
Title of the Satellite Meeting: New Horizons in Artificial Intelligence in Libraries

Date: 21-22 juillet 2022

Location: Galway, Irlande

Pour une éthique de la recommandation personnalisée à la Bibliothèque nationale de France

Céline Leclaire

Questions scientifiques et techniques, Direction des Services et des réseaux, Bibliothèque nationale de France (BnF), Paris, France

celine.leclaire@bnf.fr

Lucie Termignon

Questions scientifiques et techniques, Direction des Services et des réseaux, Bibliothèque nationale de France (BnF), Paris, France

lucie.termignon@bnf.fr



Copyright © 2022 by **Céline Leclaire and Lucie Termignon**. This work is made available under the terms of the Creative Commons Attribution 4.0 International License:

<http://creativecommons.org/licenses/by/4.0>

Résumé:

Dans le domaine de l'intelligence artificielle, les questions relatives à l'éthique font l'objet de nombreuses réflexions et débats au sein du monde de la recherche, et des institutions, chez les porteurs de projets... Nous proposons d'illustrer et nourrir ce questionnement en nous appuyant sur l'exemple concret d'un développement en cours à la Bibliothèque nationale de France (BnF) : l'implémentation d'un nouveau service de recommandation dans la bibliothèque numérique [Gallica](#), qui apparaît comme un outil de valorisation de plus en plus nécessaire eu égard à la richesse et la spécificité de ses collections, et aux attentes de ses publics (en particulier des chercheurs).

Au-delà des défis techniques, ce nouveau service soulève des risques éthiques – les bulles de filtre, la trahison de la confiance de l'utilisateur par l'exploitation de ses données – et oblige la Bibliothèque à mettre au clair la position qu'elle entend occuper dans le paysage documentaire. Il s'agit de réfléchir au sens profond de la recommandation et du rôle du bibliothécaire.

Quels principes éthiques guident l'action de la Bibliothèque alors qu'elle se lance dans la mise en place d'un service de recommandation personnalisée de contenus outillé par l'intelligence artificielle (IA) ? En quoi cette approche est-elle représentative de la politique globale de l'institution en matière d'IA ?

Mots clés: Intelligence artificielle, recommandation personnalisée, bibliothèque numérique, éthique

Introduction

Aujourd'hui, lorsqu'un internaute fait une recherche dans Gallica (la bibliothèque numérique de la Bibliothèque nationale de France – BnF – et de ses partenaires), il obtient exactement les mêmes résultats que tout autre internaute qui effectuerait la même recherche : il ne bénéficie pas de recommandations personnalisées. Cela s'explique par un principe déontologique auquel est attachée la BnF : les données personnelles des utilisateurs et en particulier leurs traces de navigation ne sont en aucune manière utilisées par le moteur de recherche pour orienter les résultats proposés. Seules sont prises en compte les métadonnées relatives aux documents.

Cependant, plusieurs facteurs sont susceptibles de faire évoluer cette position, en particulier la maturité des technologies, de plus en plus à même de traiter des millions de données en temps réel, l'étude des attentes des utilisateurs, et l'augmentation croissante des ressources culturelles disponibles, qui nécessitent l'ouverture de nouvelles portes d'entrée dans les collections, de nouveaux chemins de médiation, *a fortiori* quand il s'agit d'univers francophones n'alimentant pas de manière naturelle les corpus d'entraînement des algorithmes.

Pour mieux répondre à ces élans, la BnF s'est dotée d'un cadre général à travers la feuille de route sur l'intelligence artificielle qu'elle a publiée en 2022. Celle-ci précise en la manière dont l'IA concerne tous les domaines de la BnF. Elle vise aussi à établir des projets phares et des priorités tout en installant un solide arrière-plan stratégique et un dispositif éclairé d'accompagnement des changements à venir (via la formation, l'évaluation...).

C'est dans ce contexte favorable, enthousiasmant, mais également tout rempli de questionnements (comment concilier la déontologie du service public et les opportunités offertes par la technologie en matière d'amélioration du service ?) que s'inscrit le projet de recommandation personnalisée de contenus dans Gallica, premier exemple à la BnF d'un projet où l'éthique est une composante par défaut dès la conception du service.

Pour étudier ce projet, nous préciserons par touches successives la définition de la recommandation personnalisée en nous appuyant, à chaque fois, sur des questions d'ordre éthique : en avons-nous vraiment besoin ? la collecte des données personnelles est-elle aussi inéluctable qu'on nous le présente parfois ? de quels garde-fous disposons-nous collectivement ?

Qu'est-ce que la recommandation personnalisée par algorithme et pourquoi en avons-nous besoin ?

Plaçons-nous pour commencer du côté du concepteur, soit une bibliothèque mue par sa responsabilité de service public national.

Le principe de la recommandation

Tout travail impliquant l'intelligence artificielle conduit à regarder d'un œil nouveau des pratiques professionnelles ancrées dans le quotidien. Par exemple, lorsqu'on développe un agent conversationnel (*chatbot*) pour le service public, on se demande ce qu'est,

fondamentalement, une conversation humaine, et, en cherchant à les expliciter, on touche aux limites des règles qu'on impose aux usagers.

Dans le cas de la recommandation personnalisée, on s'aperçoit qu'elle recouvre une multitude d'opérations. Imaginons que je sois dans une salle ou que je tiens une permanence à distance, et qu'une personne me demande un conseil de lecture. Je vais lui demander de préciser, m'enquérir de ce qu'elle a déjà consulté, analyser le sujet en profondeur, je vais mobiliser ma culture générale, ma connaissance du fonds assistée par les métadonnées des catalogues, et, le cas échéant, faire appel à des bases extérieures (banques d'articles, collègues plus experts...). Enfin, j'ajusterai ma réponse à l'interlocuteur, au temps dont il dispose, et je veillerai à ce qu'il ait un accès rapide aux contenus qu'il désire, car il souhaite généralement moins des références que des accès. J'aurai eu le temps d'avoir quelques doutes : ma réponse n'est-elle pas biaisée par des présupposés que j'aurais vis-à-vis de cet interlocuteur, par exemple quant à sa capacité intellectuelle à s'emparer de telle ou telle ressource ? Où se situe la limite entre recommandation et prescription ?

