

RAMEAU, LES CATALOGUES, LE WEB

MICHEL MINGAM

Responsable du Centre national Rameau,
Département de l'information bibliographique et numérique
Bibliothèque nationale de France

Voici une dizaine d'années, un précédent article¹ tentait d'ouvrir quelques perspectives, en faisant le pari, plutôt que de céder au vertige de la table rase, que les bouleversements qui s'annonçaient (apparition des moteurs de recherche, possibilités offertes par le plein texte) ne priveraient pas nécessairement des outils documentaires plus traditionnels et, par exemple, les langages d'indexation comme Rameau. Qu'en est-il aujourd'hui? Le point d'étape que nous proposons s'efforcera de mesurer le chemin parcouru, dans un contexte en constante évolution, gros de défis mais aussi de promesses.

UN CONTEXTE PORTEUR

DES UTILISATEURS DE PLUS EN PLUS NOMBREUX

Premier constat : l'audience de Rameau ne cesse de croître. Ses utilisateurs sont en effet de plus en plus nombreux en France, au point de former aujourd'hui une nébuleuse qu'il serait vain de prétendre quantifier. Au cœur de cet ensemble, se détache toutefois un noyau facilement mesurable : le réseau que forment les établissements qui ont choisi, non seulement d'utiliser le langage Rameau, mais encore de contribuer à son enrichissement en formulant des propositions terminologiques au moyen d'un outil de dialogue en ligne, le Fichier national des propositions Rameau. Or, le nombre de ces établissements n'a lui-même cessé d'aug-

menter dans de notables proportions, passant de 186 en 2004 à 295 en 2014.

En s'étoffant, le réseau national s'est naturellement davantage diversifié. Trois grandes composantes le structurent : les départements de la Bibliothèque nationale de France; le réseau du Sudoc (bibliothèques d'universités et de grands établissements scientifiques); la lecture publique, depuis les bibliothèques des grandes villes jusqu'à des médiathèques plus modestes (du fait, notamment, de la possibilité désormais offerte à ces établissements de récupérer les notices bibliographiques de la BnF, indexées avec Rameau). Sont venus s'ajouter depuis lors d'autres profils de partenaires : bibliothèques de grandes institutions, musées, services documentaires spécialisés, organismes privés, etc. Franchissant nos frontières, le rayonnement de Rameau se manifeste également dans les pays francophones (Europe, Afrique du Nord et subsaharienne, Proche-Orient) : en témoignent la participation au Fichier national des propositions Rameau de nombreuses bibliothèques de la communauté francophone de Belgique ainsi que celle du réseau suisse romand RERO (représentant plus de 260 établissements de tous types); un certain nombre de missions de formation à Rameau ont par ailleurs été conduites ces dernières années dans les pays arabes (Maroc, Tunisie, Égypte, Syrie). Mentionnons enfin, au-delà du monde francophone, l'adaptation de Rameau en polonais et en roumain. L'accroissement et la diversification du réseau national, avec ses prolongements internationaux, témoignent donc d'un dynamisme certain, au service de notre langue et de la francophonie.

¹ Michel Mingam. « Rameau : bilan, perspectives », *Bulletin des bibliothèques de France*, 2005, n° 5. Consultable en ligne : <http://bbf.enssib.fr/consulter/bbf-2005-05-0038-001>

UN RÉSEAU STRUCTURÉ

Tout en se développant, le réseau Rameau n'a cessé de renforcer son organisation sur le plan national. Une convention, renouvelée en 2011 et associant les ministères de la Culture et de l'Enseignement supérieur ainsi que la BnF et l'Agence bibliographique de l'enseignement supérieur, confie la gestion du langage et l'administration du réseau à la BnF, à travers le Centre national Rameau, au sein du Département de l'information bibliographique et numérique. Elle prévoit également la réunion à dates régulières d'instances où siègent des représentants des tutelles et des utilisateurs (comité d'orientation, comité opérationnel). Enfin, des Journées nationales Rameau, organisées tous les trois ans à la BnF, sont l'occasion d'échanges directs entre les membres du réseau (la dernière en date a réuni 170 participants le 6 novembre 2014). Dans ce contexte, plusieurs outils ont été progressivement mis en place au service des uti-

lisateurs de Rameau : un site web Rameau, régulièrement actualisé, qui fournit un certain nombre d'informations et donne accès à différentes fonctionnalités ainsi qu'à une messagerie; la consultation du réservoir terminologique via le catalogue général de la BnF; la consultation du *Guide d'indexation RAMEAU* proposée en ligne depuis 2009; le Fichier national des propositions Rameau; un répertoire national de formateurs ainsi que des supports pédagogiques en ligne pour faciliter la formation et l'autoformation à l'indexation Rameau; la possibilité offerte aux établissements de récupérer les données Rameau pour leurs catalogues.

DE NOUVEAUX HORIZONS

À l'origine, Rameau a été conçu et développé au service des catalogues des bibliothèques et autres établissements documentaires. De fait, l'essentiel de ses utilisateurs, comme la quasi-totalité des membres du réseau Rameau dont il

vient d'être question, utilisent le langage dans ce contexte. Cela n'est bien entendu pas près de s'arrêter, alors même que les catalogues sont amenés à évoluer en profondeur : Rameau demeurera naturellement à leur service, selon sa mission première.

