

Pratiquer la musicologie dans une bibliothèque numérique encyclopédique

Practicing musicology in an encyclopedic digital library

Jacques DUCLOY (1), Thierry DAUNOIS (2)

(1) Université de Lorraine

Jacques.Ducloy@univ-lorraine.fr, Thierry.Daunois@univ-lorraine.fr

Résumé. Cet article présente une bibliothèque numérique structurée par une infrastructure encyclopédique réalisée avec des wikis sémantiques. Elle permet de mener simultanément et de façon collaborative la plupart des actions numériques de la recherche en sciences humaines. Nous présentons une application pilote en musicologie avec la rédaction d'articles sur Roland de Lassus pour laquelle on s'appuie sur des explorations de corpus et des éditions diplomatiques.

Mots-clés. Musicologie, bibliothèque numérique, exploration de corpus, changement de paradigme, édition diplomatique

Abstract. This paper introduces a digital library structured by an encyclopedic infrastructure realized with semantic wikis. It allows to conduct simultaneously and in a collaborative way most of the digital actions of research in humanities. We present a pilot application in musicology with the writing of articles on Roland de Lassus for which one relies on corpus explorations and diplomatic editions.

Keywords. musicology. digital library, corpus discovering, paradigm shift, diplomatic edition

1 Introduction

Comment vont évoluer les pratiques des chercheurs avec la connaissance dans un paysage où le numérique est omniprésent ? Cette question est au centre des préoccupations des services de soutien de la recherche, et plus particulièrement des bibliothèques. Pour apporter des éléments de réponse, nous avons lancé, en Lorraine, le projet LorExplor pour « appréhender les besoins informationnels des chercheurs dans la diversité de leurs thématiques ».

Grâce au soutien d'ISTEX, nous avons multiplié des expériences d'exploration de corpus de publications scientifiques en texte intégral, à la fois avec des enseignants-chercheurs, des étudiants et des professionnels des bibliothèques. En pratique, nous avons utilisé un ensemble cohérent de sites encyclopédiques (le réseau de wikis sémantiques nommé Wicri) pour résoudre les interfaces homme-machine et mutualiser les actions de curation de données.

De façon indépendante, nous avons également utilisé cette souche encyclopédique pour gérer des collections de publications, pour réaliser des éditions diplomatiques numérisées de documents anciens ou mettre en place des systèmes

d'information construits en mode collaboratif. Autrement dit, nous avons déposé des ouvrages numériques sur un substrat encyclopédique, comme on peut déposer des livres sur les rayons d'une bibliothèque.

Le résultat est que nous disposons d'un embryon de bibliothèque numérique où des chercheurs et praticiens peuvent mener ensemble, dans le même espace, des activités qui se faisaient autrefois sur des supports différents. Nous retrouvons ici l'image d'une salle de travail dans une bibliothèque dont les murs et rayonnages sont remplacés par une base encyclopédique qui assure la cohésion de l'ensemble.

Cette approche peut-elle être généralisée ? Plus précisément, avec LorExplor, nous avons fait des simples tests de faisabilité dans de multiples domaines. Nous voulons maintenant aller en profondeur dans un domaine spécialisé. Nous avons choisi la musique, ou, plus précisément, la musicologie. Une première étude sur la Renaissance, s'avère déjà très riche en enseignements. Par exemple, les besoins d'exploration d'une très large diversité de corpus sont omniprésents.

Nous présentons ici les premiers résultats de cette expérience d'immersion dans le département musique d'une bibliothèque totalement numérique. Dans un premier temps, nous donnerons quelques éléments de réflexion sur les changements de paradigme dans les bibliothèques de la recherche, en y situant nos expériences antérieures. Nous introduirons ensuite les grandes lignes d'un travail en cours sur la musique de la Renaissance autour de Roland de Lassus.