La recommandation par algorithme

De telles questions n'entrent pas dans le champ des algorithmes puisqu'ils ne reposent que sur du calcul et sur des paramètres organisés, priorisés. L'objectif des systèmes de recommandation est de proposer les résultats les plus adaptés aux besoins et préférences des utilisateurs afin de faciliter la recherche et l'accès à l'information.

Les données mobilisées par l'algorithme pour produire des recommandations sont de deux types : les données relatives au contenu et les données relatives aux utilisateurs. Les premières concernent les métadonnées descriptives des contenus, les données relatives à l'utilisation et à la consultation des contenus et les données d'évaluation des contenus éventuellement fournies par les utilisateurs. Les données relatives aux utilisateurs concernent quant à elles les données personnelles d'authentification, les données liées au contexte d'utilisation et les données d'usage renseignant sur les comportements des utilisateurs. Les données de ce second ensemble peuvent être collectées de manière explicite (à travers le remplissage d'un formulaire ou la réponse à un questionnaire) ou implicite via des techniques de traçage (cookies, balises web, empreinte du navigateur, ciblage par cohorte).¹

Le système de recommandation fonctionne en trois temps : la collecte des données d'utilisateurs (1) permet de générer un profil utilisateur ou modèle de préférence (2) exploité par des méthodes d'apprentissage automatique (c'est-à-dire des familles d'algorithmes) (3) afin de déterminer les recommandations adaptées à l'utilisateur. Les algorithmes mobilisés dans les méthodes d'apprentissage sont généralement regroupés en deux types distincts : le filtrage collaboratif et le filtrage par contenu.² Ces deux approches ne sont pas exclusives l'une de l'autre, et peuvent se combiner pour atténuer leurs biais et effets de bord. Une fois le choix de collecte et de traitement des données effectué et le modèle entraîné, l'intervention

¹ Le ciblage par cohorte (ou *Federated Learning of Cohorts*, FLoC) est une initiative développée par Google qui consiste à cibler un groupe d'utilisateurs aux comportements similaires et non pas un utilisateur de manière individuelle. A partir de centres d'intérêt ou traces d'usages communs, les utilisateurs sont agrégés en groupes anonymes repérés par un identifiant unique et persistant, partagé par tous les utilisateurs de la même cohorte et géré au niveau du navigateur ou du système d'exploitation.

² Le filtrage collaboratif exploite la connaissance d'une classe d'utilisateurs similaires à l'utilisateur pour prédire ses futurs agissements tandis que le filtrage par contenu permet de proposer des recommandations sur la seule base des similarités calculées entre contenus.

humaine n'est plus de mise. L'algorithme est capable, via une interface, de proposer des contenus à l'utilisateur, sans intermédiation humaine. S'établit une sorte de jeu entre l'utilisateur et l'algorithme : plus l'algorithme a de retours sur ce qu'il propose, plus il s'affine.

Le charme de l'algorithme à l'heure du *big data*

Le premier motif qui pousse la BnF à s'engager dans un projet de recommandation personnalisée dans Gallica, autrement dit à choisir d'investir de l'argent public dans un service nouveau fondé sur des algorithmes, c'est l'attention qu'elle porte aux souhaits des utilisateurs. Selon l'observatoire 2020 des publics de la BnF (Bastard, 2022, p. 15 à 20), s'ils sont plus de 90% à être globalement satisfaits de Gallica, les utilisateurs considèrent à 49% comme prioritaire l'amélioration du moteur de recherche et seuls 11% se sentent bien accompagnés dans leur recherche (41% si l'on ne considère que les doctorants, dont près d'un tiers estime que les réponses du moteur de recherche ne sont pas toujours pertinentes). Parmi leurs remarques, on note par exemple : « Par manque de temps, je ne peux assouvir ma curiosité et surtout tout l'intérêt suscité par GALLICA : une merveille. Le seul bémol : un tel fourmillement en tout qu'il m'est parfois difficile, surtout par manque de temps, de repérer mes centres d'intérêt. »³ ou « Le moteur de recherche donne souvent un très grand nombre de résultats difficiles à trier »⁴.

Un besoin de simplicité fait écho à ces attentes : le mode le plus usité pour accéder aux documents de Gallica reste la recherche simple (78% des utilisateurs y recourent « à chaque fois » ou « souvent »). Comme le précise un répondant, « naviguer dans la BnF numérique (...) semble complexe »⁵. Ces réponses signent la fin du mythe d'une aptitude innée à se servir du numérique (et de Gallica en particulier). Doublées d'une réflexion déjà ancienne sur les manières d'approcher la connaissance en général et notamment les genres fictionnels (Pernoo, 2001), elles peuvent conduire à reconnaître les limites des moteurs de recherche tels qu'ils ont été pensés classiquement en bibliothèque quand il s'agit de permettre à certains usagers de trouver ce qu'ils cherchent.

