avant la **diffusion**, c'est-à-dire, le mode non conventionnel ou non commercial de la distribution sur le marché de l'information. Par exemple, Mackenzie Owen (1997) définit la littérature grise comme "*l'information distribuée directement par son créateur*". Gelfand (1999) souligne la "*voie alternative de la distribution*" et Boekhorst et al.

¹ Le texte original de cette définition est "*(Grey literature is) that which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers i.e. where publishing is not the primary activity of the producing body*" (Schöpfel & Farace, 2010).

² En tout, nous avons analysé 320 communications et articles publiés entre 1993 et 2010 dans la revue *The Grey Journal* et dans les actes des conférences sur la littérature grise cf <http://www.opengrey.eu/>

³ <http://www.greynet.org/>

(2004) insistent sur la dichotomie "*gris vs. commercial*" comme outil cognitif pour mieux comprendre ce type de littérature scientifique. Parfois, les auteurs ajoutent un autre attribut : le nombre limité des exemplaires diffusés (Aceti et al, 1999).

D'autres auteurs mettent l'accent sur la **production**. De Blaaij (2003) par exemple considère la littérature grise comme l'information "*produite dans le domaine public et financé avec l'argent public*". Dix ans plus tôt, Chillag (1993) fait une distinction entre publications et documents. Pour lui, "*en théorie, et généralement parlant, les premières ne font pas partie de la littérature grise*". D'après Chillag, un rapport devient "blanc" en cas d'acquisition (collecte) et vente. Quant aux documents avec plusieurs versions, aux "working papers" ou tout autre document non signalé (catalogues, bases de données etc.), il les voit comme "black hole material" ("matière de trou noir"). Dans la même année, Cotter & Carroll (1993) déclarent que la littérature grise n'est pas publiée par des éditeurs établis, anticipant les définitions économiques de Luxembourg et New York.

Environ 40% des études adoptent **une approche typologique**. Dans une telle définition, la question clé est de savoir quel type de document appartient à la littérature grise. Les bibliothécaires sont généralement d'accord pour dire que les thèses et mémoires, actes de conférences, rapports et documents de travail sont gris. Mais quid des brevets et prépublications, blogs, bases de données et autres tweets ? Selon Kufa (1993), la littérature grise comprend des items non-conventionnels, "fugitifs", inédits, "non-book" etc. Luzi (1995), Luzi et al. (2003) et Ranger (2004) ont travaillé sur de nouvelles formes d'information scientifique, telles que des conférences électroniques, protocoles, sites web ou "datasets". Stock & Schöpfel (2008) ont évalué la présence de plusieurs types de documents traditionnels - thèses, rapports, documents de travail etc. - dans des archives ouvertes. Sulouff et al. (2005) comparent différents types de littérature grise en fonction de leur discipline scientifique. L'**acquisition** spécifique préoccupe un tiers des études. McDermott (1995) : "*Vous savez que vous avez de la littérature grise lorsque vous ne pouvez pas placer un ordre permanent*". Nahotko (2007) exprime l'opinion majoritaire : "*Ils sont difficiles à acquérir par les bibliothèques*". Internet change-t-il la situation? Pour Natarajan (2006), ce n'est pas le cas ; il considère la littérature grise comme "*le web gris ou caché, l'information qui n'est pas consultable ou accessible par les moteurs de recherche conventionnels ou les annuaires thématiques*".

Moins d'auteurs soulèvent la question de la **qualité**, et ils le font d'une façon controversée. Erwin (2006) observe que la qualité (de la littérature grise) reste suspecte, même pour les chercheurs, et elle n'est pas surprise qu'en raison de sa qualité variable la littérature grise continue à être absente de la plupart des politiques de développement des collections scientifiques. A l'opposé et à partir d'une analyse de la base de données SIGLE, Wessels (1997) affirme que "*beaucoup de littérature grise est publiée par des établissements prestigieux dont les noms sont une garantie de qualité*", et il insiste sur le caractère unique de cette littérature.

En marge de ces définitions, on trouve d'autres remarques et observations intéressantes : D'après de Blaaij (1999), la littérature grise peut faciliter le partage de l'information dans le domaine public car son **statut juridique** est différent de l'édition commerciale. Cornish (1999) suggère un dispositif international (= en fait, une licence) qui signale pour chaque document gris l'utilisation autorisée par le titulaire des droits d'auteur. Pavlov (2003) décrit un modèle russe pour l'attribution et la gestion des droits d'auteur, dans le cadre d'un dispositif de signalement et de commercialisation de thèses scientifiques.

Jeffery & Asserson (2006) définissent les rapports, thèses etc. dans le cadre d'un **système d'information de recherche**⁴ comme des "*objets gris intelligents et hyperactifs*", dotés de métadonnées qui déterminent leur traitement et utilisation.

Crowe & Davidson (2008) placent la littérature grise dans l'intersection de l'**open source** et de l'intelligence économique et ils décrivent comment des documents librement et légalement accessibles peuvent devenir de l'information sensible et source de renseignement.