1.1 Avant-propos

Par rapport aux objectifs de CIDE.21, la rédaction de cet article est aussi une illustration des nouvelles pratiques rédactionnelles dans la recherche. En effet, il est travaillé et rédigé dans l'environnement hypertexte que nous présentons. Par construction, le manuscrit est déjà en ligne, il sera modifié suivant les remarques des relecteurs.

Une version numérique, plus complète, permettant au lecteur d'analyser le travail en profondeur, est disponible en parallèle. En particulier, la version papier ne contient pas d'explicitation de sigles ou de notes explicatives de bas de page. Elles sont remplacées par des liens dans la version numérique. Ces liens donnent accès à des pages rédigées ou améliorées dans le processus de rédaction de cet article. De même la bibliographie n'est explicitée que dans la version numérique.

Pour les relecteurs, adresse provisoire de la version numérique :

https://loreexplor.istex.fr/Wicri/Musique/fr/index.php/Utilisateur:Jacques_Ducloy/CIDE.21_Djerba/Article_hypertexte

2 Bibliothèques et pratiques numériques dans le nouveau paradigme de la recherche

2.1 Les bibliothèques dans les changements de paradigme de la recherche

Pour alerter les chercheurs sur la révolution numérique, Jim Gray, en 2005 avait défini quatre changements de paradigme dans les pratiques de la recherche.

1. Pendant des millénaires, les premiers érudits utilisaient une méthodologie empirique basée sur l'observation.

2. Puis, depuis quelques siècles, les scientifiques utilisent des modèles théoriques et font appel aux abstractions et généralisations.
3. Depuis 1950, quelques décennies déjà, ils utilisent des ordinateurs pour modéliser des phénomènes complexes. La programmation devient un outil de travail et d'expression du chercheur.
4. Nous entrons maintenant dans une quatrième étape, celle des *big data* où les scientifiques sont confrontés au déluge des données.

Cette notion de quatrième paradigme de la recherche nous semble particulièrement riche. Nous cherchons donc à l'appliquer à la connaissance devenue numérique et traditionnellement portée par les bibliothèques.

En 2005, cette réflexion s'appliquait essentiellement aux sciences physiques ou de l'environnement (génomique notamment). Avec la mise en ligne massive de corpus numériques, les chercheurs en sciences humaines sont à leur tour concernés. Ils doivent assimiler les pratiques de création ou d'exploration de corpus de documents. Par rapport aux physiciens, les données qu'ils traitent ne sont plus exogènes. En effet, ce sont des documents qui ont été produits à des fins éditoriales, comme leurs productions scientifiques.

Sur un plan technique, l'émergence du quatrième paradigme coïncide avec celle des wikis qui rendent possible un travail réellement collaboratif. Nous entrons donc dans une époque où la plupart des pratiques de certaines communautés scientifiques peuvent se réaliser en mode coopératif dans un même espace numérique. Au niveau du document, le wiki généralise la notion d'hypertexte à rédaction collective, là où les étapes précédentes (hypercard...) étaient caractérisées par la nécessité d'un rédacteur unique.

Les collections de documents qui étaient juxtaposés dans les rayons deviennent des corpus que l'on peut maintenant explorer par des méthodes statistiques ou de filtrage. Plus encore, ils peuvent être mis en interrelation. La bibliothèque n'est plus simplement un lieu de dépôt ou de lecture, elle devient un espace de travail. Paradoxalement, grâce au numérique, elle retrouve la position stratégique qu'elle avait antérieurement à l'apparition de l'imprimerie¹ : un espace où la connaissance s'écrit, se lit et se travaille.

En pratique, en s'appuyant sur nos expériences antérieures, nous proposons un modèle de bibliothèque où les murs et les rayonnages sont remplacés par un substrat encyclopédique.

2.2 Des produits nouveaux à l'accompagnement des nouvelles pratiques

Le travail décrit ici s'appuie sur des expérimentations démarrées à l'INIST.

En 1988 l'INIST a pris la suite des centres de documentation (CDST et CDSH) créés en 1949 en même temps que le CNRS, avec la mission de communiquer sur l'essentiel² des résultats de la recherche. La création récente des «

¹ Voir par exemple la place stratégique de la bibliothèque dans le film tiré du roman *Le nom de la Rose*.