Parallèlement à ces conclusions, l'activité de médiation connaît une évolution constante et implique toujours davantage les publics et les partenaires. Ces dernières années, la BnF a ainsi ouvert les portes du blog de Gallica à ses partenaires⁶ et conçu des sélections avec eux⁷, elle a interagi avec les Gallicanautes sur les réseaux sociaux, associé des youtubeurs à la mise en valeur de ses contenus patrimoniaux... Longtemps on s'est fondé, à juste titre, sur l'excellence des métadonnées bibliographiques rassemblées⁸ et on a attendu de l'utilisateur

³ Bastard, 2022, p. 10-11. Propos d'une « Femme, 50-64 ans, réside en province, niveau d'étude Bac+3, profession artisan / commerçant / chef d'entreprise, centre d'intérêt pour la reliure / l'histoire du livre, consulte exclusivement Gallica »

⁴ Bastard, 2022, p.18. Propos d'une « Femme, entre 50 et 64 ans, réside à l'étranger, professionnelle de la recherche ; utilise Gallica et les catalogues, n'est jamais venue sur site »

⁵ Bastard, 2022, p.18. Propos d'un « Homme, plus de 65 ans, réside en province, retraité, pas d'activité amateur déclarée »

⁶ Par exemple, le Musée national d'Histoire naturelle contribue à la série de billets sur les végétaux intitulée « Herbarium de Gallica » (<https://c.bnf.fr/MrN>) et l'université de Lille, à celle qui est dédiée aux « Tissus et textiles » (<https://c.bnf.fr/Ogx>)

⁷ À l'occasion de la labellisation « 2020, année de la BD » en France, la BnF et la Cité internationale de la bande dessinée et de l'image ont mis en ligne sur sélection sur la bande dessinée (<https://gallica.bnf.fr/html/und/litteratures/bande-dessinee>)

⁸ La BnF est un acteur clé dans la diffusion des métadonnées bibliographiques, et des identifiants en particulier (ISNI, etc.). Elle a mis en ligne en 2011 un site de référence, data.bnf.fr (Bermès et al, 2016). Elle est en train de

d'un catalogue ou d'une bibliothèque numérique qu'il sache chercher. Mais on s'est rarement demandé ce qu'il avait à apporter de son côté... On lui reconnaît de plus en plus une « expertise », qu'il est par ailleurs habitué à déployer sur d'autres sites, en particulier commerciaux : nous reviendrons sur cette expertise.

Enfin, et surtout, les services de découverte qui existent déjà dans Gallica (sélections documentaires, bloc de conseils conçus manuellement dans le style du « knowledge graph » de Google, rapport de recherche, fonctionnalités de recherche par proximité ou de recherche géographique...) ne suffisent plus à honorer le très grand nombre et la très grande diversité des documents, pas plus qu'ils n'honorent le flux en constante extension qu'ils représentent. La plupart des utilisateurs ignorent que Gallica contient non seulement des documents imprimés, mais également des photographies, des enregistrements sonores, des vidéos ou encore des objets en trois dimensions.

L'enjeu est donc double : aider les utilisateurs à se repérer dans la masse des collections et améliorer la visibilité des plus méconnues, tout ceci en temps réel, quelle que soit l'heure de la journée, et au gré des envies infiniment diverses et variables des utilisateurs.

Qu'est-ce qu'une recommandation éthique ?

Le recours à l'algorithme de recommandation est justifié par la nécessité d'améliorer la satisfaction des utilisateurs mais présente en même temps le risque de leur déplaire si les résultats proposés par le système de recommandation ne sont pas pertinents, jugés opaques ou issus de pratiques intrusives. La confiance des utilisateurs dans l'algorithme repose sur des principes éthiques généraux mentionnés dans de nombreuses chartes éthiques et articles scientifiques et sujets à des définitions variées : l'explicabilité, la robustesse, l'équité, la responsabilité, la transparence, la supervision humaine (contrôlabilité), la sécurité, la sûreté (*privacy* et confidentialité)⁹. Dans le cas de la recommandation, ces principes se présentent comme une réponse aux travers identifiés depuis que l'utilisation de cette forme d'incitation s'est largement répandue dans de nombreux domaines (réseaux sociaux, plates-formes de vente, plates-formes de *streaming*...) : usage abusif des données personnelles, manipulation des comportements, discrimination, etc.

L'identification de ces écueils explique que les exemples de recours à un algorithme de recommandation personnalisée pour un site de consultation de contenus numériques culturels à but non lucratif (par exemple une bibliothèque) sont rares et, lorsqu'ils existent, franchissent rarement le cap de l'industrialisation.¹⁰ Cette rareté traduit la limite du simple

se doter d'un nouvel outil de catalogage qui augmentera encore le potentiel représenté par ces données, et d'accroître ses relations, dans ce domaine, avec les industries culturelles et créatives, notamment dans le cadre du développement du dépôt légal dématérialisé.

⁹ Les huit principes mentionnés correspondent à la liste des caractéristiques prioritaires de la confiance retenues par l'[AFNOR](#) pour la stratégie française de normalisation de l'Intelligence Artificielle. D'autres caractéristiques de la confiance telles que la traçabilité ou le nudging sont également mentionnées de manière récurrente dans les chartes éthiques émanant de différentes instances nationales ([CNIL](#), [Arcom](#)) et internationales ([Commission Européenne](#), [UNESCO](#)...).

¹⁰ Parmi les institutions détentrices de collections, on peut citer le projet de système de recommandation basé sur les notices bibliographiques MARC21 mené par la Bibliothèque et Archives nationales du Québec (BAnQ) en 2017, le site de consultation du catalogue en ligne du National Library Board de Singapour, le service Maken développé par la Bibliothèque nationale de Norvège ou encore l'outil de recommandation Bookarang dont le NBD Biblion a fait l'acquisition en 2020.

énoncé de grands principes éthiques et la spécificité des enjeux soulevés par la conception d'un tel algorithme eu égard à la singularité des publics destinataires (« grand public » et public expert), aux données qui nourrissent l'algorithme et aux enjeux de service public portés par les institutions. Elle incite à aborder de manière pratique les défis éthiques générés par l'usage de nouvelles technologies de recommandation : quelles exigences techniques sous-jacentes sont à l'œuvre pour contrer ces écueils et quelles métriques et contrôles permettent l'évaluation d'une recommandation ainsi fondée sur des algorithmes?