Mais, parallèlement, des perspectives inédites se sont récemment ouvertes : en effet, depuis 2012, grâce au site *data.bnf.fr*, la BnF propose la possibilité de récupérer et de réutiliser les autorités Rameau selon les standards du web. Étape décisive, qui offre au référentiel une visibilité sans précédent ainsi qu'un brusque élargissement de ses usages, sur la « nouvelle frontière » du web sémantique et du web de données, dont Rameau devient un acteur à part entière.

Précisons d'emblée qu'entre la vocation traditionnelle de Rameau, au service des catalogues, et ces nouveaux horizons, il ne faut voir aucune incompatibilité mais, tout au contraire, une dialectique féconde en train de s'inventer.

ENRICHISSEMENT DE RAMEAU

UN RÉSERVOIR TERMINOLOGIQUE EN CONSTANT ACCROISSEMENT

Un réseau de contributeurs en expansion ne peut que se refléter au niveau du langage lui-même, puisque ce sont ces derniers qui, à travers leurs propositions, participent au premier chef à l'enrichissement de Rameau.

De fait, d'un point de vue strictement quantitatif, l'augmentation régulière du vocabulaire ne laisse aucun doute : de 145 000 en 2004, on est ainsi passé à 176 000 descripteurs en 2014 (dont 107 000 noms communs et 59 000 noms géographiques). Un tel résultat témoigne, pour une part essentielle, de l'activité soutenue du Fichier national des propositions Rameau, puisqu'on dénombre à ce jour plus de 23 000 propositions depuis son lancement en juin 2000.

D'un point de vue plus qualitatif, il convient de noter que la diversité du réseau permet également un large éventail de propositions terminologiques, depuis les plus généralistes jusqu'aux plus spécialisées, selon les besoins d'indexation d'établissements partenaires extrêmement variés (établissements scientifiques aussi bien que bibliothèques de lecture publique, voire

musées) pour des types de documents eux-mêmes très hétérogènes (textes, sons, images, objets). Rameau se trouve ainsi en mesure de refléter assez fidèlement l'évolution conceptuelle de la documentation indexée.

Sans relever à proprement parler du web social, Rameau a développé de la sorte un modèle collaboratif efficace : le langage Rameau est bien l'« œuvre collective » de ses utilisateurs, comme le mentionne la convention nationale de 2011, et il s'enrichit à proportion de leur nombre et de leur dynamisme. Ce modèle combine réactivité (les propositions terminologiques reçoivent une réponse rapide), mutualisation (l'ensemble du réseau profite des enrichissements du vocabulaire, avec les économies afférentes), transparence (le fichier des propositions peut être consulté librement par quiconque sur le site Rameau) et garantie institutionnelle (le réseau s'appuie sur des acteurs publics).

ENCYCLOPÉDISME, EXHAUSTIVITÉ, GRANULARITÉ

Encyclopédique, le vocabulaire Rameau a vocation à couvrir tous les domaines du savoir et des pratiques. Pour autant, il n'est évidemment pas exhaustif : la prétention serait absurde pour un référentiel de cette nature. De plus, comme on vient de le voir, loin d'être constitué *a priori*, il s'enrichit au contraire progressivement en fonction des besoins d'indexation de ses utilisateurs, ce qui introduit un double filtre :

- d'une part, un filtre documentaire, puisqu'un descripteur n'est créé que s'il est justifié par l'indexation d'un document précis, quel qu'il soit : Rameau reflète donc le monde à travers les documents qui en parlent ;
- d'autre part, un filtre humain puisque la création du descripteur suppose par ailleurs que l'indexeur, en présence du document qui la justifierait, travaille dans un établissement inscrit au Fichier national des propositions Rameau et fait l'effort, ou a le temps, d'en soumettre la proposition : Rameau reflète donc également l'importance plus ou moins grande accordée à l'indexation dans les établissements documentaires.

Mais, à défaut d'une utopie exhaustive, Rameau, conformément à sa vocation encyclopédique, peut en revanche être enrichi à l'infini, et cette souplesse, cette capacité à s'adapter

à pratiquement toutes les demandes, constituent un atout précieux qui lui permet de gagner continûment de nouveaux utilisateurs. La question autrefois soulevée d'une granularité minimale en deçà de laquelle la précision du vocabulaire ne devrait pas descendre se révèle, dès lors, infondée. En effet, seule la possibilité de cerner précisément un profil de public particulier aurait pu la justifier : or, ainsi qu'on vient de le noter, Rameau est, au contraire, au service d'établissements et d'acteurs documentaires extrêmement divers, s'adressant eux-mêmes à des publics très variés. Face à un tel défi, le référentiel ne peut donc, logiquement, que s'organiser selon un dégradé qui part des termes les plus généraux formant le noyau du vocabulaire courant (celui du *Petit Larousse*) pour aboutir insensiblement aux notions les plus spécialisées, les plus scientifiques, à mesure que l'on descend dans des hiérarchies conceptuelles de plus en plus précises, chacun de ces niveaux pouvant ensuite être complété à volonté en fonction des besoins, les quelques restrictions prévues par Rameau étant en la matière extrêmement libérales.

