



HAL
open science

Ce que tout chercheur et tout bibliothécaire devraient savoir sur la recherche bibliographique : note de lecture à partir de trois articles

Frédérique Flamerie de Lachapelle

► To cite this version:

Frédérique Flamerie de Lachapelle. Ce que tout chercheur et tout bibliothécaire devraient savoir sur la recherche bibliographique : note de lecture à partir de trois articles. Bulletin des Bibliothèques de France, 2021. hal-03258344

HAL Id: hal-03258344

<https://hal.science/hal-03258344>

Submitted on 11 Jun 2021

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Ce que tout chercheur et tout bibliothécaire devraient savoir sur la recherche bibliographique : note de lecture à partir de trois articles

Frédérique Flamerie de Lachapelle

Conservatrice des bibliothèques à l'université de Bordeaux – Responsable de la bibliothèque de l'Institut de santé publique, d'épidémiologie et de développement (Isped)

Note de lecture des trois articles suivants, par ordre de parution :

Gusenbauer, M., & Haddaway, N. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed and 26 other resources. *Research Synthesis Methods*, 11(2), 181-217. <https://doi.org/10.1002/jrsm.1378>

Klopfenstein, D. V., & Dampier, W. (2021). Commentary to Gusenbauer and Haddaway 2020: Evaluating retrieval qualities of Google Scholar and PubMed. *Research Synthesis Methods*, 12(2), 126-135. <https://doi.org/10.1002/jrsm.1456>

Gusenbauer, M., & Haddaway, N. (2021). What every researcher should know about searching – clarified concepts, search advice, and an agenda to improve finding in academia. *Research Synthesis Methods*, 12(2), 136-147. <https://doi.org/10.1002/jrsm.1457>

La seule lecture des titres et du nom des auteurs suffit à comprendre que ces trois articles sont liés, et ce, à plusieurs titres. Ils partagent en premier lieu un thème central, l'évaluation d'outils de recherche bibliographique, qui ne peut que retenir l'attention du bibliothécaire, à plus forte raison si cette évaluation concerne plus d'une vingtaine de ressources y compris des incontournables telles que Google Scholar et PubMed. On repère en deuxième lieu des relations éditoriales : Klopfenstein et Dampier (2021) commentent un article de recherche de Gusenbauer et Haddaway (2020), ces derniers répondant à leur tour à leurs commentateurs dans un autre commentaire (Gusenbauer & Haddaway, 2021).

Ces seuls liens suffiraient à expliquer le choix d'une analyse groupée.

Une autre raison motive ce choix, celle de pouvoir ainsi proposer un ordre de lecture inverse à celui de la parution des articles. Avec cette lecture à rebours, de (Gusenbauer & Haddaway, 2021) à (Gusenbauer & Haddaway, 2020), on aborde la série d'articles par une clarification des concepts et un cadre de compréhension, éclairant les analyses plus techniques et détaillées présentées dans les deux autres articles.

Si l'on cherche un point de vue général concernant la recherche bibliographique, incluant le rôle des bibliothèques universitaires dans un plan d'action à mener, on pourra

même s'en tenir à la lecture de (Gusenbauer & Haddaway, 2021). Nous espérons toutefois que la présente note incitera à la lecture des trois articles le plus grand nombre de leurs lecteurs potentiels, des bibliothécaires concernés par la recherche d'articles scientifiques dans un contexte universitaire à ceux intéressés plus largement par le rôle des bibliothèques dans le développement des compétences informationnelles, au-delà même du contexte académique.

Que l'on choisisse un ordre de lecture chronologique ou à rebours, le point de départ demeure les deux auteurs N. Haddaway et M. Gusenbauer. Neal Haddaway est membre de l'organe de gouvernance (*board of trustees*) de la *Collaboration for Environmental Evidence* et coordonnateur du centre CEE du Stockholm Environment Institute. Il est auteur non seulement de revues systématiques mais également d'articles de méthode concernant les revues systématiques, et notamment les aspects de recherche bibliographique ; citons à cet égard un article de référence concernant Google Scholar (Haddaway *et al.*, 2015). Une revue systématique peut être définie brièvement comme une synthèse des faits avérés, produite pour répondre à une question de recherche précise et structurée. L'application d'une méthode rigoureusement définie vise à limiter les biais et les erreurs à toutes les étapes, de la recherche bibliographique pour identifier les études pertinentes à l'analyse et à la synthèse de leur contenu.