Défis éthiques relatifs aux usages d'un système de recommandation personnalisée en bibliothèque

Pour fonctionner, l'algorithme associe les métadonnées liées aux documents et leurs contenus aux données d'usage des utilisateurs. La collecte, le stockage et l'exploitation des données liées aux utilisateurs posent de nombreux questionnements éthiques dont l'enjeu principal repose sur l'établissement d'un lien de confiance entre l'utilisateur et le système de recommandation. Ces questionnements répondent à une double exigence de pertinence documentaire des résultats proposés par l'algorithme et de respect des données des utilisateurs. Il s'agit d'une part d'éviter d'enfermer l'utilisateur dans des bulles fermées de résultats et d'autre part de garantir le respect de la confidentialité afin de préserver l'intelligence des recherches individuelles.

Collecte, gestion et utilisation des données personnelles : l'enjeu de la sûreté et de la transparence

Comme évoqué précédemment, les algorithmes de filtrage déployés dans les systèmes de recommandation reposent sur la construction de « profils utilisateurs ». Ces profils sont plus ou moins personnalisés selon les données à partir desquelles ils sont élaborés (données personnelles remplies par l'utilisateur dans le cadre d'un questionnaire, données d'usage, traces de navigation voire données émanant de sources extérieures liées à l'utilisateur). À partir des données des utilisateurs collectées de manière explicite ou implicite, une matrice est construite pour mettre les données d'un utilisateur en regard des données liées aux contenus et aux usages des autres utilisateurs afin de lui proposer une liste de suggestions en fonction de sa requête. Différentes techniques et dispositifs ont été développés relevant de solutions algorithmiques (cryptographie, anonymisation des données¹¹), de protocoles ou de normes (RGPD¹² en Europe) pour garantir la protection des données personnelles aux différents niveaux du système de recommandation contre des usages malveillants. L'enjeu des expérimentations à mener est d'optimiser la pertinence des prédictions calculées par algorithme tout en garantissant la protection des données des utilisateurs.¹³

Au-delà de la question de la protection des données personnelles, l'exploitation des données liées au parcours de recherche des utilisateurs peut poser un problème de confidentialité. La singularité des publics de la BnF tient à la diversité de leurs usages, allant de la recherche

¹¹ Dans le cadre d'un projet de recherche menée avec Télécom ParisTech qui consistait à expérimenter l'analyse de traces d'usage de Gallica, la BnF a eu l'occasion d'être confrontée aux enjeux et aux solutions techniques de l'anonymisation (Nouvellet, Adrien et Beaudouin, Valérie, et al., « Analyse des traces d'usage de Gallica : Une étude à partir des logs de connexions au site Gallica », 2017)

¹² Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016, relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données)

¹³ La littérature relative aux techniques de *privacy preserving collaborative filtering* et *privacay preserving data publishing* est abondante et exige un examen approfondi (Majeed et al., 2020, Paraschakis, 2017).

extrêmement ciblée d'un document que l'utilisateur sait faire partie des contenus accessibles dans Gallica à l'exploration tous azimuts des collections guidée par la seule curiosité de l'usager. Alors que les profils chercheurs et experts seront attachés à la confidentialité de leur parcours de recherche afin de garantir le caractère inédit de leurs travaux, d'autres publics plus novices sont guidés dans leurs découvertes par les différentes formes de recommandation existantes et en ont un réel besoin. L'intégration d'un système de recommandation par algorithme au moteur de recherche d'un site tel que Gallica doit satisfaire la condition de répondre aux exigences de ces deux types de publics opposés.

Par ailleurs l'effet « boîte noire » des algorithmes dont l'explicabilité n'est pas toujours garantie suscite la méfiance et la crainte des usagers vis-à-vis d'un système qui pourrait les manipuler en utilisant leurs données à leur insu, au-delà même des informations personnelles dont la plate-forme a véritablement besoin pour fonctionner. La liberté de choisir les critères sur lesquels reposent la recommandation de l'algorithme est identifiée comme un moyen de pallier le manque de transparence et d'augmenter la satisfaction et la confiance des usagers dans le système de recommandation (Paraschakis, 2017).

Bulles de filtre, enfermement et diversité

Selon la manière dont l'algorithme est conçu, le système de recommandation peut conduire à dissimuler une partie de l'offre auprès des utilisateurs et à les enfermer dans leurs goûts supposés par l'historique de leur navigation. L'enfermement dans des bulles de filtre entraîne un risque de standardisation et d'appauvrissement de la recherche. En effet, si les usagers, selon leur profil, se voient proposer les contenus les plus consultés, les chercheurs aux champs de recherche similaires auront les mêmes propositions, peu à même de les emmener vers d'autres domaines de pensée que celui dans lequel ils ont l'habitude d'évoluer (Farchy, Méadel et Anciaux, 2017, p. 194). Les observations de terrain¹⁴ et plusieurs études ont montré que la diversité a un impact positif sur la satisfaction des utilisateurs. Elle peut cependant susciter leur méfiance et l'incompréhension si le lien avec les préférences exprimées se fait trop distant (Castagnos et al. 2014).

Par ailleurs les pratiques de recherche et les attentes d'un chercheur peuvent évoluer d'une connexion à l'autre : en fonction du stade d'avancement de ses recherches un chercheur peut par exemple avoir un besoin accru de diversité plutôt que de pertinence dans ses résultats de requêtes. L'enfermement des utilisateurs dans des profils types non évolutifs présente le risque de cesser très vite de satisfaire leurs attentes. L'évolutivité des pratiques des utilisateurs doit pouvoir être prise en compte par les algorithmes à l'œuvre dans les systèmes de recommandation.