Remarquons à ce propos que l'enrichissement du vocabulaire dans le sens de la précision n'est pas le seul fait des bibliothèques universitaires et autres contributeurs scientifiques. Bien entendu, l'indexation des thèses, notamment, est l'occasion de propositions terminologiques portant sur des notions extrêmement fines

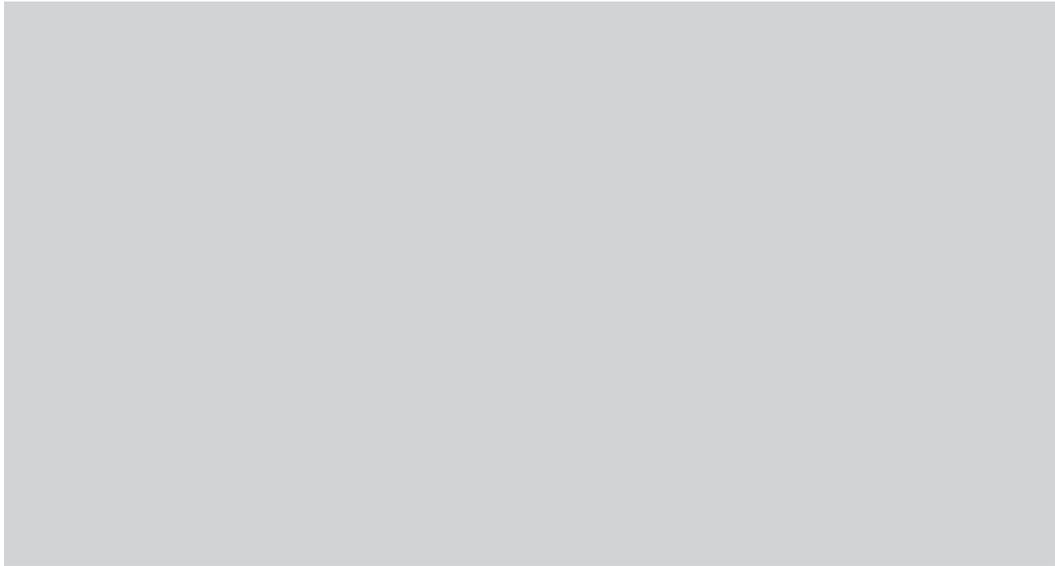
(par exemple, des virus, des bactéries, des enzymes, etc.). Mais les établissements de lecture publique peuvent également, à l'occasion, faire des propositions conduisant à détailler un domaine : ainsi, les demandes des discothèques publiques ont-elles conduit à décliner dans Rameau de très nombreux genres de musique rock (country-rock, punk rock, rock garage, rock planant, rockabilly, etc.).

MUTATION DE RAMEAU

LE CHANTIER DE RÉVISION SYSTÉMATIQUE DU VOCABULAIRE

Rameau est donc un langage qui se développe et s'enrichit. Mais c'est également un langage qui se transforme en profondeur depuis quelques années : le Centre national Rameau a en effet ouvert en 2007 un ambitieux chantier de révision générale des autorités Rameau, centré sur le vocabulaire.

Traditionnellement, depuis l'origine, le caractère précoordonné de Rameau avait prédominé : Rameau était avant tout conçu comme un ensemble de règles permettant de construire syntaxiquement des vedettes-matière ; l'essentiel de la réflexion, des débats et des efforts était mobilisé par cette question, au détriment des autres dimensions du langage. Or, Rameau, loin de se réduire à la syntaxe, se présente



également, et peut-être surtout, comme un très riche réservoir terminologique contrôlé et structuré. Il importait donc de rééquilibrer le point de vue au profit du vocabulaire, atout essentiel de Rameau, pour peu que la qualité soit au rendez-vous. D'où la décision prise de faire porter désormais l'effort principal, sans négliger pour autant la syntaxe, sur les aspects terminologique (gestion des ambiguïtés de la langue naturelle) et sémantique (gestion des relations conceptuelles entre les descripteurs), en se concentrant sur les notices d'autorité qui constituent le référentiel (car ce dernier ne se réduit pas à une simple liste de termes), dans le cadre d'une révision systématique du vocabulaire susceptible d'en améliorer significativement les performances.

La réflexion sur Rameau a ainsi contribué à faire bouger les lignes ; mais il se trouve que l'évolution rapide du contexte général est ensuite venue conforter cette nouvelle orientation. D'une part, en effet, la réflexion sur les catalogues et le catalogage, elle-même en plein développement dans le cadre des travaux de modélisation conduits sous l'égide de l'IFLA, tend aujourd'hui à mettre en valeur la notion de notice d'autorité. D'autre part, comme on vient de le noter, un nouvel

horizon s'ouvre depuis peu pour Rameau du côté de la communauté du web : or, il est clair que, pour cette dernière, ce sont les dimensions terminologique et sémantique de Rameau qui importent principalement.

Réflexion interne et environnement externe se conjuguent donc pour expliquer, et justifier, le chantier de révision systématique du vocabulaire engagé par le Centre national Rameau.

MÉTHODE

Le travail entrepris s'inscrit dans le cadre d'un programme à long terme, rétrospectif et méthodique : il s'agit de réviser progressivement l'ensemble des autorités Rameau en procédant d'une manière systématique, discipline par discipline, domaine après domaine, les noms communs étant traités en premier lieu (jusqu'alors, la plupart des interventions sur le vocabulaire, lorsqu'elles avaient lieu, étaient menées d'une manière dispersée et aléatoire). Seule une telle approche est susceptible, en effet, de donner les moyens d'améliorer véritablement, avec quelque chance de succès, la cohérence intellectuelle et formelle du vocabulaire, tout en permettant de mesurer avec préci-

sion l'avancement du travail. Or, un tel effort est d'autant plus nécessaire que le vocabulaire s'est développé d'une manière profuse au cours des années, alors même, on l'a dit, que l'attention était accaparée par les questions de syntaxe.