² Pierre Auger avait repris les ambitions des listes de grandeurs physiques : *Nous relevions l'essentiel de ce qui se faisait dans toutes les langues intéressantes à l'époque*. Cité par Jean Astruc dans le CNRS et l'information Scientifique et technique en France (Revue Solaris 1997)

essentiels de Gallica » à la BnF montre que les missions initiales de l'INIST sont encore d'une furieuse actualité. Conçu en 1985, le modèle de fonctionnement choisi pour l'INIST était écartelé entre la modernisation des acquis des centres de documentation du CNRS et la perception d'un changement de paradigme non encore verbalisé.

La renommée du CDST et du CDSH reposait sur les bulletins signalétiques du CNRS qui alimentaient les bases Pascal et Francis. Cette activité, notamment en sciences humaines, reposait sur un vaste réseau de coopérations avec les laboratoires. Ancré dans le paradigme numérique des années 80, la modernisation a pris la forme d'une « usine à transformer le savoir par des chaînes de production ». Mais un indéniable succès initial a aussi ouvert la voie aux difficultés actuelles de l'établissement. En effet, deux chemins s'ouvraient : celui de l'industrialisation, dans une perspective de volume et de chiffres - et, en particulier, de chiffre d'affaires -, ou celui de la connaissance. Le contexte de l'époque, et des choix dont il est aisé de dire aujourd'hui qu'ils étaient discutables, ont amené à privilégier le premier chemin.

Le fondateur de l'INIST, Goéry Delacôte, souvent qualifié de visionnaire, avait mis en place un département stratégique de Recherche et Développement, dont la durée de vue a été éphémère, mais qui a permis de dégager des options fondamentales qui sont encore au cœur de notre démarche. En effet, une orientation était de penser « station de travail du chercheur ». Cette directive, soutenue par l'expertise de Nathalie Dusoulier sur les standards d'échanges de documents, avait conduit à la création d'une première « boîte à outils SGML » pour construire cette station de travail, perçue comme très différenciée suivant les disciplines scientifiques visées. Ces résultats, acquis en 91, ont été largement validés par la suite au sein du monde du numérique. Au sein de l'INIST, ils ont été vécus comme une remise en cause du modèle informatique de production.

Cette approche s'est donc poursuivie dans un autre cadre, au Loria plus précisément, avec la boîte à outils Dilib, qui a intégré le standard XML et fait émerger la notion de serveur d'exploration. Elle est revenue à l'INIST en 2000, où elle a permis la création d'un service des prestations de veille et le lancement d'un plan de formation nommé « mutation technologique ». Mais une situation redevenue conflictuelle après un nouveau changement de direction a conduit à une mise en sommeil de cette activité. Le soutien d'ISTEX pour le projet LorExplor a permis de reprendre cette idée de boîte à outils, mais pour aller au-delà des métadonnées et traiter maintenant des corpus hétérogènes en texte intégral.

Cette « mise en sommeil » a dégagé du temps pour des actions éditoriales. D'abord à l'INIST, avec le projet Artist/Ametist, nous avons créé un ensemble éditorial multiforme (papier, numérique simple, numérique structuré, blog). À partir de 2008, au sein des universités de Lorraine et avec le soutien de la DRRT, nous avons lancé un réseau de wikis sémantiques, nommé Wicri, basé sur la technologie MediaWiki+Semantic MediaWiki, avec l'objectif initial de construire des observatoires de l'innovation. Nous avons par exemple constitué un inventaire très détaillé des projets européens dans lesquels des acteurs lorrains étaient impliqués.