Principes éthiques de conception du système de recommandation : différents positionnements possibles pour répondre à ces enjeux

Ethique par conception

La prise en compte des enjeux éthiques est complexe et nécessite une vision globale du projet dans la mesure où les écueils se présentent à chaque étape de développement du système de recommandation : profilage des utilisateurs, conception de l'algorithme, création

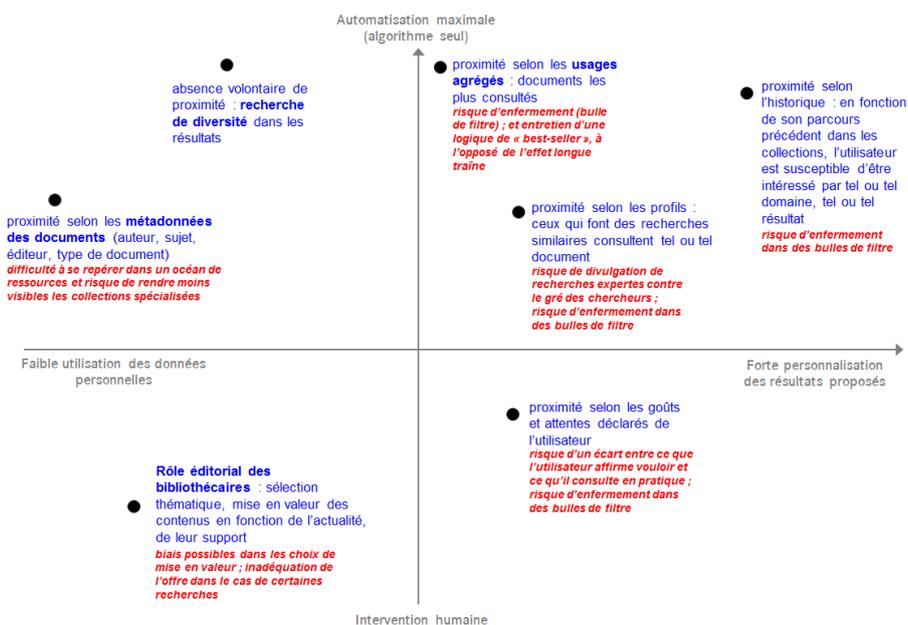
¹⁴ La femme citée en première partie, qui exprimait un « bémol » quant à son exploration de Gallica, le nuancit aussitôt en ajoutant « C'est l'occasion de découvrir des secteurs méconnus ou inconnus ».

de l'interface utilisateur, expérimentations et tests. Le travail sur l'anonymisation des données, le paramétrage de l'algorithme en lui-même (insertion d'une part de hasard, de « pas de côté » dans les suggestions faites, voire de propositions à l'opposé des goûts supposés de l'utilisateur ; exploitation des données de consultation à un niveau agrégé¹⁵) sont autant de moyens de se prémunir des risques éthiques à l'œuvre aux différents stades de développement du système de recommandation. L'intégration de métriques éthiques dans les indicateurs d'évaluation des systèmes de recommandation fait également partie d'une démarche éthique « by design ». Par exemple, le ministère de la Culture a testé une forme de recommandation à l'occasion de l'expérimentation d'un pass Culture¹⁶ à destination des jeunes de 18 ans. Pour évaluer l'algorithme de ce pass, une métrique de diversité a été associée à une métrique de pertinence. Les quatre indicateurs de cette métrique permettent de quantifier la découvrabilité et évaluer l'efficacité de l'algorithme en matière de diversification des contenus proposés.

La place de l'humain dans les systèmes de recommandation : l'utilisateur objet et sujet de la recommandation

Dans le cas de l'implémentation d'un système de recommandation sur un site de diffusion de contenus culturels, le recours à un algorithme de recommandation n'a pas vocation à se substituer au rôle de conseil éditorial du bibliothécaire. Dans les calculs de similarité utilisés par l'algorithme pour proposer des recommandations à l'utilisateur, plusieurs métriques de proximité peuvent être prises en compte que l'on peut situer sur un diagramme à deux axes en fonction du degré d'autonomie laissé à l'algorithme et de la personnalisation de la recommandation proposée à l'utilisateur (voir figure ci-dessous).

Diagramme représentant plusieurs leviers possibles pour paramétrer un algorithme de recommandation, ainsi que leurs principales limites d'un point de vue éthique
en abscisse : le degré d'utilisation des données personnelles des usagers
en ordonnée : le degré d'intervention humaine dans le processus de recommandation



¹⁵ Nuancer par exemple l'application des règles : si seul un utilisateur adopte une pratique, fait un choix donné parmi une liste de résultats, on considérera qu'il s'agit d'une pratique particulière. En revanche, si vingt personnes ont le même comportement, on considérera qu'il s'agit d'un comportement susceptible d'être partagé

¹⁶ Ministère de la Culture, *Synthèse relative au bilan de l'expérimentation du pass Culture*, mars 2021, https://pass.culture.fr/wp-content/uploads/2021/05/0321_SYNTHESE_RAPPORT-EXPERIMENTATION.pdf

Ce graphique met en évidence le fait que le risque d'enfermement dans des bulles de filtre est plus grand lorsque l'algorithme utilise le plus les données personnelles de l'utilisateur et quand le rôle éditorial des bibliothécaires n'est pas mobilisé. Il suggère aussi que ce risque est d'autant moins perçu par l'utilisateur que ses données sont collectées sans qu'il lui ait été demandé de déclarer des préférences. Au-delà d'une dichotomie entre la recommandation « manuelle » des bibliothécaires par sélections thématiques et la recommandation automatisée par algorithme qui rend possible une personnalisation très poussée si les données requises sont mobilisées, il pourrait aussi être envisagé un entre-deux hybride où les bibliothécaires se serviraient des recommandations algorithmiques issues de données agrégées comme aide à l'éditorialisation des contenus sur la plate-forme¹⁷, sans que l'utilisateur soit directement confronté à l'algorithme.