Une fois la nouvelle orientation validée, un vade-mecum méthodologique très détaillé a été mis au point collectivement par le Centre national Rameau en 2006-2007, afin d'encadrer et de guider le travail à venir sur les notices d'autorité². Le chantier a pu alors commencer, en juin 2007, avec une montée en charge progressive au sein de l'équipe du Centre national Rameau, chacun des huit experts étant responsable des disciplines correspondant à son profil et à sa formation. Dans chaque domaine à traiter, le vocabulaire est d'abord recensé le plus exhaustivement possible, ensuite mis à plat d'un point de vue conceptuel, enfin révisé systématiquement d'une manière rigoureuse et approfondie : cela peut se traduire par un remaniement du domaine, la modification d'intitulés de vedettes, la création, la suppression ou la fusion de notices, etc. Chaque notice revue reçoit une note à usage interne qui permet de signaler son nouveau statut et d'effectuer des comptages. Lorsqu'un domaine a été intégralement révisé, il est relu par un autre expert, donc avec un regard neuf, dans un souci de contrôle de la qualité et de vérification intellectuelle de la cohérence obtenue. Les domaines revus sont ensuite signalés sur le site Rameau.

Bien entendu, les difficultés ne manquent pas : volumétrie et caractère encyclopédique du vocabulaire (il est certainement moins difficile de travailler sur un répertoire spécialisé) ; nécessité de répondre à des besoins divers et parfois contradictoires (selon les types d'établissements et de documents) ; impossibilité de revenir sur certains choix antérieurs (notamment parce que cela impliquerait trop de corrections dans les indexations des catalogues) ; obligation de tenir compte de la précoordination ; enfin, *last but not least*, tension entre l'exigence de qualité et la nécessité d'avancer. Bien entendu, des choix ont dû être faits, qui peuvent être discutés. Bien entendu, il s'agit d'un travail humain, donc éminemment perfectible. Bien entendu, enfin, la progression du chantier doit être rapportée à l'effectif du Centre national Rameau : même motivées, huit personnes ne sauraient faire de miracles, d'autant moins qu'elles

doivent par ailleurs se partager entre différentes tâches. Néanmoins, les premiers résultats obtenus sont loin d'être négligeables.

PREMIERS RÉSULTATS

D'un point de vue quantitatif, 44 000 autorités Rameau ont été revues à ce jour dans le cadre de ce programme (soit plus de 40 % du total des noms communs). À ce chiffre correspondent de nombreux domaines disciplinaires achevés : architecture, arts décoratifs, peinture, musique, jeux et sports, philosophie, anthropologie et ethnologie, archéologie et préhistoire, linguistique et langues, anatomie humaine et animale, taxonomie des bactéries, des mammifères, des oiseaux, etc. Dans chacun d'entre eux, les notices revues apparaissent bien plus riches et rigoureuses lorsqu'on les compare à leur état antérieur (le résultat étant plus spectaculaire encore lorsqu'on considère les ensembles conceptuels eux-mêmes). Coexistent donc désormais, au sein du référentiel, deux états qualitatifs bien distincts, selon que les domaines ont été révisés ou qu'ils sont encore, en quelque sorte, « en friche ».

On le voit, un tel travail de fond, qui s'inscrit dans la durée, vise à améliorer d'une manière significative la richesse et la cohérence des données, à leur faire franchir un véritable seuil qualitatif, en se plaçant résolument dans la double perspective, désormais solidaire, des catalogues et du web. Dès à présent, les résultats obtenus permettent d'esquisser le nouveau « visage » de Rameau, tel qu'il se dessine progressivement sous l'effet du travail entrepris, dans ses dimensions terminologique et sémantique.

ÉVOLUTION TERMINOLOGIQUE

En tant que vocabulaire contrôlé, Rameau s'efforce de pallier autant que possible les ambiguïtés de la langue naturelle (homonymies, synonymies, polysémies) : l'objectif est de parvenir à un vocabulaire exact dans son contenu, homogène dans sa forme et univoque dans son utilisation (où un descripteur = un concept et un concept = un descripteur), tout en multipliant les points d'accès aux formes retenues à partir de formes rejetées grâce à des renvois d'équivalence. Le document méthodologique mis au point par le Centre national Rameau pour enca-

² Mémento RAMEAU : notices d'autorité noms communs. Consultable en ligne : http://rameau.bnf.fr/utilisation/docs_reference.htm

drer le travail de révision systématique du vocabulaire a fortement précisé les exigences sur ce point.

Le choix de la forme retenue obéit ainsi à des règles strictes : la vedette doit être à la fois exacte, pertinente, non ambiguë et, dans le respect de la langue française, conforme à l'usage tel qu'il est attesté en priorité dans la documentation en français (documents de référence généraux et spécialisés, études particulières...). La consultation de sources documentaires imprimées et/ou électroniques (d'abord françaises et francophones, mais également étrangères si nécessaire) est donc systématique : ce sont elles qui font foi en dernière instance, et non les répertoires auxquels Rameau est par ailleurs relié (Library of Congress Subject Headings, Répertoire de vedettes-matière Laval). Dans certains domaines, tels que la taxonomie, le souci scientifique peut être poussé très loin.