Sur cette base, de façon un peu « opportuniste », nous avons multiplié les expérimentations à caractère éditorial. Nous avons d'abord procédé à la mise en

< [https://lorexplor.istex.fr/Wicri/Ticri/fr/index.php/Solaris_\(1997\)_Astruc](https://lorexplor.istex.fr/Wicri/Ticri/fr/index.php/Solaris_(1997)_Astruc) >

ligne de collections d'articles scientifiques dans un environnement hypertexte : les trois numéros de la revue *Ametist* et quelques articles de *Solaris*. Nous avons également ouvert des wikis pour les communautés CIDE, H2PTM et VSST. Suivant ces exemples, des chercheurs de l'Inra ont lancé une publication originale, « Les mots de l'Agronomie ». Nous avons également procédé à l'édition diplomatique de textes anciens. Et, tout récemment, dans la dynamique ISTEEX, nous avons complètement repensé l'interface avec les serveurs d'exploration.

En résumé, sur des maquettes, nous avons pu montrer que pratiquement toutes les activités entrant dans le champ de vision d'une bibliothèque ou de l'INIST pouvaient être menées dans un environnement collaboratif unifié.

Pendant toute la durée du troisième paradigme, un slogan était martelé : l'ambition affichée était de « décharger le chercheur des tâches documentaires ». Avec le quatrième paradigme, il faut faire exactement le contraire et accompagner le chercheur dans la maîtrise des outils et pratiques de la connaissance numérisée !

2.3 Des sciences de l'information, à la musicologie, en passant par l'environnement

Après avoir pensé « outils et prestations pour les chercheurs », il nous paraissait donc important de nous immerger dans une activité scientifique afin de mieux comprendre les interrelations des pratiques.

Les orientations données au réseau Wicri concernaient les sciences de l'environnement. Il serait effectivement stratégique de mener une expérimentation dans une des disciplines où nous avons déjà posé des jalons, comme l'eau, le bois, la forêt, les matériaux ou la santé. Nos premières expériences ont montré qu'une expertise solide dans les fondements théoriques étaient indispensables pour interpréter les résultats d'une expérimentation et émettre de nouvelles hypothèses.

Nous avons naturellement mené des travaux d'investigation scientifique dans les approches numériques des sciences de l'information. Mais ce domaine présente une spécificité paradoxale. En effet, il n'est pas nécessaire de lire un article pour comprendre la démarche algorithmique d'un collègue. Une discussion suffit ! Un colloque comme CIDE offre un cadre unique pour la veille scientifique. Le fait d'y soumettre un article est fondamental pour clarifier des concepts. La présentation publique ou la démonstration sont importantes. Mais la lecture n'est que rarement indispensable et le besoin d'exploration de corpus est donc un peu marginal.

Nous avons donc cherché un domaine impliquant des traitements de corpus, où nous n'avions pas forcément de fortes compétences mais une forte motivation pour acquérir une expertise et une érudition. Les premières investigations lancées sur la musique de la Renaissance ont mis en évidence le fait que, dans ce domaine de la musicologie, les besoins d'exploration de corpus dépassent nos espérances...

3 Écrire sur Roland de Lassus dans une bibliothèque numérique

Nous avons une expertise initiale en ingénierie numérique et une faible érudition musicale. Pour nous immerger dans une pratique scientifique, nous avons choisi de publier sur des sujets sur lesquels nous étions bétotiens. Nous avons choisi comme cible éditoriale les choristes qui cherchent à mieux comprendre le sens de ce qu'ils chantent. En effet, nous avons fait le pari que ce positionnement devrait nous amener à une intense activité de découverte et donc d'exploration de corpus. Nous

espérons ainsi améliorer progressivement le contenu et les outils de cet ensemble pour y attirer à terme des spécialistes, compositeurs, ou musicologues.

Voici maintenant quelques retours d'expérience sur un travail éditorial en cours : montrer comment la musique renaissance était composée et chantée.

3.1 La salle d'écriture pour la musique : une base encyclopédique ouverte sur le monde

En poursuivant l'analogie entre le réseau Wicri et une bibliothèque, le cadre de ce travail, le wiki Wicri/Musique, devient au maintenant la salle de lecture, ou plus précisément la salle d'écriture et de travail.

