

Ouvrir les bibliographies pour le futur – Un modèle de bibliographies collaboratif et mû par la recherche

French Translation of the Original Paper: “Opening up the bibliographies for the future – A collaborative research-driven model for bibliographies”.

Translated by: Mathilde Koskas, Bibliothèque nationale de France, Paris, France.

Hege Stensrud Høsøien

Recherche et collections, Bibliothèque nationale de Norvège Oslo, Norvège.

Hege.Hosoiien@nb.no

Irene Hole

Acquisitions et catalogage, Bibliothèque nationale de Norvège Oslo, Norvège.

Irene.Hole@nb.no



This is a French translation of “Opening up the bibliographies for the future – A collaborative research-driven model for bibliographies” Copyright © 2013 by **Hege Stensrud Høsøien and Irene Hole**. This work is made available under the terms of the Creative Commons Attribution 3.0 Unported License: <http://creativecommons.org/licenses/by/3.0/>

Résumé :

Bien que la production et la médiation de différentes sortes de données bibliographiques représente une des activités fondamentales de la bibliothèque, aujourd’hui un certain nombre de pratiques, de technologies et de disciplines influence ce vaste domaine. Même si la bibliographie est habituellement comprise comme le produit de modèles bibliothéconomiques, le domaine inclut aussi certains types de recherches sur des bases empiriques, la philologie et l’histoire du livre, entre autres disciplines.

Au sein de la communauté des bibliothèques, on s’est concentré sur la situation actuelle de changements technologiques constants, l’émergence de nouvelles technologies de recherche/interfaces utilisateurs, la fusion des catalogues et des bibliographies, l’intégration de catalogues de bibliothèques individuels dans des méta-catalogues, et les programmes de numérisation à grande échelle. On s’est aussi concentré sur la réalisation et l’implémentation des FRBR, du Web de données et des technologies sémantiques.

Jusqu’à récemment, la production de bibliographies spécialisées s’est concentrée sur la consolidation de l’infrastructure de recherche nationale. Aujourd’hui, toutefois, les bibliographies spécialisées cherchent aussi à incorporer des données de recherche et des modèles théoriques du monde universitaire.

Cette communication traite du modèle de production de bibliographies spécialisées de la Bibliothèque nationale de Norvège. Les exemples de trois bibliographies spécialisées, Hamsun,

Pour aborder le défi proposé par le comité IFLA et par le titre de cette session, **Ouvrir les bibliographies pour l'avenir**, il semble juste de demander, avant de commencer à réfléchir aux moyens d'ouvrir les bibliographies pour l'avenir, qu'est-ce que la bibliographie aujourd'hui pour nous, pour les bibliothécaires et les bibliothèques ? Le mot bibliographie lui-même vient du grec βιβλιογραφία, *bibliographia*, littéralement « écriture d'un livre ». Les auteurs grecs ont d'abord utilisé le mot pour décrire la copie manuscrite des livres. Au XII^e siècle, on a commencé à utiliser bibliographie pour « l'activité intellectuelle de composer des livres », mais ce n'est pas avant le XVII^e siècle que « la description des livres » a émergé comme sens moderne du mot.

En réfléchissant au titre de cette session, **Ouvrir les bibliographies pour l'avenir**, je me rends compte que ce titre lui-même et l'usage de la forme particulière « la bibliographie », au singulier, sont, en fait, exactement ce que je voudrais traiter. Je crois que cette session a d'abord eu pour titre **Des bibliographies pour l'avenir**. Les deux titres correspondent au mot « bibliographie » compris comme un nom, un ou des produits. Peut-être aussi un sens du mot qui reflète la forme ou la convention bibliographique que nous sommes le plus habitués à rencontrer dans la bibliothèque, la bibliographie énumérative, la liste méthodique de livres, par exemple une bibliographie nationale. D'un point de vue interne à la bibliothèque, il est facile de voir que nous nous entourons, nous immergeons dans les bibliographies et les activités bibliographiques : enregistrement et médiation de métadonnées bibliographiques, discussions sur leurs normes et leurs règles, parler MARC, MODS, Dublin Core et RDF.