Les problèmes de synonymie sont l'objet d'une attention particulière. Si l'on admet que la fonction première des descripteurs dans un langage documentaire n'est pas tant d'« exprimer » le contenu d'un document que de le désigner, d'y donner accès, afin que l'utilisateur final puisse ensuite le trouver à partir des termes de sa requête, les formes rejetées acquièrent une importance égale à celle de la forme retenue, au sein d'un ensemble indissociable que l'on peut appeler la « grappe » ou le « bouquet » terminologique. Le travail de révision du vocabulaire entrepris par le Centre national Rameau s'efforce donc de multiplier ces relations d'équivalence, en essayant d'anticiper les types de requêtes susceptibles d'être formulées ; c'est également un moyen de concilier les exigences diverses des utilisateurs potentiels de Rameau, selon qu'ils relèvent du grand public ou de chercheurs spécialisés (les variantes communes et savantes pouvant devenir, selon les cas, des formes retenues ou rejetées). Bien entendu, cet effort est complété par les suggestions des contributeurs du réseau, attentifs à proposer des formes alternatives qu'ils repèrent dans la documentation indexée.

L'objectif est donc clairement défini, sinon toujours entièrement atteint : proposer un vocabulaire autant que possible rigoureusement établi, univoque dans ses acceptions, ménageant de multiples possibilités d'interrogation dans ses accès.

ÉVOLUTION SÉMANTIQUE

Vocabulaire contrôlé, le langage Rameau se présente également comme un réseau sémantique de descripteurs reliés entre eux d'une manière structurée : l'objectif est de permettre à l'utilisateur de naviguer au sein du vocabulaire entre les autorités, afin d'élargir (termes génériques), d'affiner (termes spécifiques) ou de réorienter (termes associés) sa recherche. Le fait n'est pas nouveau, mais le travail de révision en cours l'a infléchi dans deux directions.

D'une part, l'établissement des relations sémantiques entre les autorités Rameau épouse désormais plus étroitement les règles des thésaurus. Cela signifie qu'ont été bannies les relations hiérarchiques bancales (A ne peut plus avoir pour termes génériques B et C si B a lui-même pour terme générique C), à l'inverse de ce que pratiquent à l'occasion les LCSH et le répertoire Laval. D'une manière générale, les relations hiérarchiques génériques/spécifiques sont limitées au profit des relations associatives.

D'autre part, dans le respect de ces règles plus contraignantes, le réseau des relations sémantiques est développé au maximum, de telle sorte que le référentiel s'apparente progressivement à un ensemble de « thésaurus » partiels (puisque Rameau ne saurait être un réservoir terminologique exhaustif) reliés entre eux, au sein desquels il sera loisible de descendre toute la hiérarchie conceptuelle, du plus général au plus spécifique, et entre lesquels il deviendra à tout moment possible de circuler grâce aux termes associés (ou, lorsqu'elles existent, aux polyhiérarchies). Les domaines revus autorisent dès à présent ce type de parcours : ainsi est-il possible d'accéder par navigation à l'essentiel du vocabulaire concernant, par exemple, la philosophie ou la musique à partir des seules autorités « Philosophie » et « Musique » (à chaque fois, sont concernés plusieurs centaines de descripteurs) ; de même, les autorités « Mammifères » ou « Oiseaux » permettent de parcourir intégralement les taxonomies correspondantes dans Rameau. On peut donc parler d'un « devenir-thésaurus » de Rameau, du moins comme horizon régulateur.

Le Centre national Rameau souhaite aller plus loin encore en intégrant à terme la totalité ou quasi-totalité des autorités noms communs



dans ce dispositif, de telle sorte qu'il devienne possible d'entrer dans le référentiel par l'intermédiaire d'une quinzaine d'autorités correspondant aux grands domaines du savoir et de l'activité (philosophie, religion, droit, sciences sociales, sciences humaines, sciences, techniques, médecine, arts, littérature...) et formant autant de têtes de hiérarchies à partir desquelles se déploiera selon des axes verticaux et horizontaux l'ensemble du vocabulaire. Si, comme on vient de le voir, le fonctionnement sémantique de Rameau s'apparente d'une certaine manière à celui des thésaurus, il rejoindra aussi, de la sorte, celui des classifications, en s'en distinguant toutefois par une plus grande souplesse dans la gestion des réseaux sémantiques grâce aux relations associatives et aux polyhiérarchies.

Bien entendu, Rameau ne saurait prétendre se substituer ni aux thésaurus (spécialisés par nature, alors que lui-même est encyclopédique) ni aux classifications (dont la visée est davantage synthétique, là où lui-même privilégie une approche analytique). Il n'en reste pas moins que le développement en cours des potentialités sémantiques de Rameau ouvre des perspectives heuristiques extrêmement intéressantes pour l'utilisateur final, dans le contexte des catalogues aussi bien que dans celui du web.

ET LA SYNTAXE ?

La priorité accordée aux dimensions terminologique et sémantique du langage ne signifie pas que la syntaxe soit totalement oubliée (même si les autorités Rameau peuvent être utilisées

sans elle dans certains contextes, en particulier sur le web) : Rameau demeure un langage documentaire précoordonné qui propose, non seulement un « vocabulaire », mais encore une « grammaire » de règles syntaxiques permettant de construire, à l'indexation, des vedettes-matières à partir des autorités Rameau et d'autres types d'autorités (noms de personnes, collectivités, titres) ; et c'est bien ainsi qu'il est utilisé dans les principaux catalogues (BnF, Sudoc), la précoordination y apportant une aide précieuse pour limiter le bruit documentaire.

Simplement, les marges de manœuvre pour faire évoluer cette dimension du langage sont désormais fort réduites. En effet, on ne saurait envisager aujourd'hui de modifier profondément les règles syntaxiques en usage sans compromettre gravement la cohérence des catalogues, compte tenu de la masse documentaire considérable qui a été indexée au fil du temps selon ces principes : la précoordination s'inscrit dans une continuité déjà ancienne, qu'elle doit nécessairement respecter. À l'inverse du chantier entrepris sur le vocabulaire, le Centre national Rameau se limite donc à un travail de maintenance pour ce qui relève de la syntaxe : actualisation des règles exposées dans le *Guide d'indexation Rameau* ; enrichissement du stock des subdivisions lorsque nécessaire ; introduction de notes d'application dans les notices d'autorité pour aider les indexeurs.