Pour finir, posons la question de la fonction ou de la finalité de la littérature grise. Pourquoi existe-t-elle ? A quoi sert-elle ? Nous avons comparé les définitions de la littérature grise avec les fonctions classiques de l'édition scientifique, telles qu'elles ont été décrites par Henry Oldenbourg (cf. Mabe, 2009). Le résultat n'est pas vraiment surprenant. La plupart des documents soulignent son rôle dans la diffusion des savoirs, des résultats de la recherche etc. D'autres fonctions sont considérées comme moins importantes comme par exemple l'enregistrement pour établir la propriété intellectuelle, la conservation pérenne ou encore le contrôle de qualité par les pairs.

Kufa (1993) a identifié un rôle intéressant et spécifique pour la littérature grise dans les pays africains dans la mesure où elle peut être plus pertinente et adaptée aux besoins et conditions locales. D'après Kufa, une production scientifique non-commerciale peut être utile pour la communication et la diffusion des contenus à caractère local qui ne sont pas publiés ailleurs. Cette observation rejoint notre propre constat concernant le rôle de la littérature grise pour les communautés linguistiques régionales (Schöpfel, 2008).

Les problèmes de la définition actuelle

Que pensent les experts au sujet de la définition actuelle de la littérature grise, celle dite de New York ? Est-elle utile ou pas ? Devrait-elle être révisée ? Voici quelques éléments issus de notre enquête.

L'opinion générale est globalement positive. Plus de 60% des réponses disent qu'elle reste utile et pertinente. Toutefois, seul un tiers des personnes interrogées croit que la définition actuelle est précise et adaptée aux nouvelles technologies, et seulement 11% pensent qu'elle n'a pas besoin de révision. L'avis partagé est "*révision oui, mais pas d'abandon de la définition actuelle*", c'est-à-dire, pour un changement dans la continuité mais pas de rupture. L'enquête confirme le mode de diffusion et de production (= l'économie documentaire) comme attributs-clés de la littérature grise, suivis par les questions de qualité et d'acquisition. Plus surprenant : seulement 37% considèrent l'approche typologique ("*la littérature grise, ce sont les thèses, rapports etc.*") comme pertinent pour la future définition tandis que 23% pensent le contraire.

Une bonne définition devrait indiquer les conditions nécessaires et suffisantes qu'un document doit remplir pour appartenir à la littérature grise. Elle devrait proposer une conjonction de type "si-et-seulement-si" qui correspond à une relation d'équivalence entre les deux éléments.

Notre enquête et l'état de l'art révèlent que la définition actuelle de la littérature grise est considérée comme utile et pertinente, et elle semble globalement acceptée comme référence. Alors pourquoi dans le même temps notre étude reflète aussi une conviction générale que cette définition doit être révisée ? La raison en est que la définition souffre de plusieurs points faibles. Les voici.

Les impasses de la définition de New York

Suite à l'analyse des publications et à l'enquête, nous pouvons identifier quatre problèmes de la définition de New York.

⁴ Current Research Information System, cf. <http://www.eurocris.org>

La définition de New York est trop large. Elle ne permet pas de distinguer clairement entre la littérature grise et d'autres "objets", et elle s'applique à des éléments qui ne font pas partie de l'extension de la définition. Trop d'études n'arrivent pas à faire la différence entre littérature grise et d'autres types de littérature et/ou d'objets qui sont peut-être gris mais qui n'appartiennent sûrement pas à la littérature.⁵

La définition de New York manque d'attributs essentiels de la littérature grise. La définition met en évidence l'aspect économique spécifique de la littérature grise, le mode de diffusion et de production. Elle échoue cependant à introduire d'autres attributs qui, depuis la première édition du livre d'Auger en 1975, sont devenus essentiels pour la compréhension de la littérature grise, tels que la nature documentaire, la propriété intellectuelle, la qualité, la collecte, la conservation, etc.

La définition de New York contient un risque de circularité. Cela signifie qu'elle présuppose au moins partiellement une connaissance a priori de ce que la littérature grise représente. Par exemple, elle favorise des interprétations telles que "*gris est ce qui n'est pas blanc (ni noir)*".⁶ Cette circularité peut être utile pour la cohésion et l'acceptation communautaire mais ne facilite pas la compréhension.

La définition de New York porte aussi un risque d'obscurité. Elle permet des interprétations avec des termes ambigus tels que "fugitif", "éphémère" etc. ou avec des concepts négatifs ou d'exclusion.⁷

Pour résumer : la définition de New York a été et reste utile pour la compréhension économique de la production et la diffusion de la littérature grise mais elle est d'une utilité limitée dans le contexte des nouvelles technologies et nouveaux usages car elle ne fait pas de distinction entre la littérature grise publiée en ligne et d'autres "digital objects" à caractère non-commercial.

Par la suite, nous allons décrire plusieurs aspects qui devraient faire partie d'une nouvelle définition de la littérature.