Michael Gusenbauer combine également l'intérêt pour la recherche fondée sur les faits avérés (*evidence-based*) dans son domaine de recherche, le management de l'innovation, et l'intérêt pour les outils de recherche bibliographique ; citons un article concernant à nouveau Google Scholar (Gusenbauer, 2019).

Gusenbauer & Haddaway, 2021 : clarification des concepts, conseils de recherche et programme d'actions

« One of the most critical factors that can easily limit the quality of our work is the belief that how we search academically is perfectly fine. »

« L'un des facteurs les plus critiques qui peuvent facilement limiter la qualité de notre travail est la conviction que la façon dont nous effectuons nos recherches [bibliographiques] académiques est parfaitement acceptable. »

Cet article définit à la fois un cadre conceptuel et de compréhension de la recherche bibliographique et un plan d'action, mentionnant explicitement les bibliothèques universitaires et le rôle qu'elles pourraient jouer dans les actions auxquelles il appelle.

Les auteurs rappellent tout d'abord et caractérisent les trois types de recherche bibliographique que peuvent avoir à mener les chercheurs : la recherche de vérification (*look-up*), la recherche exploratoire et la recherche systématique. Cette description détaillée permet de comprendre les mécanismes de la recherche bibliographique à des fins de recherche, et met également en lumière les mésusages que des confusions entre ces trois types peuvent susciter. À partir de cette étape de clarification des concepts, les auteurs

définissent un cadre d'analyse et de compréhension, le « triangle de la recherche », dont les trois pointes sont constituées par les objectifs de l'utilisateur, les heuristiques de recherche appropriées (c'est-à-dire les méthodes appliquées) et les systèmes de recherche appropriés (c'est-à-dire les outils interrogés). Le tableau 1 fournit ainsi une vue synthétique et une mise en correspondance de ces trois éléments, en fonction du type de recherche considéré. Les objectifs sont illustrés par des cas d'usage, tandis que les systèmes de recherche sont traduits en caractéristiques techniques requises pour ces systèmes.

Une autre clarification importante concerne la différence entre les deux grands types de système de recherche disponibles, à savoir les bases de données bibliographiques traditionnelles (Web of Science, Scopus) « exhaustives-transparentes » « *comprehensive-transparent* », et les nouveaux moteurs de recherche (Google Scholar, Semantic Scholars) « efficaces-rapides » « *efficient-slick* ». Ces différences sont illustrées dans l'article de Klopfenstein et Dampier (2021) avec la comparaison entre Google Scholar et PubMed. Un risque de mésusage de Google Scholar est par exemple de renforcer le biais de confirmation, en faisant dériver la recherche exploratoire vers une sélection rapide d'articles ou « *cherry-picking* ».

Le second aspect majeur de cet article est l'appel à l'action pour une amélioration de la recherche bibliographique et plus largement documentaire, dans lequel le monde des bibliothèques est explicitement invité à jouer deux rôles. Le premier concerne les bibliothèques universitaires et leur rôle dans la formation à la « *search literacy* », à la connaissance et au choix éclairé des outils de recherche. Le second concerne les spécialistes des sciences des bibliothèques, et leur rôle d'acteur au côté des chercheurs appliquant les méthodes de recherche fondée sur les faits avérés (*evidence-based*), dans le dialogue avec les fournisseurs des systèmes de recherche pour améliorer ces derniers.

Soulignons enfin le lien fort de cet article avec l'actualité, la crise de la Covid-19 mettant en lumière l'urgence à se repérer dans la masse croissante des publications.

Klopfenstein & Dampier, 2020 : comparatif de Google Scholar et PubMed

« But the GS search interface has severe deficiencies that make literature searches laborious and, most importantly, unreproducible. »

« Mais l'interface de recherche de GS présente de graves lacunes qui rendent la recherche documentaire laborieuse et, surtout, non reproductible. »

On retrouve dans cet article intermédiaire, centré sur deux outils de recherche, des apports techniques et méthodologiques autant qu'une contribution à une réflexion plus large sur la recherche bibliographique.