Seule exception à cette règle : l'élargissement du champ d'application des subdivisions, qui permet de faciliter leur utilisation et donc le travail des indexeurs sans bouleverser l'économie des catalogues, doit être envisagé sans

timidité à chaque fois que possible, et aussi loin que possible, tout ce qui concourt à l'assouplissement et à la simplification de la syntaxe étant éminemment souhaitable, sans préjuger d'autres évolutions à venir, en relation avec la modélisation des données dans les catalogues.

USAGES DE RAMEAU

UTILISER RAMEAU

Tel se présente aujourd'hui, dans ses grandes lignes, Rameau. Reste la question de son utilisation : qu'en est-il des usages concrets ou potentiels des métadonnées Rameau (et des indexations qu'elles permettent de générer) dans les catalogues et sur le web? Dans quelle mesure la triple dimension de l'outil (langage contrôlé permettant la recherche par mots-clés; langage structuré permettant la navigation sémantique dans le vocabulaire; langage précoordonné permettant le feuilletage d'index) est-elle exploitée? À quelles conditions pourrait-elle l'être davantage?

Les réponses à ces questions n'engagent pas seulement Rameau : il appartient en effet aux différents dispositifs techniques existant ou à venir d'exploiter la matière première que constituent le référentiel et/ou les indexations, dans le cadre de choix politiques définis ou au gré d'initiatives innovantes. Cela revient à distinguer plusieurs ordres de réalités et plusieurs catégories d'intervenants qui, bien qu'en relation, ne sauraient pour autant être confondus :

- le référentiel lui-même, dont il vient d'être question, qui relève de la responsabilité des contributeurs du réseau pour son enrichissement et de celle du Centre national Rameau pour sa maintenance et son amélioration ;
- les indexations dans les catalogues, qui résultent du travail des catalogueurs, en fonction de politiques d'indexation définies par leurs établissements ;
- les dispositifs techniques susceptibles d'exploiter ce référentiel et/ou ces indexations, lesquels relèvent du savoir-faire de l'ingénierie documentaire et informatique (concepteurs de systèmes de gestion pour les bibliothèques, spécialistes du web sémantique...);
- enfin, les choix stratégiques présidant à l'utilisation de Rameau et à son exploitation technique, lesquels sont du ressort des ins-

titutions et des réseaux concernés, ou bien d'entreprises et d'individus particuliers, en fonction de leurs objectifs propres.

Ce préalable étant posé, essayons d'esquisser un état des lieux en nous plaçant successivement dans la perspective des catalogues puis du web, étant entendu que ce nouvel horizon ne saurait rester sans conséquence sur les catalogues eux-mêmes.

LES CATALOGUES

Les catalogues des établissements documentaires utilisent Rameau dans le cadre de la recherche par sujets proposée par les systèmes intégrés de gestion de bibliothèque (SIGB). Une part essentielle du service final rendu dépend donc des performances de ces derniers : idéalement, un bon système de gestion devrait être à même d'exploiter les trois dimensions complémentaires du langage, même si elles peuvent être utilisées indépendamment les unes des autres dans une certaine mesure. Dans ce contexte, les interfaces publiques d'interrogation revêtent une grande importance car elles déterminent, pour une part essentielle, les modes de recherche des utilisateurs, selon que telle ou telle option est mise en valeur ou même simplement proposée sur l'écran d'accueil (sans doute est-il prudent, en ce domaine, de ne pas préjuger trop vite des usages en restreignant *a priori* le choix offert aux utilisateurs).

Au niveau terminologique, la gestion des formes rejetées constitue le strict minimum obligatoire, en deçà duquel les requêtes par mots les plus élémentaires sont condamnées à l'échec, sauf à tomber par hasard sur la forme retenue du descripteur. L'existence de SIGB assurant la seule gestion des formes retenues s'apparente donc à une aberration, au même titre qu'un système de prêt qui n'enregistrerait pas le retour des documents. Un tel service minimal devrait aller de soi, sans dépendre de son éventuelle inscription dans un cahier des charges.

Bien entendu, les catalogues ont tout intérêt à exploiter également la navigation sémantique dans les autorités, par exemple sous la forme d'arbres heuristiques, dans la mesure où elle permet de guider intelligemment l'utilisateur dans sa recherche, de lui en élargir l'horizon, en l'amenant même – qui sait? – à trouver cela qu'il ignorait chercher (ce qui est le propre de

la découverte). À terme, la possibilité d'entrer dans l'ensemble du vocabulaire Rameau à partir d'un nombre réduit d'autorités très génériques ne pourra qu'améliorer sensiblement ce type de service, sur le modèle du référentiel médical MeSH (Medical Subject Headings) qui offre déjà cette navigation dans l'arborescence de ses descripteurs à partir d'un nombre limité d'entrées proposées sur la page d'accueil.