La littérature grise comme médium dans le contexte : la propriété intellectuelle

Depuis le début, la question de la typologie a mis la communauté grise dans l'embarras. Avec le développement d'internet, le problème est devenu encore plus compliqué. En appliquant strictement la définition de New York qui est centrée sur la production et la diffusion, la plupart des "items" sur le web pourraient être décrits comme grise - publiés en nombre limité (ou uniques), mal contrôlés, avec une diffusion non commerciale (ou non-lucrative). Or, est-ce réaliste ? D'autant que, à certains égards internet rend les contenus plus largement accessibles que les documents imprimés par les maisons d'édition.

Les objets numériques tels que les ensembles de données, courriels, copies techniques, wikis, blogs, tweets ou même des sites web etc., sont-ils gris ? Relèvent-ils au moins de la littérature ?

Lors de notre enquête, certains experts ont plaidé en faveur d'un large concept "fourre-tout" : "*Définir la littérature grise en opposition à l'édition commerciale (à but lucratif) est une erreur (...) car de nombreux organismes sans but lucratif publient sur exactement la même base que les éditeurs commerciaux (par exemple l'American Chemical Society) - et leurs*

⁵ Avec les paroles d'une personne interrogée : "*Attention - maintenant que la littérature est devenue «sexy», les gens vont trop loin. Les données brutes ne sont pas de la littérature grise. Ephemera non plus. L'ensemble du Web non plus. Je pense que si on va trop loin on risque de perdre toute crédibilité (...)*".

⁶ "*On doit définir la littérature grise dans sa relation aux publications officielles ou blanches (...)*" (Artus, 2003). Mais alors, qu'est-ce qui est blanc, et où commence ce blanc ?

⁷ Un exemple de l'enquête : "*La littérature grise est l'absence d'éléments clés, tels que l'apparition dans la littérature traditionnelle. La littérature grise est difficile à trouver, difficile à classer, difficile à acquérir, etc. Elle se trouve dans des zones périphériques (petites revues, des archives isolées, des formats inhabituels).*"

publications ne sont pas grises. À mon avis, la littérature grise est tout travail scientifique publié sans 'peer review' (ou équivalent) en dehors de la revue traditionnelle et de la filière du livre. Par conséquent, la littérature grise comprend des 'datasets', outils d'analyse, des logiciels ainsi que les documents plus courants comme le préprint et le document de travail."

Il paraît douteux que les "datasets", outils d'analyse ou logiciels doivent ou peuvent être considérés comme travail scientifique en termes de littérature. Peut-être nous devrions envisager la question sous un autre angle. Dans l'ère post-Gutenberg, le réseau français interdisciplinaire *RTP-DOC* a élaboré une théorie en 3D décrivant des documents numériques en tant que signe (le contenu), la forme (structure) et le médium (contexte) (Pédauque, 2003). Selon l'approche *RTP-DOC*, les anciens critères et attributs liés à l'impression ne peuvent pas être transposés à des objets numériques.

Définir la littérature grise uniquement comme un contenu ou une structure spécifique (par exemple, comme type de document) semble impossible et voué à l'échec. Par contre, le concept apparaît clairement comme une valeur sur la troisième dimension, le médium, et devrait être défini dans son contexte, lié à des attitudes, la compréhension et prise de conscience, les normes sociales, et par rapport aux utilisateurs, à l'utilisation et aux procédures.

Retenons que dans un premier temps, le concept de littérature grise devrait être limité à la signification spécifique de littérature, non pas comme un contenu ou une structure/type mais dérivé de sa nature sociale ou morale : un objet numérique fait partie de la littérature grise si et seulement s'il est un objet protégé par des droits de propriété intellectuelle. En d'autres termes, la littérature grise implique la paternité et un caractère d'œuvre de l'esprit.

Les droits d'auteur peuvent être détenus par l'organisme producteur et/ou par le ou les auteur(s) mais en aucun cas par un fournisseur commercial. Récemment, Kansa et al. (2010) a soulevé la question des licences (Creative Commons etc.). Cela peut être une option adéquate pour les documents gris diffusés sur le web par leurs producteurs et/ou auteurs, mais n'est en aucun cas une condition suffisante pour définir la littérature grise.

En revanche, une définition fonctionnelle ne paraît pas pertinente. Nos résultats indiquent que la littérature grise ne remplit pas d'autres fonctions que les revues ou les livres du commerce mais se limite plutôt à certaines fonctions telles que la diffusion et la préservation. Comme attribut, ce n'est pas suffisant pour différencier les documents gris d'autres documents.

De la typologie à la qualité

Certains auteurs ont tendance à définir la littérature grise à travers une liste des différents types de documents. Une telle définition en extension ou par dénotation est vouée à l'échec, pour au moins deux raisons : (a) il y a trop de catégories différentes de la littérature grise, en particulier dans un format numérique, et (b) une telle définition ne pourra pas faire face à l'évolution de nouveaux types et formats.