La sensibilisation passe par le design de l'interface

La réponse aux écueils liés à la collecte, au stockage et à l'exploitation des données personnelles et la confiance que l'utilisateur accorde au système de recommandation passent en partie par le design de l'interface et son ergonomie. Le fait de nommer explicitement les critères de recommandation (« les personnes intéressées par ce titre ont aussi consulté ceci... »¹⁸, « des suggestions surprenantes », « à l'opposé de vos recherches », « ce que la BnF conseille à tout le monde »), de laisser à l'utilisateur la liberté de choisir le degré de pertinence ou diversité qu'il souhaite obtenir dans ses résultats ou de paramétrer et mettre à jour ses préférences (possibilité de nuancer le poids de ses préférences dans la liste des résultats qui lui sont proposés) voire de proposer deux modes de navigation avec ou sans recommandation sont autant de pistes à explorer qui mettent en jeu l'ergonomie d'utilisation dans l'interface.

La technologie des systèmes de recommandation soulève des interrogations éthiques liées à la conception et l'utilisation de ces systèmes. Ne pas tenir compte de ces interrogations risque de susciter la crainte et la méfiance de l'utilisateur. C'est pourquoi les enjeux éthiques doivent être pris en compte dès la phase de conception du projet et soulèvent des questionnements relevant aussi bien du domaine de l'ingénierie et de la science des données, que du design d'interface et des sciences humaines.

En pratique, à la BnF : entre changement de paradigme et exemplarité

La BnF prévoit de se lancer, à partir de 2023, dans l'élaboration d'un prototype testable utilisant des mécanismes éthiques de recommandation par algorithme. Elle déploiera trois approches pour y parvenir : un travail d'ingénierie (constitution des données, entraînement des algorithmes), le prototypage à proprement parler (en appliquant les principes du design utilisateur – ou UX design), et une analyse des premiers usages s'appuyant sur un regard sociologique.

¹⁷ C'est vers cette solution que s'oriente actuellement l'INA en mobilisant des technologies d'intelligence artificielle pour renforcer ses missions de valorisation et de conservation du patrimoine audiovisuel. Cf Poupeau, Gautier, « En quoi l'intelligence artificielle constitue pour l'INA une opportunité pour renforcer ses missions de valorisation et de conservation du patrimoine audiovisuel ? », Présentation lors de la conférence « Futurs fantastiques », décembre 2021, <https://www.bnf.fr/fr/mediatheque/en-quoi-lintelligence-artificielle-constitue-pour-lina-une-opportunit%C3%A9-pour-renforcer>

¹⁸ Préférable à « autres titres susceptibles de vous intéresser », qui est moins clair pour l'utilisateur

Comment les enjeux énoncés précédemment peuvent-ils se traduire concrètement à la Bibliothèque ? Comment envisager le développement de ce projet du point de vue du service rendu, des citoyens qui sont susceptibles d'en profiter, mais aussi de son personnel et de ses partenaires ? De même que la mise en ligne massive de contenus en mode texte a considérablement modifié les pratiques de recherche, l'implémentation d'un système de recommandation conduit en effet la Bibliothèque à changer de paradigme en matière d'offre de recherche et à bouleverser en profondeur les stratégies des chercheurs. Plusieurs solutions seront conjuguées, d'ordre institutionnel, organisationnel, technique, législatif.

Une inscription dans une politique globale

Tout d'abord, comme nous l'avons expliqué, la recommandation personnalisée ne saurait faire disparaître les autres formes de médiation, pas plus qu'elle ne se veut un substitut à la mise en œuvre de nouvelles méthodes de fouille ou à l'évolution du moteur de Gallica vers la recherche visuelle.

En outre, elle figure dans plusieurs feuilles de route de la Bibliothèque qui garantissent le développement d'une politique cohérente et respectueuse des missions fondamentales de l'institution : le programme MISAOA (Mutualisation et innovation pour la sauvegarde et l'accès aux œuvres audiovisuelles), qui concerne le dépôt légal dématérialisé, inclut une amélioration de l'environnement numérique de l'utilisateur et de son expérience via des recommandations « intelligentes » de documents audiovisuels ; le Contrat d'objectifs et de performance 2022-2026 de la Bibliothèque répond aux grandes lignes de sa [feuille de route sur l'intelligence artificielle](#), dont le projet de recommandation est un projet majeur. Cette feuille de route a déjà donné lieu à quelques ateliers et réunions internes sur l'éthique dans le but de prendre du recul, de mieux identifier les points de vigilance communs aux projets, et de rassembler des bonnes pratiques. Cette sensibilité s'inscrit dans une politique responsable à l'égard du numérique en général (sobriété, développement d'indicateurs de mesure de l'impact du numérique, politique vertueuse en matière d'usage des cookies, absence de demande d'authentification pour accéder à la très grande majorité des services et contenus en ligne...).

Au-delà de ses préoccupations internes, de gestion, la BnF traduit également sa mission de service public par un engagement dans la formation de l'esprit critique et dans l'éducation aux médias (conférences tous publics sur l'intelligence artificielle¹⁹, collaboration avec les partenaires pédagogiques...). Faire preuve d'éthique dans la conception des services numériques en tant que tels ne suffit pas : la confiance des publics envers l'institution est renforcée par l'ensemble de ces actions, et le recours à un système de recommandation personnalisée par une institution publique répond à un enjeu citoyen de sensibilisation aux effets de l'algorithme sur la manière de découvrir des contenus. L'ergonomie d'une plateforme de service public et le paramétrage de l'algorithme constituent des leviers de sensibilisation en ce qu'ils rendent explicites, explicables et contrôlables les objectifs assignés au système de recommandation qui peuvent être multiples et évolutifs (gagner du temps en allant au résultat le plus pertinent ou prendre le temps du pas de côté pour « flâner » dans les collections).