Nous avons souvent présenté le réseau Wicri (Wikis pour les communautés de la Culture, de la Recherche et de l'Innovation). Il utilise la même souche logicielle que celle de Wikipédia avec des extensions sémantiques (Semantic MediaWiki).

Pour permettre l'expression de travaux scientifiques, nous devons dépasser l'obligation de citation et imposer en retour une identification explicites des auteurs. Seules les personnes enregistrées sont autorisées à contribuer.

Pour favoriser un travail scientifique en profondeur, chaque thématique peut disposer de son propre wiki, avec ses règles terminologiques propres. Par exemple, le wiki Wicri/Santé utilise le MeSH comme base terminologique. Pour la musique, compte tenu de son absence dans Francis, nous allons probablement devoir mener un travail spécifique. Il est provisoirement initialisé avec une première compatibilité avec le Web Sémantique via Wikipédia (et WikiData).

Dans cette démarche nous avons fait un test intéressant avec le Trésor de la Langue Française. Nous avons déjà comme pratique de l'utiliser pour donner une définition lors de la création d'un nouvel article de la base encyclopédique. Nous allons plus maintenant plus loin en reprenant des articles complets et en les adaptant. Voici un exemple de retombées : pour le terme « Musique » nous avons trouvé une citation qui a joué un rôle de déclencheur :

[...] *musique et paroles d'une chanson. Même sans la musique le Victimae pascali laudes est un admirable poème en vers libres* (GOURMONT, *Esthét. lang. fr.*, 1899, p.250).

Nous sommes alors partis de cette citation pour rédiger un article sur la séquence *Victimae pascali laudes*, et une entrée encyclopédique sur Remy de Gourmont. Ces nouvelles pages sont reliées par des liens sémantiques (par exemple « A pour auteur cité » entre l'article musique et Remy de Gourmont.

Ce mode de fonctionnement est généralisé avec les rééditions diplomatiques.

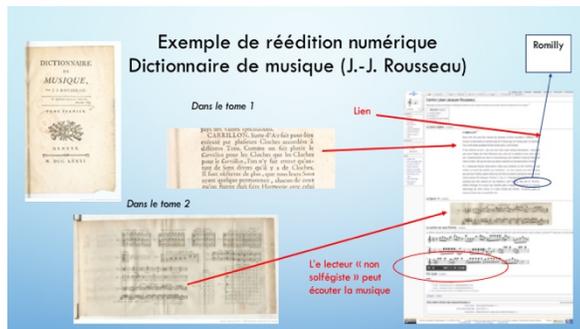
Rééditions diplomatiques en numérique

Le TLF définit l'édition diplomatique comme celle « *Qui reproduit scrupuleusement le texte primitif d'un ouvrage* ». Nous utilisons régulièrement ce procédé pour mettre en valeur un article fondateur qui devient alors un élément structurant de l'hypertexte et une source de définitions pour l'encyclopédie.

S'agissant de textes anciens, une transcription numérique permet de rendre le texte plus lisible pour des lecteurs contemporains. Ceci est particulièrement significatif pour les noms propres, et notamment les noms de lieux qui évoluent au cours de l'histoire. Le texte original peut de même être complété par une iconographie complémentaire. En musique ancienne, le fait religieux est omniprésent et souvent totalement inconnu des lecteurs nés après Vatican 2. Les liens hypertextes permettent de conserver la vision originale et de le rendre lisible.

Il est parfois possible d'offrir au lecteur une version qui aurait été celle de l'auteur s'il avait disposé du numérique. Par exemple, nous offrons une version de la *Chanson de Roland* écrite par Francisque Michel en 1869 et annotée ensuite par Paul Meyer ou le lecteur peut découvrir le résultat de ces annotations.