Pendant ce temps, dans les universités et le milieu universitaire, on s'est plutôt intéressé à une autre forme de bibliographie, la bibliographie descriptive (et ses différentes manifestations : analytique, historique, etc.). Cette forme est la description systématique et souvent très précise des livres comme objets physiques, qui intéresse les philologues de toutes sortes et, de plus en plus, le monde universitaire, dans le domaine pluridisciplinaire de ce que nous pouvons appeler histoire du livre ou *book history*.

Dans l'avant-propos de son influent essai *La bibliographie et la sociologie des textes*, (McKenzie, 1984), le bibliographe néo-zélandais D. F. McKenzie (1931-1999), qui fut ensuite maître de conférences et professeur de bibliographie et de critique textuelle à l'université d'Oxford, affirme :

« Les processus historiques familiers selon lesquels, au cours des siècles, les textes ont changé de forme et de contenu ont à présent accéléré au point de rendre la définition et la localisation de l'autorité textuelle à la manière traditionnelle à peine possible. Les bibliothécaires professionnels, sous la pression de changements technologiques et sociaux irrésistibles, redéfinissent leur discipline afin de décrire, conserver et donner accès à du son, des images statiques et animées, avec ou sans paroles, et à un flux d'information sur ordinateur. À titre de comparaison, ce n'est que récemment que la bibliographie universitaire a commencé à trouver une motivation nouvelle dans ces évolutions et à exploiter l'expérience et l'intérêt nouveaux d'étudiants pour lesquels les livres ne représentent qu'une forme de texte parmi d'autres. » (McKenzie, 1984, 1).

En quelques phrases, McKenzie dit plusieurs choses qui donnent matière à réflexion. À propos de la bibliographie, il inclut le genre énumératif et le genre descriptif, mais il la considère moins comme un produit que comme une ou des disciplines. Il décrit l'accélération du processus de changement de forme et de contenu des textes (un sujet dans lequel quiconque travaille dans une bibliothèque peut se reconnaître) et la redéfinition de leur discipline par les bibliothécaires afin de « décrire, conserver et donner accès à du son, des images statiques et animées, avec ou sans paroles, et à un flux d'information sur ordinateur ». Cela traduit un élargissement du champ de ce que nous avons l'habitude de considérer comme l'objet de la bibliographie, à savoir le livre ou du moins un document écrit d'une sorte ou d'une autre.

Par opposition avec la bibliothécaire professionnelle, désireuse de redéfinir sa discipline face à des évolutions technologiques « irrésistibles », McKenzie montre une bibliographie universitaire qui n'a que récemment « commencé à trouver une motivation nouvelle dans ces évolutions ». En identifiant deux disciplines différentes, l'une bibliothéconomique et l'autre universitaire, il paraît logique que les deux écoles aient des perceptions différentes non seulement de ce qui constitue la bibliographie comme produit ou discipline, mais aussi de la manière « d'ouvrir la bibliographie pour l'avenir », selon le thème de notre session.

Pour McKenzie, la conséquence du rythme accéléré des changements de forme et de contenu textuels est que la définition et la localisation de l'autorité textuelle « à la manière traditionnelle » est « à peine possible ». Pour comprendre son commentaire, il faut le mettre en regard du genre de discipline bibliographique où des éditeurs philologues cherchent à établir l'autorité textuelle en essayant de comprendre l'intention de l'auteur et d'établir quel texte peut servir d'exemplaire. Le but est de déterminer quel texte est un original fiable auquel comparer d'autres textes et sur lequel publier de nouvelles éditions, qu'il s'agisse du manuscrit livré par l'auteur à l'imprimeur, de la première édition publiée ou de la dernière édition publiée dans le processus éditorial de laquelle l'auteur a été impliqué.