Aussi, le même type de document peut être gris ou blanc, à différents moments de son cycle de vie et en fonction du contexte. Par exemple, les thèses de doctorats sont-elles grises ? Sont-elles toujours grises ? Un mémoire de Master est-il gris ? Qu'en est-il des mémoires en premier cycle (licence) ? Quelle est la différence ? Certains rapports sont commercialisés par leurs producteurs et/ou via les filières de distribution habituelles (librairies, grossistes...), certains documents gris sont des rapports. Comment un rapport devient-il gris ?

Notre explication est la qualité et la valeur scientifique et/ou technique du document en question. Complémentaire au mode de diffusion et à la propriété intellectuelle, ce critère permet de distinguer la littérature grise des documents "gris foncés" sans intérêt pour la recherche et l'enseignement. En effet, alors que la diffusion et la production sont des attributs essentiels pour déterminer la ligne de partage entre les documents commerciaux (blancs) et gris, la propriété intellectuelle et la qualité sont utiles pour le "côté obscur" de la littérature

grise, par exemple pour faire une distinction nette entre les travaux du premier cycle universitaire et les mémoires de Master ou thèses de doctorat, ou entre un document de travail ("working paper") et un manuscrit dans une version préliminaire, provisoire et transitoire. La notion de qualité implique l'assurance et le contrôle de qualité et une valeur scientifique intrinsèque. Une étude récente constate que l'idée qu'il n'existe pas de "peer review" pour les rapports est fautive, et l'auteur observe qu'un rapport subit souvent plusieurs niveaux de contrôle et de validation interne avant d'être diffusé dans sa version finale (Seymour, 2010). En 2005, l'initiative GLISC⁸ ("groupe de Nancy") a établi un label de qualité pour les rapports. D'autres documents gris comme les thèses, mémoires ou communications sont soumis à des procédures de "labelling" et/ou de validation qui garantissent un niveau minimum de qualité. D'après nos propres études sur la littérature grise en France et en Belgique, ces procédures concernent jusqu'à 60% de la littérature grise.

La qualité devient une question cruciale de la littérature grise. Dans notre enquête, plus de la moitié des personnes interrogées considèrent la qualité comme essentielle dans la définition de la littérature grise. Farace (2010) compare les procédures de "peer review" pour la littérature grise et pour les publications en série. En termes de gestion de la qualité, certaines procédures sont plutôt conçues pour détecter la non-qualité (par exemple, l'évaluation et l'examen des thèses de doctorat) tandis que d'autres sont censées empêcher une non-conformité ou une qualité non acceptable (comme les recommandations GLISC).

Cela étant, la qualité n'est pas un absolu, une valeur en soi mais elle est liée aux normes, à l'utilisation ou au but spécifique, et elle est basée sur l'évaluation et/ou l'expérience d'un "client". Cela nous amène au dernier attribut essentiel de la littérature grise.

La littérature grise sans bibliothèque ?

Qui est le "client" de la littérature grise ? Le plus souvent, les internautes ne sont pas conscients des catégories particulières de produits et canaux de diffusion. Leur problème est l'accès, pas la production, distribution ou collecte. Les documents non publiés ou semi-publiés ne signifient rien pour eux. Gris ou blanc leur importe peu ; par contre, ils n'aiment pas la littérature "trou noir", c'est-à-dire les documents qu'ils ne peuvent pas obtenir.

En fait, depuis l'invention du concept, le vrai client de la littérature grise est et a toujours été le bibliothécaire, le professionnel de l'information qui essaie de l'acquérir. Quand on relit les articles et communications sur la littérature grise des deux dernières décennies, cet aspect saute aux yeux. "*Chercher et obtenir la littérature grise (...) a toujours été un grand défi pour les bibliothécaires et documentalistes*" (Natarajan, 2006). Le point crucial est qu'il ne s'agit pas d'un problème pour les chercheurs, étudiants ou autres lecteurs mais uniquement pour les professionnels.

En d'autres termes, le problème avec la littérature grise existerait-il toujours si les bibliothécaires arrêtaient d'essayer de l'acquérir ? Un "working paper" serait-il gris si personne ne s'y intéressait ? Une pré-publication envoyée par son auteur à un collègue est-elle grise ? Sans doute, il s'agit là d'une mauvaise question. Voici une meilleure question : une bibliothèque voudrait-elle acquérir cette pré-publication pour ses collections, en raison de sa qualité, de sa valeur scientifique, de sa rareté et/ou de son caractère patrimonial ? Une autre : est-ce qu'un document obtient le label "gris" parce qu'il a été produit et validé par une institution ?

La définition de la littérature grise est intimement conditionnée par le fait que c'est un objet de collection et d'acquisition. Un document devient gris, non seulement parce qu'il est une œuvre de l'esprit et non commercialisé, mais dans la mesure où quelqu'un - une institution, une bibliothèque, un service d'information, un professionnel - montre de l'intérêt pour l'obtenir. Il

⁸ <http://www.glisc.info>

devient gris car il est censé avoir de la qualité, de la valeur scientifique, de l'originalité ou un caractère patrimonial.