Avec les anciens musicologues, le lecteur peut maintenant écouter les musiques proposées dans leurs articles. Par exemple, Jean-Jacques Rousseau a écrit un dictionnaire en 2 volumes où les planches contenant des partitions dans le deuxième tome. Sur Gallica, le lecteur doit trouver le tome 2, y localiser les planches pour tomber sur une partition qui contient une clé qui n'est plus utilisée. Sur Wicri/Musique, tout est regroupé sur une seule page ; la musique est retranscrite avec une clé plus actuelle et être écoutée. En un clic, le lecteur peut consulter dans la souche encyclopédique une biographie d'un horloger cité dans l'article.



En multipliant de telles rééditions dans une perspective éditoriale, nous constituons un corpus, dans lequel chaque instance a été décidée par un être humain, un bibliothécaire, et non seulement par un algorithme de moissonnage.

Serveurs d'exploration

Dans notre approche les algorithmes sont essentiellement destinés à explorer des corpus de documents qui ne sont pas « encore rangés, ni même commandés dans la bibliothèque ». Grâce à des sources d'information comme ISTEEX, le bibliothécaire peut explorer les contenus de ce qu'il va mettre dans les rayons.

Nous ne développerons pas ici nos travaux sur les serveurs d'exploration. Voici cependant quelques résultats de l'expérience LorExplor.

L'utilisation des wikis sémantiques a permis un saut considérable dans la réalisation des interfaces homme-machine (paramétrage, présentation et explications des résultats). Un résultat plus important, (moins attendu au départ) est venu de la

couche encyclopédique qui devient un outil pour mutualiser les règles de curation. Elle permet également de diffuser les résultats d'une exploration de corpus.

Concernant la musique, nous avons mené une quinzaine d'expériences, en majorité sur des compositeurs (Mozart, Monteverdi, Debussy), mais également sur des types d'œuvre (comme le Magnificat ou l'opéra). Nous avons également traité des exemples pluri-disciplinaires autour de la « danse-thérapie » qui sont également exprimés dans l'espace santé (Wicri/Santé) de notre bibliothèque numérique.

Sur le plan des analyses statistiques, nous avons pu montrer que dans la grande majorité de nos études (tous domaines confondus) les corpus « bruts » étaient très fortement bruités ou biaisés. Voici quelques observations en musique :

- La requête « *Heinrich Schütz* » donne des résultats catastrophiques compte tenu de la fréquence d'usage du patronyme Schütz.
- Concernant Mozart, nous avons pu rencontrer des affiliations de laboratoires localisés dans des avenues Mozart, ou des « Mozart de l'informatique », ou encore des projets dont l'acronyme est Mozart.
- Toujours sur Mozart, les articles médicaux le concernant sont rédigés par des équipes bien rodées sur la déclaration des affiliations, là où les musicologues réputés se contentent de donner leur nom... Les résultats statistiques sur la musicologie sont alors insignifiants et masqués par des problématiques de santé.

Nous avons montré que des traitements de curation permettaient de donner une meilleure version globale, mais encore très imprécise pour des traitements statiques. En revanche, le serveur d'exploration devient alors très intéressant pour des détections d'acteurs, soit par navigation, soit par l'utilisation de techniques de filtrage. Il se révèle également comme un puissant moteur de sérendipité.

Des ressources numériques spécifiques pour la musique

Comme nous l'avons déjà mentionné, notre activité en musicologie est récente et nous ne présentons ici que des premières pistes de réflexions.

Concernant l'écriture de la musique, et plus précisément la notation numérique, nous utilisons le logiciel LilyPond, solution retenue par Wikipédia. Il bénéficie d'une bonne réputation et offre un avantage fonctionnel majeur pour notre expérimentation : la possibilité d'une écriture collaborative. Wicri/Musique devrait à terme disposer d'un dossier sur le sujet, avec un investissement sur les normes comme la notation "abc" qui permet de faire des recherches sur des thèmes musicaux, et bien entendu la norme MusicXml.

Pour pouvoir présenter des interprétations, nous utilisons pour l'instant une solution basée sur des fenêtres YouTube.

L'identification des sources musicales est un volet important sur lequel nous progressons. Nous avons cité ISTEEX qui présente deux limites. D'une part la couverture de la musicologie est essentiellement anglophone (*Cambridge University Press* et *Oxford University Press*). Les musicologues francophones sont sous-représentés. D'autre part, seules les métadonnées sont exploitables en accès public.