Nous ne reviendrons pas sur l'aspect réflexif, qui concerne les trois types de recherche présentés de façon très approfondie par Gusenbauer et Haddaway (2021). Cet aspect réflexif

met aussi en perspective la recherche bibliographique avec certaines exigences de la science ouverte, qu'il s'agisse des principes FAIR ou de la reproductibilité.

S'agissant des aspects techniques et méthodologiques, l'apport de cet article concerne tant les résultats, le comparatif, et notamment sa figure récapitulative, que la méthode utilisée pour produire ce comparatif, fondé sur six critères de comparaison. Au-delà de leurs propres résultats, les auteurs invitent en effet tous les chercheurs, afin de rendre leurs recherches plus efficaces et plus rapides, à évaluer leurs systèmes de recherche au moyen des cinq critères principaux proposés. Ces cinq critères sont repris de Boeker *et al.* (2013) : caractère reproductible de la recherche, export, historique de recherche, documentation de la stratégie de recherche (documentation fournie par l'éditeur de l'outil de recherche), recherche avancée (nombre de champs de recherche et filtres). Les auteurs ajoutent un sixième critère correspondant à une fonctionnalité phare de Google Scholar en partie seulement présente dans PubMed, à savoir l'exploitation des liens de citations vers les articles citants (*forward*). Les auteurs proposent de fait et mettent à disposition un outil pour pallier ce manque dans PubMed. Les six critères de comparaison sont passés en revue de façon détaillée ; des boîtes de texte concernant les limites de Google Scholar, créées à partir de l'aide de Google Scholar, complètent et illustrent ce texte. Un comparatif de la couverture et du mode de sélection des ressources dans ces deux outils parachève le comparatif des fonctionnalités.

Soulignons enfin le résumé précis et étayé d'exemples des limites de Google Scholar présenté en introduction, de laquelle nous avons extrait la citation mise en exergue, qui parlera y compris aux non-utilisateurs de PubMed.

Gusenbauer & Haddaway, 2020 : évaluation des performances de recherche de 28 ressources

« We advocate educated use of these systems, so users have the right tool for the right purpose fully aware of its strength and weakness. »

« Nous prôtons une utilisation éduquée de ces systèmes, afin que les utilisateurs disposent du bon outil pour le bon usage, tout en étant conscients de ses forces et de ses faiblesses. »

Nous voici arrivés à l'article de recherche qui a initié cette série d'articles. À nouveau, nous considérons en premier lieu les apports techniques et méthodologiques, pour aborder en second lieu sa contribution à une réflexion plus large concernant la recherche documentaire.

Mentionnons tout d'abord que cet article a une portée immédiate plus restreinte que les deux commentaires qu'il a suscités, puisqu'il concerne un type de recherche particulier, la recherche systématique, sans considération des autres types de recherche. Du point de vue des domaines couverts, bien que les auteurs indiquent avoir effectué une sélection multidisciplinaire des outils, on trouve surtout de fait des ressources en informatique, en dehors des outils utiles pour les domaines des trois grandes collaborations de recherche

systematique (Cochrane¹ pour la médecine, CEE² pour les sciences de l'environnement, Campbell³ pour les sciences humaines).

Les apports techniques et méthodologiques concernent bien sûr le comparatif des différents outils. À nouveau la méthode est aussi intéressante que le résultat en lui-même, c'est-à-dire les tableaux comparatifs exhaustifs de tous les systèmes analysés. Les vingt-sept critères de comparaison appliqués constituent une traduction en termes de fonctionnalités et de comportements objectivement observables des exigences d'une recherche bibliographique systématique ; les modes d'évaluation sont également très précis et documentés. Ces exigences concernent l'exhaustivité, la reproductibilité et la transparence. L'exhaustivité se traduit bien sûr en couverture (disciplinaire, géographique, temporelle et par type de document), mais aussi en termes de capacités d'un système à rendre compte de cette couverture. Une couverture importante sans les fonctionnalités de recherche permettant de l'explorer pour en extraire le contenu pertinent voit ainsi sa valeur considérablement amoindrie. Ces critères se voient compléter par une caractéristique permettant d'affiner le classement, celle du caractère nécessaire ou désirable d'une réponse positive au critère. Le comparatif fournit ainsi une liste de ressources pouvant être utilisées comme ressources principales pour une recherche systématique, les autres ressources étant considérées comme secondaires.