Refusant de se concentrer uniquement sur l'établissement de l'autorité textuelle comme but principal de la bibliographie descriptive, McKenzie présente le concept de bibliographie comme une sociologie des textes :

« [...] la bibliographie est la discipline qui étudie les textes comme formes enregistrées et leurs processus de transmission, y compris leur production et leur réception [...] j'inclus dans ma définition du texte les données verbales, visuelles, orales et numériques [...] » (McKenzie, 1984, 12).

Ce glissement au sein de la bibliographie universitaire de préoccupations philologiques centrées sur l'autorité textuelle vers la « sociologie des textes » a été aiguillonné par des bibliographes et des historiens du livre comme McKenzie lui-même, mais aussi par Philip Gaskell, Robert Darnton et Roger Chartier et leurs pareils, pour citer quelques-uns des plus connus. On pourrait dire que ce nouveau point de vue a constitué un changement de paradigme pour la bibliographie universitaire, changement qui a considérablement contribué à revitaliser le domaine.

Au sein de la communauté des bibliothèques, d'un autre côté, les attentes pour l'avenir de la bibliographie, en plus d'inclure dans la notion de texte les données verbales, visuelles, orales et numériques, semblent se concentrer principalement sur les avancées pratiques et

technologiques : l'émergence de nouvelles technologies de recherche, interfaces utilisateurs et applications, l'intégration des catalogues et des bibliographies, la fusion de catalogues individuels de bibliothèques dans des méta-catalogues, et la mise en place de nouveaux modèles de représentation des métadonnées, comme RDF (Resource Description Framework). Et, ainsi que quiconque a assisté à un congrès de bibliothèques d'une certaine ampleur ces dernières années peut en témoigner, nous rêvons de la réalisation et de l'implémentation des FRBR (Functional Requirements for Bibliographic Records – Fonctionnalités requises des notices bibliographiques), du Web de données et des technologies sémantiques.

McKenzie a écrit *La bibliographie et la sociologie des textes* en 1984, il y a presque 30 ans. Lorsque nous avons commencé à travailler avec notre nouveau modèle de production de certains types de bibliographies à la Bibliothèque nationale de Norvège, il y a quelques années, la situation était à peu près la même. Tandis que les bibliothécaires cataloguaient et produisaient des métadonnées bibliographiques et des bibliographies comme la Bibliographie nationale, les professionnels des communautés universitaires produisaient des bibliographies dans Word et les publiaient sous forme de livres. À l'approche du tournant du siècle, nous avons commencé à convertir les fichiers Word en HTML et à les publier sur des sites internet. En particulier, s'il y avait à manipuler un vaste corpus de textes, nous avons commencé à entrer les données dans une base de données propriétaire, ce qui a demandé beaucoup de travail aux bibliothécaires.