Enfin, le caractère précoordonné de Rameau intéresse les catalogues en ce qu'il permet d'afficher les résultats des requêtes sous la forme d'index alphabétiques pouvant être feuilletés : au lieu d'une avalanche confuse de réponses, l'utilisateur obtient par ce moyen une présentation ordonnée des différents aspects sous lesquels est envisagé le sujet recherché, les subdivisions intervenant comme autant d'éléments de tri entre différents points de vue. Il s'agit donc d'un outil précieux pour réduire le bruit inhérent à la recherche par mots : c'est notamment le seul moyen d'isoler les documents d'ensemble portant sur un thème abondamment traité (par exemple, la France), sans qu'il nous soit répondu que la requête amène un trop grand nombre de réponses. Remarquons au passage que le niveau terminologique est également concerné, dans la mesure où la

consultation d'un index permet de distinguer aisément les homographes, ainsi « Ampoules (archéologie) » et « Ampoules (pharmacie) ».

Les bibliothèques et autres établissements documentaires ont donc intérêt à choisir, pour leurs catalogues, les systèmes d'interrogation les plus capables d'exploiter les possibilités du langage documentaire sur les plans technique et ergonomique. De ce point de vue, la récupération des notices d'autorité Rameau fournies par la BnF constitue un atout décisif, surtout si le SIGB sait utiliser leurs numéros identifiants pérennes plutôt que des mécanismes plus aléatoires de reconnaissance de caractères. De toute façon, comme l'évolution de la modélisation dans le domaine du catalogage tend de plus en plus à placer les données d'autorité au centre du dispositif, les SIGB seront nécessairement conduits à en tenir compte à l'avenir.

WEB, WEB SÉMANTIQUE, WEB DE DONNÉES

En 2008, dans le cadre du programme européen TEL (The European Library), le vocabulaire Rameau a été traduit à titre expérimental selon le formalisme SKOS (Simple Knowledge Organisation System = Système simple d'organisation

des connaissances) : ce dernier, conforme au standard RDF (Resource Description Framework = Cadre de description des ressources), permet de représenter et d'exploiter tous les types de vocabulaires contrôlés sur le web³. En 2012, la BnF s'est chargée de mettre cette traduction à la disposition des utilisateurs d'une manière à la fois visible et pérenne, avec la garantie de mises à jour régulières, grâce au site *data.bnf.fr*⁴. Les modalités juridiques de ce nouveau service ont été précisées au 1^{er} janvier 2014, la BnF, ayant choisi, comme l'Abes, de placer ses métadonnées descriptives sous le régime de la « licence ouverte » de l'État qui en permet la libre réutilisation, avec obligation de mentionner la source.

Alors que l'horizon du web venait ainsi de s'ouvrir à Rameau, et que le référentiel pouvait désormais être moissonné par les moteurs de recherche, de nouveaux usages n'ont guère tardé à se manifester. En premier lieu, au sein même de l'application *data.bnf.fr* puisque cette dernière, outre la mise à disposition du référentiel en format SKOS, utilise les autorités Rameau pour ses besoins propres et la création de pages thématiques, selon une démarche innovante. Ensuite, les autorités Rameau, récupérées via *data.bnf.fr*, ont fait l'objet de réutilisations par des acteurs extérieurs dans les domaines les plus divers (enseignement, recherche, etc.). Un certain nombre de problématiques précédemment explorées, dont Rameau était partie prenante, se trouvent ainsi relancées à une nouvelle échelle dans le contexte du web :

- l'interopérabilité, c'est-à-dire la possibilité de mettre en relation des référentiels différents afin de les rendre capables de dialoguer tout en préservant – exigence capitale – leur intégrité et leur autonomie respectives ;
- la recherche thématique géolocalisée, moyennant l'alignement des autorités géographiques Rameau avec des données géolocalisables (c'est-à-dire disposant des coordonnées géographiques qui font défaut aux notices Rameau) ;
- le multilinguisme, qui vise à permettre des interrogations multilingues dans des réfé-

rentiels employant des langues différentes, selon la perspective novatrice ouverte par le programme MACS (Multilingual ACcess to Subjects = Accès multilingue par sujet) qui associait la BnF, la Bibliothèque nationale suisse et la Deutsche Nationalbibliothek de Francfort ;

- enfin, le repérage d'entités nommées dans des textes numérisés, dans le prolongement d'une très intéressante expérimentation menée dans le cadre de TEL par plusieurs sociétés spécialisées dans le web sémantique (principalement Mondeca) : il s'agissait de vérifier si la recherche plein texte dans Gallica pouvait être améliorée avec un référentiel terminologique comme Rameau. Le test, effectué sur *l'Histoire de France* de Michelet, ainsi que sur ses cours au Collège de France, s'est révélé très prometteur : le niveau de pertinence des concepts repérés dans le texte (et restitués sous forme de nuages de mots) fait en effet un bond spectaculaire lorsqu'on utilise les autorités Rameau, par rapport aux résultats obtenus par les seules procédures d'extraction automatique. Le fait que les documents numérisés intègrent, de par leur nature même, des métadonnées endogènes ne semble donc pas rendre nécessairement superflu le recours à des métadonnées exogènes pour optimiser leur consultation.