¹⁹ Voir par exemple la conférence intitulée « Y a-t-il une intelligence artificielle ? », avec Serge Abiteboul, Nicolas Sabouret, et Marie-des-Neiges Ruffo de Calabre, BnF, 20 janvier 2022, <https://www.bnf.fr/fr/agenda/y-t-il-une-intelligence-artificielle>

Enfin, la BnF s'inspire de la politique gouvernementale, en particulier telle qu'elle est déclinée par son ministère de tutelle, le ministère de la Culture. Dans l'énoncé de sa stratégie numérique (Ministère de la Culture, 2022), le ministère insiste sur l'intérêt d'une attention constante aux usages²⁰, sur l'enjeu de la découvrabilité des contenus culturels en ligne²¹, et alerte sur les risques que certains algorithmes font courir à la diversité culturelle²², dont dépend l'avenir de la création artistique. Sur le plan international, les actions à prioriser identifiées par la France et le Québec – pour ne citer que cet exemple - dans leur stratégie commune en matière de découvrabilité des contenus culturels francophones consistent d'une part à « renforcer les actions visant à développer les compétences numériques et à vulgariser le concept de découvrabilité » et d'autre part à « sensibiliser les citoyens, afin qu'ils deviennent des consommateurs avertis, aux enjeux posés par le fonctionnement des grandes plateformes sur la diversité de leurs pratiques culturelles »²³. Sur les questions de l'éthique et de la responsabilité numérique, de nombreux organismes internationaux sont également sources d'inspiration : l'UNESCO²⁴, l'Union européenne²⁵, etc.

Expérimenter, évaluer

Conduites en toute transparence sur une interface dédiée à cet effet, avec des utilisateurs informés et sans générer de préjudice pour ces utilisateurs²⁶, les expérimentations permettront de mesurer la pertinence des résultats obtenus par le système de recommandation, l'intelligibilité du service pour l'internaute, son adéquation, et d'anticiper ses effets sur le long terme (est-il par exemple susceptible d'avoir des répercussions sur la politique de numérisation en tant que telle ?). Elles ne chercheront pas à tout englober à la fois mais se concentreront sur quelques aspects, suffisamment divers pour couvrir l'ensemble des problématiques susceptibles de se poser : les algorithmes, l'interface... Une réflexion sur les métriques d'évaluation des modèles et plus généralement l'introduction d'étapes de validation à différents moments du développement du projet constitueront un garde-fou essentiel.

²⁰ Ministère de la Culture, 2022, p. 9 : « La compréhension des usages et pratiques est (...) une démarche continue : elle encourage à penser et anticiper les comportements numériques à venir dans une logique d'amélioration constante et en complémentarité avec l'évolution des offres des lieux culturels. »

²¹ Ministère de la Culture, 2022, p. 12 : « La numérisation et la mise en ligne ne suffisent (...) pas à elles seules à donner accès aux œuvres : l'enjeu, dans certains secteurs, n'est plus de trouver, mais de choisir. L'abondance de certaines offres et la fragmentation des audiences nécessitent de capter l'attention et de faciliter le repérage. »

²² Ministère de la Culture, 2022, p. 18 : « Les méthodes et algorithmes de sélection sur les plateformes de diffusion représentent un éventuel frein à la diversité culturelle alors même qu'ils peuvent être une opportunité pour la soutenir : une attention portée à la définition et la régulation de ces algorithmes est donc une condition de la protection de la diversité culturelle et d'une meilleure rémunération des créateurs. »

²³ Ministère de la Culture (France) et Ministère de la Culture et des Communications (Québec), *Mission franco-québécoise sur la découvrabilité en ligne des contenus culturels francophones : stratégie commune*, 2020, <https://cdn-contenu.quebec.ca/cdn-contenu/adm/min/culture-communications/publications-adm/strategie/Decouvrabilite-Strategie-commune.pdf?1606736672> (objectif 1.3 : sensibiliser le grand public)

²⁴ UNESCO, *L'Intelligence artificielle au service du développement durable: défis et perspectives pour les programmes de science et d'ingénierie de l'UNESCO*, mai 2019, https://unesdoc.unesco.org/ark:/48223/pf0000368028_fre.locale=fr

²⁵ Voir *The Artificial Intelligence Act*, <https://artificialintelligenceact.eu/>

²⁶ Pour en savoir plus sur l'éthique des expérimentations, voir Paraschakis, 2017, en particulier p. 5-6 (« Online experiments »)

Partager, et s'ouvrir à la diversité

Dans l'attente du moment magique où, une fois opérationnels et fiables, les algorithmes de recommandation ainsi mis en œuvre pourront eux-mêmes être diffusés à tous selon une logique d'ouverture, de transparence et de mutualisation des développements, et bien que son début officiel ne soit pas prévu avant 2023, ce projet est déjà placé sous le signe du partage et se met en quête de partenaires intéressés par tout ou partie de ses développements : chercheurs confirmés et jeunes professionnels, professeurs et post-doctorants, bibliothécaires techniciens et non techniciens, responsables de bibliothèques numériques, partenaires de Gallica, juristes, garants de la sécurité informatique (pour éviter que les données personnelles, si elles sont collectées, ne soient piratées)... La BnF a la chance de rassembler un grand nombre de ces profils, mais elle a également besoin de regards extérieurs qui partagent cet enjeu de la découvrabilité des contenus, en particulier des contenus non anglophones. Les projets européens, les projets internationaux représentent un cadre très favorable pour cultiver une éthique de la recommandation.