Cela nous aide à clarifier deux aspects. Tout d'abord, "collection" ne signifie pas "stockage". Une collection est synonyme de politique de sélection et d'acquisition, est dynamique, tient compte de l'utilisation et reflète des objectifs culturels ou besoins de la communauté. Un document "difficile à obtenir" devient gris quand il est jugé utile (ou supposé être utile à l'avenir) pour un scientifique, une équipe de recherche, un laboratoire, une institution ou une communauté.

Deuxièmement, la littérature grise n'est pas cachée, invisible ou confidentielle, mais "open source". Les documents qui ne peuvent être collectés ne sont pas gris. Encore une fois, le point crucial n'est pas l'invisibilité ou la confidentialité, mais la collecte, la mise à disposition et la préservation pour l'avenir.

Aujourd'hui, les archives institutionnelles ont commencé à prendre en charge certaines des fonctions traditionnelles des collections de la bibliothèque. En termes de fonction, elles présentent une certaine équivalence avec la littérature grise elle-même dans la mesure où leur rôle principal consiste dans la diffusion et, dans une moindre mesure, la préservation. Cela peut expliquer l'intérêt croissant de la communauté de la littérature grise pour ce type d'archives ouvertes (cf. Luzi, 2010). Une nouvelle définition devrait ou pourrait le mentionner.

L'une des questions de notre enquête était : "*pouvez-vous imaginer un monde sans littérature grise ?*" La plupart des experts interrogés ne peuvent pas. Mais pouvons-nous imaginer la littérature grise sans professionnels de l'information ? Non. Le défi de la littérature grise est un défi de médiation. La littérature grise n'existe pas dans un espace vide. "*Grey lies in the eye of the beholder*".⁹ Elle est "medium" au sens de *RTP-DOC*.

Proposition d'une nouvelle définition

La définition actuelle de la littérature grise - la définition de New York - demeure utile et ne doit pas être remplacée mais adaptée au nouveau contexte.

Comme nous l'avons montré ci-dessus, l'approche typologique ne fournit pas une liste exhaustive et explicite des documents. L'approche économique de la définition de New York, d'autre part, est une définition intensionnelle qui précise la condition nécessaire pour un document faisant partie de la littérature grise. Mais la même définition n'est pas suffisante dans le contexte de l'édition sur internet, et nous avons besoin d'autres attributs pour différencier clairement la littérature grise d'autres types de documents.

Notre proposition est d'ajouter quatre attributs à la définition de New York :

1. La nature documentaire de la littérature grise (dans le sens du concept de RTP-DOC).
2. Sa nature juridique en tant qu'œuvre de l'esprit, c'est-à-dire, la protection par la propriété intellectuelle.
3. Un niveau minimum de qualité (contrôle par les pairs, label, validation).
4. La médiation, c'est-à-dire l'intérêt des documents gris pour une collection (et non pas pour l'utilisateur final).

Voici notre proposition pour une nouvelle définition de la littérature grise ("*définition de Prague*") :

"La littérature grise correspond à tout type de document produit par le gouvernement, l'administration, l'enseignement et la recherche, le commerce et l'industrie, en format papier ou numérique, protégé par les droits de propriété intellectuelle, de qualité suffisante pour être

⁹ En adaptant un proverbe anglais : Le gris est dans l'œil de celui qui regarde...

collecté et conservé par une bibliothèque ou une archive institutionnelle, et qui n'est pas contrôlé par l'édition commerciale."¹⁰

La littérature grise comprend toutes sortes de matériel documentaire de qualité et d'impact qu'une bibliothèque souhaiterait collecter mais qui lui posent un problème à cause des modes de distribution non-conventionnels. Elle n'est pas seulement une question de production et de diffusion, mais aussi de qualité et de collection. Sans médiation par les bibliothèques, pas de littérature grise. C'est un cas pour les professionnels de l'information, un défi qui rassemble les communautés de la littérature grise et des collections spéciales.

Il y quelques années, le principal problème avec la littérature grise semblait être l'économie. Ainsi, Simpson (1995) observait que le "matériel périphérique", y compris la littérature grise, augmentait sans cesse. Pour lui, les bibliothèques ayant des difficultés à acquérir les documents habituels avaient peu d'espoir d'acquérir la périphérie.

Aujourd'hui, en raison de l'immense succès de l'édition en ligne et de l'accès aux documents, le débat s'est déplacé vers la qualité, la propriété intellectuelle et la médiation. Sans une révision qui inclut ces attributs, la définition de New York risque de devenir de plus en plus incapable de différencier la littérature grise d'autres types de documents.

Notre proposition d'une nouvelle "*définition de Prague*" ajoute les nouveaux attributs à l'ancienne approche économique. La prochaine étape devrait être de tester cette définition par rapport au terrain (= bibliothèques) et aux différents types de documents (gris et autres).