Ayant décidé de faire les choses différemment à l'avenir, nous avons commencé à réfléchir à un nouveau modèle de production de certaines bibliographies à la Bibliothèque nationale, un modèle qui prendrait les meilleurs aspects des deux traditions bibliographiques. Le modèle a été produit en collaboration par les départements de la recherche, du catalogage et des technologies de l'information de la Bibliothèque nationale. Nous voulions une solution numérique qui soit pragmatique, technologiquement dynamique et durable, et qui puisse être utilisée à la fois par la Bibliothèque nationale et par d'autres bibliothèques. Après avoir examiné différentes options, nous avons décidé de produire les bibliographies dans BIBSYS, le méta-catalogue collaboratif norvégien, qui rassemble plus de 120 grandes bibliothèques, parmi lesquelles toutes les bibliothèques universitaires, la plupart des bibliothèques de recherche et la Bibliothèque nationale. Nous avons utilisé BIBSYS comme outil de production des bibliographies, donnant ensuite à chaque nouvelle bibliographie une interface utilisateur personnalisée sur notre propre site internet. De plus, l'utilisation de BIBSYS présentait un avantage significatif par rapport à d'autres solutions. Il nous a permis d'utiliser le MARC (Machine Readable Cataloguing – Catalogage lisible par machine), une norme internationale pour les notices bibliographiques bien connue, qui rend la coopération, la conversion et la migration des données, etc., relativement faciles. BIBSYS contient des données concernant un grand nombre de collections de bibliothèques, avec les titres et les exemplaires de chacune, ce qui nous a permis d'identifier un grand nombre de documents à inclure dans la bibliographie. Après les avoir identifiés, nous avons pu rediffuser des métadonnées de grande qualité provenant de la Bibliographie nationale (BIBSYS contient la Bibliographie nationale depuis 2011) et d'autres bibliothèques, en ajoutant une balise à des articles de la bibliographie déjà existants. BIBSYS contient aussi un fichier d'autorité virtuel des noms, développé et maintenu en coopération avec la Bibliothèque nationale, et le choix de BIBSYS nous a donc permis de nous servir du fichier national d'autorité des noms et d'y contribuer. La solution intégrée a donné aux usagers des informations sur la localisation de l'exemplaire d'un livre et l'accès aux options de prêt.

Enfin, et cela s'est avéré de plus en plus important, BIBSYS contient les données et les liens aux fichiers numérisés provenant du programme de numérisation de la Bibliothèque nationale. Commencé quelques années auparavant, le programme de numérisation, dont le but est de numériser l'intégralité des collections de la Bibliothèque nationale, prenait de l'ampleur, et le gouvernement était désireux d'augmenter le taux de numérisation et de consultation. En 2009, un accord de licence collectif étendu appelé Bokhylla (L'étagère) entre la Bibliothèque nationale et l'organisation Kopinor, représentant les auteurs, les éditeurs et autres ayants-droit, a donné à tous les utilisateurs dont l'adresse ip est en Norvège l'accès libre à 50 000 titres norvégiens publiés dans les années 1990. En 2012, les deux parties ont fait une évaluation positive du contrat et l'accord a été étendu aux livres imprimés publiés en Norvège jusqu'à l'année 2000 incluse. En 2013, 110 000 livres sont numérisés et librement accessibles, et l'accord prévoit 250 000 titres en 2017. Lorsque nous avons travaillé sur le corpus d'une bibliographie, tous les titres en ont été numérisés.

Au sein du nouveau modèle, la bibliographie Hamsun est la première réalisée par la Bibliothèque nationale. Dans le cadre des célébrations nationales du cent-cinquantième, la Bibliothèque nationale a présenté une bibliographie sur le lauréat norvégien du prix Nobel Knut Hamsun (1859-1952). La bibliographie a été produite en coopération avec la bibliothèque universitaire de Tromsø et a apporté la preuve des possibilités du modèle pour la production collaborative de bibliographies. En mettant à niveau et en complétant des articles bibliographiques existants, notre travail a amélioré le méta-catalogue collaboratif et la Bibliographie nationale, ainsi qu'ajouté beaucoup de nouvelles entrées dans le fichier d'autorité national des noms de personnes et d'organisations. En travaillant de manière systématique sur un corpus de textes, nous nous sommes rendu compte des capacités du modèle à identifier des lacunes dans les collections de la Bibliothèque nationale, des documents disparus sans que l'on s'en rende compte au cours des années, et des documents qui pour une raison ou une autre n'avaient jamais fait l'objet d'un dépôt légal.