Conclusion

Au cours de cette réflexion, nous avons retrouvé un couple classique, celui de la demande et de l'offre, qui préside à l'élaboration des services en bibliothèque. L'introduction de la notion de personnalisation assistée par l'intelligence artificielle permet de le considérer sous un jour nouveau. En effet, si la mission d'une bibliothèque – et de la BnF en particulier – est de mettre en valeur et de développer la diversité, l'ouverture, la curiosité, elle ne peut se contenter d'une logique de demande et doit également développer son offre. Ne pas le faire expose la bibliothèque au risque de perdre ce qui représente l'intérêt même de la constitution de ses collections, si exhaustives et encyclopédiques soient-elles, et de sa politique documentaire (ou de numérisation), soit le fondement même de son existence. Ne pas le faire rend inutile l'entreprise de patrimonialisation qui vise à ouvrir les usages à des terres inconnues : dans le sillage des publications éphémères ou négligées des siècles passés, patiemment collectées par la bibliothèque, les nouveaux objets qui entrent aujourd'hui dans son périmètre (tweets, billets de blogs, archives de l'internet en général, NFT...) sont susceptibles de rejoindre un jour les collections courantes de la bibliothèque numérique sans qu'une demande explicite ou généralisée ne les appelle dans un premier temps. Dans une vision étriquée de la bibliothèque, focalisée sur la demande et sur la satisfaction de souhaits clairement exprimés, le bibliothécaire se changerait en simple statisticien, analyste de traces d'usage et de comportements.

Dans ce couple de l'offre et de la demande, la place de l'utilisateur prend un relief nouveau. L'intégration d'un système de recommandation au moteur de recherche permet à l'offre de s'adapter en temps réel à la demande de chaque utilisateur spécifique. L'enjeu est alors de rendre l'utilisateur conscient de son nouveau statut, désormais partenaire actif et indispensable d'une médiation de qualité dont il cesse d'être uniquement le destinataire. Il doit être respecté à ce titre, revêt une humanité riche, nous oblige. L'histoire qui commence s'annonce riche, mais aussi exigeante : au nom de la même diversité qui s'applique au champ documentaire, il s'agit de ne pas accentuer la fracture numérique entre les utilisateurs experts qui ont beaucoup lu, beaucoup exploré, et ceux qui débutent, lesquels doivent, tout autant que les premiers, se voir ouvrir les portes d'une richesse documentaire enthousiasmante.

Bibliographie

Bastard, Irène, *Les publics de la BnF - Synthèse de l'Observatoire 2020*, 2022, https://www.bnf.fr/sites/default/files/2022-01/Les_publics_de_la_BnF_Synthese_de_l_observatoire_2020.pdf

Bermès, Emmanuelle et al., « Améliorer l'accès aux données des bibliothèques sur le web : l'exemple de data.bnf.fr », 2016, <http://library.ifla.org/id/eprint/1447/1/081-bermes-fr.pdf>.

Bibliothèque nationale de France, *La BnF et l'intelligence artificielle*, page du site bnf.fr, 2021, <https://www.bnf.fr/fr/feuille-de-route-ia>

Cardon, Dominique, *A quoi rêvent les algorithmes : nos vies à l'heure des big data*, Seuil, 2015

Castagnos. Sylvain, Brun. Armelle, Boyer. Anne, La diversité : entre besoin et méfiance dans les systèmes de recommandation. *Revue I3 - Information Interaction Intelligence*, Cépaduès, 2014. <https://hal.inria.fr/hal-01108998>

Comité national pilote d'éthique du numérique, *Pour une éthique du numérique*, PUF, 2022

DeLeon,Haley, "The Ethical and Privacy Issues of Recommendation Engines on Media Platforms", *Towards Data Science*, 2019, <https://towardsdatascience.com/the-ethical-and-privacy-issues-of-recommendation-engines-on-media-platforms-9bea7bcb0abc>

Farchy, Joëlle, Méadel, Cécile et Anciaux, Arnaud, « Une question de comportement. Recommandation des contenus audiovisuels et transformations numériques », *tic&société* , Vol. 10, N° 2-3 | 2ème semestre 2016 - 1er semestre 2017, <http://journals.openedition.org/ticetsociete/2136>

Harris, Jeremie, "Ethical challenges of recommender systems: Silvia Milano on the tough questions we need to explore to build beneficial algorithms", *Towards Data Science*, 2021, <https://towardsdatascience.com/ethical-problems-with-recommender-systeems-398198b5a4d2>

Majeed, Abdul, Lee, Sunchang, "Anonymization Techniques for Privacy Preserving Data Publishing: A Comprehensive Survey," in *IEEE Access*, vol. 9, pp. 8512-8545, 2021, doi: 10.1109/ACCESS.2020.3045700.

Ministère de la Culture, *Politiques culturelles : la stratégie numérique du ministère de la Culture*, janvier 2022, <https://www.culture.gouv.fr/Media/Thematiques/Innovation-numerique/fichiers/La-strategie-numerique-du-ministere-de-la-Culture>

Nouvellet, Adrien, Beaudouin, Valérie, et al., « Analyse des traces d'usage de Gallica : Une étude à partir des logs de connexions au site Gallica », 2017, <https://hal.archives-ouvertes.fr/hal-01709264> et document : <https://hal.archives-ouvertes.fr/hal-01709264/document>

Paraschakis, Dimitris, "Towards an Ethical Recommendation Framework", Conference: 11th IEEE International Conference on Research Challenges in Information Science (best paper award), 2017, https://www.researchgate.net/publication/316966805_Towards_an_Ethical_Recommendation_Framework

Pernoo, Marianne, « Quelles classifications et quels classements pour les oeuvres de fiction dans les bibliothèques ? : la question des frontières », *Bulletin des bibliothèques de France (BBF)*, 2001, n° 1, p. 47-53, <https://bbf.enssib.fr/consulter/bbf-2001-01-0047-003> ISSN 1292-8399

Roth, Camille et Poiroux, Jérémie, « L'écriture guidée du code : Le cas des algorithmes de recommandation », *RESET*, 2022, <https://journals.openedition.org/reset/3429>