À partir des résultats et de leurs discussions, les auteurs développent des points de réflexion générale qui seront bien plus présents dans leur commentaire de 2021, et que l'on retrouve également dans la version vulgarisée de l'article publiée sur le blog « Impact of Social Sciences » (Haddaway & Gusenbauer, 2020). Ces points concernent à la fois la sensibilisation aux biais et limitations des outils de recherche (Google Scholar notamment et plus largement des moteurs de recherche) et la sensibilisation à l'importance de la connaissance et du choix éclairé des outils de recherche documentaire, « *search literacy* ».

Conclusion

Quelles sont les pistes ouvertes par ces articles ? Elles nous semblent nombreuses pour les bibliothécaires, quel que soit leur niveau d'intervention, et plus encore à l'heure où Google Scholar apparaît dans l'enquête nationale la plus récente comme l'outil d'accès privilégié à la documentation scientifique par les chercheurs (Baligand *et al.*, 2021).

- Une meilleure compréhension du cadre pratique et théorique dans lequel s'inscrit la recherche bibliographique de niveau recherche.
- Une meilleure connaissance des outils de recherche bibliographique, tant de ceux qu'ils utilisent que de ceux qu'ils ne connaissent pas encore.
- Une méthode de comparaison des outils de recherche, applicable à des outils non considérés dans ces articles.

¹ <http://www.cochrane.org/>

² <https://environmentalevidence.org/>

³ <https://www.campbellcollaboration.org/>

- Des contenus à transmettre auprès des chercheurs et des étudiants : le commentaire de Klopfenstein et Dampier se prête bien à cet usage, qu'il s'agisse du texte entier ou de reprendre les figures et/ou les boîtes « Google Scholar ».
- Des arguments et des exemples pour soutenir l'importance de la « *search literacy* » et le rôle que peuvent jouer les bibliothèques en la matière.

Références citées

- Baligand, M. P., Colcanap, G., Harnais, V., Rousseau-Hans, F., & Weil-Miko, C. (2021). *Les pratiques de recherche documentaire des chercheurs français en 2020* [Rapport technique]. Couperin. <https://doi.org/10.5281/zenodo.4562180>
- Boeker, M., Vach, W., & Motschall, E. (2013). Google Scholar as replacement for systematic literature searches: good relative recall and precision are not enough. *BMC Medical Research Methodology*, 13(1), 131. <https://doi.org/10.1186/1471-2288-13-131>
- Gusenbauer, M. (2019). Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. *Scientometrics*, 118(1), 177-214. <https://doi.org/10.1007/s11192-018-2958-5>
- Gusenbauer, M., & Haddaway, N. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed and 26 other resources. *Research Synthesis Methods*, 11(2), 181-217. <https://doi.org/10.1002/jrsm.1378>
- Gusenbauer, M., & Haddaway, N. (2021). What every researcher should know about searching – clarified concepts, search advice, and an agenda to improve finding in academia. *Research Synthesis Methods*, 12(2), 136-147. <https://doi.org/10.1002/jrsm.1457>
- Haddaway, N., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The Role of Google Scholar in Evidence Reviews and Its Applicability to Grey Literature Searching. *PLOS ONE*, 10(9), e0138237. <https://doi.org/10.1371/journal.pone.0138237>
- Haddaway, N., & Gusenbauer, M. (2020). A broken system – why literature searching needs a FAIR revolution. *Impact of Social Sciences*. <https://blogs.lse.ac.uk/impactofsocialsciences/2020/02/03/a-broken-system-why-literature-searching-needs-a-fair-revolution/>
- Klopfenstein, D. V., & Dampier, W. (2021). Commentary to Gusenbauer and Haddaway 2020 : Evaluating retrieval qualities of Google Scholar and PubMed. *Research Synthesis Methods*, 12(2), 126-135. <https://doi.org/10.1002/jrsm.1456>