Rameau devient donc un acteur à part entière dans l'univers en gestation du web sémantique et du web de données. Ces systèmes, conçus pour structurer le web et permettre aux machines de répondre aux requêtes humaines en tenant compte de leur sens, ont besoin de s'appuyer sur des entités conceptuelles et non plus seulement sur la reconnaissance de chaînes de caractères. Dans cette perspective, les métadonnées structurées, comme le sont les autorités Rameau, deviennent des outils indispensables et les référentiels, « la clef de voûte du web de données⁵ ». De fait, dès à présent, les spécialistes français du web sémantique ont commencé à intégrer Rameau dans leur trousse à outils⁶, et ce, d'autant plus volontiers que, loin d'être un handicap, l'assise institutionnelle de Rameau, son inscription dans la longue durée, lui confèrent paradoxalement une véritable légitimité dans l'univers mouvant du web, comme autant de gages de neutralité, de continuité, de

³ Antoine Isaac et Thierry Bouchet, « RAMEAU et SKOS », Arabesques n° 54, 2009. Consultable en ligne : <http://www.abes.fr/Arabesques/Arabesques-n-54>

⁴ Agnès Simon et Romain Wenz, « Des outils automatiques pour le signalement en bibliothèque : expérimentations autour du projet *data.bnf.fr* », *Bulletin des bibliothèques de France*, 2012, n° 5. Consultable en ligne : <http://bbf.enssib.fr/consulter/bbf-2012-05-0039-008>

⁵ Emmanuelle Bermès, *Le Web sémantique en bibliothèque*, Éditions du Cercle de la Librairie, 2013.

⁶ Bernard Vatant, *RAMEAU : un référentiel pour le Web de données*, Journée nationale RAMEAU du 10 juin 2011. Consultable en ligne : http://rameau.bnf.fr/informations/pdf/journees2011/rameau2011_vatant.pdf

stabilité, de pérennité, en un mot : de confiance. Le défi pour Rameau, dans les années à venir, sera de s'en montrer de plus en plus digne.

Peut-on, pour autant, aller jusqu'à affirmer que Rameau offre aux praticiens du web l'équivalent de ce que l'ingénierie des connaissances nomme « ontologie », c'est-à-dire un ensemble structuré de spécifications capables de rendre compte avec précision de la conceptualisation d'un domaine donné ? En d'autres termes : Rameau propose-t-il un modèle du monde ou un modèle des connaissances sur le monde ? Nous ne le pensons pas, ne serait-ce qu'en raison de sa dimension encyclopédique : sauf à prendre la notion dans un sens très large, une ontologie ne saurait en effet s'appliquer qu'à un secteur délimité du savoir à partir d'une approche elle-même circonscrite (il n'est pas certain, par exemple, que la représentation du nerf optique dans une ontologie médicale sera identique, selon que l'on se place du point de vue du neurologue ou de celui de l'ophtalmologue). À défaut de prétendre refléter idéalement le réel ou le système des connaissances sur le réel, Rameau s'efforce en revanche, d'une manière pragmatique (nominaliste plutôt que réaliste), d'introduire autant d'ordre que possible dans le langage naturel (niveau terminologique) et dans le réseau des concepts (niveau sémantique). Si l'objectif est assurément plus modeste, il n'en demeure pas moins ambitieux, compte tenu de l'immensité potentielle du champ d'application concerné : de fait, les référentiels encyclopédiques de ce type ne sont guère légion sur le web, et c'est aussi ce qui y fait leur prix.

Notons, avant de conclure, que cette nouvelle visibilité de Rameau sur le web participe également de la présence de la langue française sur ce média et donc, à son niveau et dans son ordre, du combat pour la diversité linguistique et culturelle, au sein d'une géopolitique des langues qu'il serait naïf de méconnaître.

CONCLUSION

Un rayonnement qui s'accroît, un réseau d'utilisateurs et de contributeurs dynamique, un référentiel qui s'enrichit en permanence et évolue en profondeur, des usages qui se diversifient dans des directions insoupçonnées voici seulement quelques années : en dépit de ses limites,

mais fort de ses atouts, Rameau semble bien faire la preuve de son utilité et n'avoir pas dit son dernier mot.

Mais, on l'a souligné, c'est un *work in progress*, dans une histoire longue déjà d'une quarantaine d'années de travail collectif (mais le temps respecte-t-il ce qui se fait sans lui ?), scandée par plusieurs étapes majeures : au commencement, il a fallu procéder à l'« accumulation primitive » du vocabulaire, à partir des répertoires québécois et américain, avant d'être en mesure de s'en émanciper progressivement ; puis est venu le temps de construire l'environnement institutionnel (organisation du réseau) et technique (élaboration des outils et services) nécessaires au déploiement de Rameau ; enfin, fort de ces acquis, le Centre national Rameau a pu entreprendre récemment l'ambitieux chantier de révision systématique du vocabulaire actuellement en cours, lequel n'aura de sens que s'il est mené à son terme.

L'objectif en est clair : améliorer significativement la qualité des autorités Rameau, alors que leur exposition n'a jamais été aussi grande ni leurs réutilisations aussi nombreuses et diverses, pour un meilleur service de l'utilisateur final, quel qu'il soit. L'enjeu ne l'est pas moins car, qu'il s'agisse des SIGB ou du web, il est impossible de nier l'étroite complémentarité qui existe entre, d'une part, la qualité des procédures techniques exploitant les métadonnées et, d'autre part, la qualité de ces dernières elles-mêmes : le service rendu par les métadonnées est en effet d'autant plus satisfaisant que les dispositifs techniques sont adaptés, performants, inventifs ; mais, inversement, la technique la plus élaborée ne saurait exceller si la richesse ni la structuration des métadonnées ne sont au rendez-vous (tout au contraire, plus les dispositifs sont performants, plus impitoyablement ils font ressortir, lorsqu'elles existent, les insuffisances des données qui leur sont fournies).

D'où l'impératif de travailler en ce sens, étant entendu que l'expertise proprement humaine apparaît ici incontournable, la machine ne pouvant se substituer au cerveau humain, ni pour concevoir et développer les dispositifs techniques, ni pour créer et améliorer les métadonnées.

B:F