Gagnant en confiance et en compétence, nous avons ensuite produit une bibliographie sur le lauréat norvégien du prix Nobel Bjørnstjerne Bjørnson (1832-1910) dans le cadre des commémorations nationales du centenaire de sa mort. En nous affranchissant des procédures antérieures de la bibliothèque et en considérant à présent la bibliographie non seulement comme une infrastructure de recherche mais aussi comme une recherche collaborative, et, en fin de compte, comme une médiation et une dissémination de la recherche, nous avons travaillé en étroite collaboration avec les chercheurs et le biographe de Bjørnson, Edvard Hoem. Modifiant les procédures de travail qui consistent à examiner un corpus de manière alphabétique ou chronologique, nous avons adopté le pragmatisme comme nouveau mode de travail et avons donné la priorité aux points sur lesquels notre travail était le plus nécessaire. Nous avons ajouté des types de documents qui intéressaient les chercheurs, tels des articles de journaux et des discours publiés, identifié des documents qu'ils recherchaient et les avons mis à disposition, ajoutant de nouvelles couches de métadonnées. Par exemple, nous avons ajouté des balises pour la géolocalisation sur une carte, et adapté et exporté des données dans une application Google maps pour permettre la visualisation des déplacements et de l'engagement de Bjørnson. En contrepartie, les chercheurs ont apporté leurs connaissances de spécialistes en identifiant les auteurs et les pseudonymes, en proposant des mots-clés, en identifiant des documents pertinents, etc. Ainsi, les chercheurs ont apporté leur contribution non seulement en métadonnées utilisateurs, mais aussi en métadonnées de la recherche. Leurs connaissances ont enrichi la Bibliographie nationale, le Fichier d'autorité national et le catalogue collaboratif.

Pour le travail sur Hamsun et Bjørnson, les corpus de textes étaient vastes et nous avons procédé à une stricte évaluation de ce qui était à inclure. En 2011, toutefois, nous avons commencé à travailler à une bibliographie pour célébrer un des auteurs contemporains les plus populaires de Norvège, Dag Solstad (né en 1941). Pour la première fois, nous avons pu tenter d'inclure tout ce qui avait été écrit par un auteur et à son sujet. Examiner de plus près la production d'un auteur a montré une importante limite à ce modèle de production de bibliographies : il ne contenait aucune information sur les relations entre les documents, ce qui rendait difficile d'obtenir une vue d'ensemble et une orientation à la façon préférée des chercheurs et du grand public. Les normes bibliographiques de l'époque où les bibliographies étaient publiées sous forme de livres, et où les notices des collections des bibliothèques étaient sur fiches, de simples fonctionnalités comme pouvoir suivre la publication d'une œuvre dans de nombreuses éditions différentes au cours des années, sont abandonnées dans le catalogue informatisé. En ajoutant de nouvelles couches de métadonnées et en travaillant sur l'algorithme de recherche et l'interface utilisateur, nous avons pu recréer certaines de ces relations et donner à l'utilisateur un moyen de naviguer dans la production d'un auteur (les nombreuses éditions et traductions des œuvres, et pour finir leur réception). C'est une réalisation partielle et pragmatique de certains des principes qui sous-tendent les FRBR.

Étudier en détail un corpus textuel d'œuvres et ses manifestations nous a aussi fait prendre conscience que, bien que nous ayons utilisé les modèles bibliographiques pour identifier des lacunes dans nos collections, nous pouvions maintenant identifier des lacunes d'un autre type. Ces quelques dernières années, le modèle nous a permis d'améliorer nos compétences en histoire du livre et sur l'aspect matériel des livres. Partant, nous avons pu faire le tri dans le corpus, nous intéresser aux différentes reliures d'une même édition, relever les cas où la bibliothèque possédait un livre non dans sa reliure originale, mais dans une reliure de bibliothèque, et identifier les livres avec une provenance intéressante. Ainsi, nous avons pu nous rapprocher de notre ambition de posséder une collection complète d'imprimés norvégiens dans toutes leurs variantes matérielles, physiques et philologiques.