Ensuite, la valeur de la définition pourrait être évaluée sur la base des réponses à deux questions :

1. La nouvelle définition inclut-elle tout type de documents généralement considéré par les professionnels de l'information comme littérature grise, y compris les objets qui aujourd'hui sont difficiles à acquérir et à traiter ?
2. Est-ce qu'elle contribue à une meilleure différenciation ou à une meilleure compréhension de la littérature grise par rapport à d'autres formes de littérature ?

Une attention particulière devrait être accordée à trois grands défis auxquels nous sommes confrontés au début du 21^e siècle.

(1) Le développement des archives institutionnelles comme un service complémentaire et parfois en concurrence avec les collections traditionnelles de la bibliothèque, et la place et le traitement de la littérature grise dans ces archives.

(2) La tendance à la désintermédiation de la chaîne de valeur traditionnelle de l'information scientifique et technique. Le risque de la littérature grise n'est pas la technologie web, mais l'affaiblissement du rôle de médiation des bibliothèques et des professionnels de l'information entre les auteurs, les éditeurs et producteurs d'information, et l'utilisateur final.

(3) Le "quatrième paradigme" (Hey et al., 2009), c'est-à-dire l'usage intensif de données scientifiques et l'accès massif aux "datasets" qui génèrent ensemble une tendance à transformer et/ou marginaliser la littérature (documents).

Par rapport à la littérature grise, plusieurs personnes interrogées ont répondu à notre enquête qu'elle était importante pour la connaissance, qu'il s'agissait d'une question de liberté ou d'édition alternative. L'avenir nous dira si notre conception de la littérature grise demeure "éphémère" ou si elle contribue à une meilleure compréhension et un meilleur traitement de cette catégorie si particulière de l'information scientifique et technique.

¹⁰ En version originale : "*Grey literature stands for manifold document types produced on all levels of government, academics, business and industry in print and electronic formats that are protected by intellectual property rights, of sufficient quality to be collected and preserved by library holdings or institutional repositories, but not controlled by commercial publishers i.e., where publishing is not the primary activity of the producing body.*"

Addendum

La communication de cette proposition à la bibliothèque nationale technique à Prague en décembre 2010 a suscité deux réactions, l'une par rapport au rôle des archives institutionnelles, l'autre en lien avec la propriété intellectuelle.

Concernant les **archives institutionnelles**, la première réaction provenait de Claudia Marzi du Conseil National de la Recherche (CNR) italien qui avait présenté à la même conférence une analyse terminologique du concept de la littérature grise (Marzi et al., 2010). L'approche italienne met en avant le format numérique et l'impact du libre accès pour la définition de la littérature grise. Cette définition est en même temps plus large (elle parle d'information et de connaissance), plus restrictive (elle limite la définition aux seules archives institutionnelles et à l'information numérique), et plus prescriptive. Néanmoins, une ré-analyse des deux conceptions révèle leur convergence : le caractère documentaire de la littérature grise, la nature "protégeable" des documents gris, l'importance de leur qualité, la compréhension de ces documents comme "médium". Le débat a été poursuivi dans le groupe GreyNet du réseau social LinkedIn. Il n'a pas abouti à une définition de convergence ou compromis mais s'est déplacé à la question de savoir si les nouvelles formes numériques d'information scientifiques ("datasets" etc.) relèvent de la littérature grise ou pas.

Propriété intellectuelle : La deuxième réaction, formulée par Blane K. Dessy de la Library of Congress, était de faire remarquer que la protection d'une œuvre de l'esprit par le droit d'auteur était une conception continentale (française) et n'avait pas de sens ailleurs, notamment aux Etats-Unis. Entretemps, ce débat a évolué vers une analyse comparative des aspects juridiques de la littérature grise dont les résultats seront publiés en 2012¹¹. Deux constats préliminaires : les différences entre les deux systèmes semblent moins significatives qu'au premier vu ; en fait, la question juridique est surtout liée à la divulgation et diffusion du document. Et il paraît plus approprié sous l'aspect juridique d'appliquer une approche différentielle et de distinguer plusieurs cas de figures à la place de parler de la littérature grise en général.

Mes remerciements à tous les collègues qui ont contribué à la synthèse documentaire et à l'enquête, et tout particulièrement à Dominic Farace, Leonid P. Pavlov, Deni J. Seymour et Hélène Prost pour leurs conseils utiles.

Bibliographie

- Aceti, A., Castriotta, M., Di Cesare, R., Luzi, D. A profile of GL producers in the field of safety and health at workplaces in Italy: results of a sample survey. In: *Fourth International Conference on Grey Literature: New Frontiers in Grey Literature*. Washington, 4-5 October 1999. <<http://opensigle.inist.fr/handle/10068/697892>>
- Artus, H. M. Old WWWine in new bottles? Developments in electronic information and communication: structural change and functional inertia. In: *Fifth International Conference on Grey Literature: Grey Matters in the World of Networked Information*. Amsterdam, 4-5 December 2003.
- Auger, Charles P. *Use of Reports Literature*. Butterworths, 1975.
- Auger, Charles P. *Information Sources in Grey Literature*. Second edition. Bowker Saur, 1989.
- Boekhorst, A. K., Farace, D. J., Frantzen, J. Grey literature survey 2004 - a research project tracking developments in the field of grey literature. In: *Sixth International Conference on*

¹¹ Une première communication a été faite à Washington en décembre 2011 (Braun et al., 2011).

- Grey Literature: Work on Grey in Progress*. New York, 6-7 December 2004.
<<http://opensigle.inist.fr/handle/10068/697831>>
- Braun, M., Dessy, B. K., Schöpfel, J. Legal aspects of grey literature: a discussion. In: *Thirteenth International Conference on Grey Literature: The Grey Circuit. From Social Networking to Wealth Creation*. Washington D.C., 5-6 December 2011.
- Chillag, J. P. From Weimar to Maastricht and Beyond: Half a Century with Grey Literature. In: *First International Conference on Grey Literature: Weinberg Report 2000*. Amsterdam, 13-15 December 1993. <<http://hdl.handle.net/10068/698060>>
- Cornish, G. P. Copyright: Black and white or just making you see red? In: *Fourth International Conference on Grey Literature: New Frontiers in Grey Literature*. Washington, 4-5 October 1999. <<http://hdl.handle.net/10068/697897>>
- Cotter, G. A., Carroll, B. C. A new generation of grey literature: the impact of advanced information technologies. In: *First International Conference on Grey Literature: Weinberg Report 2000*. Amsterdam, 13-15 December 1993.
<<http://opensigle.inist.fr/handle/10068/698059>>
- Crowe, J., Davidson, T. S. The 'grey' intersection of open source information and intelligence. In: *Tenth International Conference on Grey Literature: Designing the Grey Grid for Information Society*. Amsterdam, 8-9 December 2008.
<<http://opensigle.inist.fr/handle/10068/697994>>
- de Blaaij, C.. Intellectual property on the move. In: *Fourth International Conference on Grey Literature: New Frontiers in Grey Literature*. Washington, 4-5 October 1999.
<<http://hdl.handle.net/10068/697895>>
- de Blaaij, C. Grey literature from invisibility to visibility: The quest for grey content in the domain of the invisible web. In: *Fifth International Conference on Grey Literature: Grey Matters in the World of Networked Information*. Amsterdam, 4-5 December 2003.
<<http://opensigle.inist.fr/handle/10068/697856>>
- Di Cesare, R., Sala, C. The use of grey literature in the agricultural economics field: a quantitative analysis. In: *Second International Conference on Grey Literature: Grey Exploitations in the 21st Century*. Washington, 2-3 November 1995.
<<http://opensigle.inist.fr/handle/10068/698018>>
- Erwin, P. A. The messy world of grey literature in cyber security. In: *Eighth International Conference on Grey Literature: Harnessing the Power of Grey*. New Orleans, 4-5 December 2006. <<http://opensigle.inist.fr/handle/10068/697767>>
- Farace, D. Peering through the review process: Towards transparency in grey literature. In: *Twelfth International Conference on Grey Literature: Transparency in Grey Literature. Grey Tech Approaches to High Tech Issues*. Prague, 6-7 December 2010.
- Farace, Dominic J.; Joachim Schöpfel. *Grey Literature in Library and Information Studies*. De Gruyter Saur, 2010.
- Gelfand, J. Grey literature: Plausible applications for distance education practices. In: *Fourth International Conference on Grey Literature: New Frontiers in Grey Literature*. Washington, 4-5 October 1999. <<http://opensigle.inist.fr/handle/10068/697902>>
- Gokhale, P. Grey literature varieties - definitional problems. In: *Third International Conference on Grey Literature : Perspectives on the Design and Transfer of Scientific and Technical Information*. Luxemburg, 13-14 November 1997.
<<http://opensigle.inist.fr/handle/10068/697961>>
- Hey, Tony; Stewart Tansley; Kristin Tolle. *The fourth paradigm. Data-intensive scientific discovery*. Microsoft Corp., 2009.