Nous continuons à utiliser notre modèle bibliographique comme outil de développement. Nos projets pour l'avenir sont de travailler à la création d'un fichier d'autorité des œuvres. Avec le fichier d'autorité des noms, déjà publié dans VIAF et dans le Web de données, cette évolution nous rapprochera encore de l'utilisation des technologies sémantiques. De plus, nous travaillons actuellement à une filmographie, selon les mêmes principes. Les métadonnées concernant les films se trouvent dans notre autre grand réservoir de métadonnées, MAVIS, qui contient aussi les métadonnées de la photographie, de la radio, de la télévision et de la musique. Rapprocher ces deux réservoirs et donner une interface utilisateur aux jeux de données nous permettra de faire à l'avenir de véritables bibliographies multimédia.

Enfin nous sommes aussi conscients que la bibliographie peut jouer un rôle petit mais significatif dans le futur des humanités numériques et de la culturomique. Schreibman et al., dans *A companion to digital humanities*, décrivent la situation :

« Le World Wide Web a incité toute une génération de pseudo-publications : des documents plus largement disponibles qu'aucune publication imprimée ne l'a jamais été pouvaient à n'importe quel moment arriver sur des millions d'appareils. Cependant, les mêmes documents tournaient souvent sur des comptes séparés, beaucoup d'URL étant changées ou pointant vers des documents qui n'étaient plus en ligne, ou, ce qui est sans doute pire, qui avaient fait l'objet de modifications substantielles depuis l'ajout du lien original.

Divers réservoirs de bibliothèques commencent à présent à être mis en service». (Schreibman et al., 2004).

À mesure que les bibliothèques numérisent leurs collections et que les chercheurs passent de la lecture détaillée de quelques textes canoniques à la lecture à distance d'énormes bibliothèques numériques, selon la description de Franco Moretti, la demande de corpus de textes numériques accompagnés de métadonnées fiables va augmenter. Les bibliothèques peuvent jouer le rôle de réservoirs de confiance, garants non seulement de la qualité de la numérisation, mais de ce que la numérisation représente précisément ce que l'utilisateur recherche, par exemple la première édition de *Sult* de Knut Hamsun, et pas une quelconque édition plus tardive ou une version abrégée (Moretti, 2011). Compte tenu de ce nouveau rôle des bibliothèques, nous pourrions être témoins d'un nouveau glissement de la bibliographie : de la question originelle de l'autorité textuelle à la question de la sociologie des textes et retour à une nouvelle forme d'autorité textuelle.

Revenons à présent au défi du titre de la session : Ouvrir les bibliographies pour l'avenir. Nous pouvons répondre qu'il nous est possible d'élargir notre notion du texte pour considérer la bibliographie énumérative aussi bien que descriptive – la bibliothéconomie aussi bien que la bibliographie universitaire – comme un seul grand domaine interdisciplinaire. Nous pouvons répondre qu'il est bénéfique de faire entrer des données de recherche et des modèles théoriques provenant de la communauté universitaire dans l'appareil de la bibliothèque.

Bibliographie

Bjørnstjerne Bjørnson bibliography. Newspaper articles and speeches by Bjørnson; literature on Bjørnson <http://www.nb.no/bibliografi/bjornson/search.jsf>

Hamsun bibliography : literature on Knut Hamsun
<http://www.nb.no/bibliografi/hamsun/search.jsf>

McKenzie, D.F. 1999. *Bibliography and the sociology of texts*. Cambridge University Press.

Traduction française (de l'édition originale) :

McKenzie, D.F. 1991. *La bibliographie et la sociologie des textes*. Éd. du Cercle de la librairie.

Moretti, Franco. 2011. *Network theory, plot analysis*. Palo Alto: Stanford University.
(Pamphlet/Literary Lab; 2)

Schreibman, Susan; Siemens, Ray; Unsworth, John. 2004. *A companion to digital humanities*. Blackwell.

Solstad bibliography : literature by and on Dag Solstad
<http://www.nb.no/bibliografi/solstad/search.jsf>