- Jeffery, K. G., Asserson, A. Hyperactive grey objects. In: *Eighth International Conference on Grey Literature: Harnessing the Power of Grey*. New Orleans, 4-5 December 2006. <<http://opensigle.inist.fr/handle/10068/697770>>
- Kansa, Eric; Kansa, Sarah; Burton, Margie; Stankowski, Cindy. Googling the Grey: Open Data, Web Services, and Semantics. *Archaeologies*, 2010, vol. 6, no. 2, p. 301-326. <<http://dx.doi.org/10.1007/s11759-010-9146-4>>
- Kufa, J. C. Grey literature in academic research in developing countries: Problems and prospects. In: *First International Conference on Grey Literature: Weinberg Report 2000*. Amsterdam, 13-15 December 1993. <<http://opensigle.inist.fr/handle/10068/698066>>
- Luzi, D. Internet as a new distribution channel of scientific grey literature: the case of Italian WWW servers. In: *Second International Conference on Grey Literature: Grey Exploitations in the 21st Century*. Washington, 2-3 November 1995. <<http://hdl.handle.net/10068/698027>>
- Luzi, Daniela. Grey Documents in Open Archives. *The Grey Journal*, 2010, vol. 6, no. 3, p. 137-144.
- Luzi, D., Castriotta, M., Manco, M., Di Cesare, R., Libutti, L. The communication flow of research projects results. In: *Fifth International Conference on Grey Literature: Grey Matters in the World of Networked Information*. Amsterdam, 4-5 December 2003. <<http://opensigle.inist.fr/handle/10068/697861>>
- Mabe, Michael A. Scholarly Publishing. *European Review*, 2009, vol. 17, no. 1, p. 3-22. <<http://dx.doi.org/10.1017/S1062798709000532>>
- Mackenzie Owen, J. Expanding the horizon of grey literature. In: *Third International Conference on Grey Literature: Perspectives on the Design and Transfer of Scientific and Technical Information*. Luxemburg, 13-14 November 1997. <<http://hdl.handle.net/10068/697967>>
- Marzi, C., Pardelli, G., Sassi, M. A terminology-based re-definition of grey literature. In: *Twelfth International Conference on Grey Literature: Transparency in Grey Literature. Grey Tech Approaches to High Tech Issues*. Prague, 6-7 December 2010.
- McDermott, J. C. Defining grey literature for the 21st century. In: *Second International Conference on Grey Literature: Grey Exploitations in the 21st Century*. Washington, 2-3 November 1995. <<http://opensigle.inist.fr/handle/10068/698028>>
- Nahotko, M. Some types of grey literature: A polish context. In: *Ninth International Conference on Grey Literature: Grey Foundations in Information Landscape*. Antwerp, 10-11 December 2007. <<http://opensigle.inist.fr/handle/10068/697883>>
- Natarajan, Munusamy. Grey literature: problems and prospects for collection development in e-environment. *The Grey Journal*, 2006, vol. 2, no. 2, p. 100-105.
- Pavlov, L. P. The commercialization of research findings documented in grey literature. In: *Fifth International Conference on Grey Literature: Grey Matters in the World of Networked Information*. Amsterdam, 4-5 December 2003. <<http://hdl.handle.net/10068/697864>>
- Pédauque, Roger T. *Document : forme, signe et médium, les re-formulations du numérique*. CNRS (RTP-Doc), 2003. Rapport. <http://archivesic.ccsd.cnrs.fr/docs/00/06/21/99/PDF/sic_00000511.pdf>
- Ranger, S. L. Grey literature in special libraries: Access and use. In: *Sixth International Conference on Grey Literature: Work on Grey in Progress*. New York, 6-7 December 2004. <<http://opensigle.inist.fr/handle/10068/697848>>

- Schöpfel, Joachim. Grey literature on bilingualism in Belgium. *Cahiers de la Documentation / Bladen voor Dokumentatie*, 2008, vol. 62, no. 2, p. 22-30.
- Schöpfel, J. Towards a Prague definition of grey literature. In: *Twelfth International Conference on Grey Literature: Transparency in Grey Literature. Grey Tech Approaches to High Tech Issues*. Prague, 6-7 December 2010.
<<http://invenio.nusl.cz/record/41997?ln=cs>>
- Schöpfel, Joachim; Farace, Dominic J. Grey literature. In Bates, Marcia. J., Maack, Mary N. (ed.) *Encyclopedia of Library and Information Sciences, Third Edition*. CRC Press, 2010, p. 2029–2039.
- Seymour, Deni. Sanctioned Inequity and Accessibility Issues in the Grey Literature in the United States. *Archaeologies*, 2010, vol. 6, no. 2, p. 233-269.
<<http://dx.doi.org/10.1007/s11759-010-9144-6>>
- Simpson, D. B. Grey literature: the challenges for an increasingly important body of research literature. In: *Second International Conference on Grey Literature: Grey Exploitations in the 21st Century*. Washington, 2-3 November 1995. <<http://hdl.handle.net/10068/698033>>
- Stock, C., Schöpfel, J. Grey literature in French digital repositories: a survey. In: *Tenth International Conference on Grey Literature: Designing the Grey Grid for Information Society*. Amsterdam, 8-9 December 2008. <<http://opensigle.inist.fr/handle/10068/698006>>
- Sulouff, Pat; Bell, Suzanne; Briden, Judi; Frontz, Stephanie; Marshall, Ann. Learning about grey literature by interviewing subject librarians. *The Grey Journal*, 2005, vol. 1, no. 3, p. 131-136.
- Wessels, R. H. A. An eagle's eye-view of grey literature research. In: *Third International Conference on Grey Literature: Perspectives on the Design and Transfer of Scientific and Technical Information*. Luxemburg, 13-14 November 1997.
<<http://hdl.handle.net/10068/697979>>
- Wood, David N.; Smith, A. W. SIGLE: A Model for International Co-operation. *Interlending & Document Supply*, 1993, vol. 21, no. 1, p. 18-22.
<<http://dx.doi.org/10.1108/02641619310154331>>