Le domaine de la musique ancienne permet une utilisation conséquente de sources libres de droit via Gallica, Persée ou Internet Archive. Pour les études actuelles, le problème est plus délicat. Seuls les articles publiés avec des licences permettant une réutilisation sont intégrables. On peut donc mettre en accès public une version hypertexte d'articles venant de HAL, mais pas d'OpenEdition.

Citons également, ChoralWiki, connu pour ses partitions, qui contient des inventaires, des biographies et des liens vers les manuscrits ou éditions originales.

3.2 Écrire dans l'hypertexte en explorant des corpus

Nous présentons maintenant quelques expériences d'écriture et notamment un dossier en cours sur Roland de Lassus.

Nous avons signalé la séquence médiévale *Victimae Paschali Laudes* repérée grâce au TLF. Deux articles lui ont été dédiés. Au-delà de la présentation de ce chant, le premier article donne un exemple de transcription grégorienne simple. Un travail d'investigation (sur Internet) a permis de repérer un manuscrit à 2 voix présentant une difficulté d'interprétation. Nous avons pu la résoudre grâce à une discussion qui est visible sur le site (à la façon d'un cahier de laboratoire).

Un exemple plus conséquent est donné par le dossier sur Roland de Lassus. Il a été initialisé à la un stage de chant choral organisé en Grèce (Kos) par la fédération belge du mouvement A cœur joie. Le « morceau de bravoure » était un *Regina Coeli* à quatre voix composé par Roland de Lassus. Il était accompagné par une version grégorienne de ce chant religieux, et par une chanson « *Mon cœur se recommande à vous* » plus simple à chanter. En nous appuyant sur l'expérience précédente, nous voulons rédiger quelques articles autour des partitions de la Renaissance : comment sont-elles transposables ? comment étaient-elles chantées ?

Nous avons commencé par « planter le décor » avec un serveur d'exploration sur Roland de Lassus, et un démarrage de réécriture diplomatique d'une biographie. Non musicologues, nous avons alors découvert à la fois la richesse de production de Roland de Lassus (entre 1500 et 2000 compositions) et la difficulté pour trouver des exemples simples et complets pour alimenter les articles envisagés.

Par exemple, en consultant ChoralWiki (et d'autres sources) nous avons découvert que la chanson sur laquelle nous voulions nous appuyer était probablement composition du XVIII^e siècle attribuée à tort à Lassus. Nous nous sommes donc repliés sur une chanson plus authentique « *Quand mon mary revient de hors* ». Nous avons alors découvert que les transcriptions actuelles étaient légèrement différentes des originales (un do dièse à la place d'un do naturel). En voulant comprendre ces différences nous avons découvert des contrafacta réalisés par des imprimeurs protestants qui donnaient la trace du do dièse mais sur des paroles totalement différentes (à destination des oreilles chrétiennes !).

Au niveau des pratiques, comme pour un travail de recherche, avons été confronté à une succession de phases de « début de rédaction - découverte d'un problème - recherche d'information - actions éditoriales dans le wiki - retour à la rédaction ». Ces diverses phases ont mobilisé en lecture et écriture toutes les ressources de la bibliothèque : serveur d'exploration, base encyclopédique,

biographies en réédition diplomatique, inventaire des compositions ou des ouvrages de l'auteur, etc.

Par exemple, l'approfondissement des rééditions diplomatiques de biographies a mis en évidence des multiples dénominations de Roland de Lassus : *Roland de Lattre, Orlando di Lasso etc.* Ceci a permis d'améliorer les critères de sélection et la volumétrie du serveur d'exploration qui est passée de 450 à 1300 documents.

Dans cet exemple, une action éditoriale avec un objectif assez bien cerné et sans fortes contraintes de temps nous amène à un enrichissement très significatif de la bibliothèque et révèle un travail numérique multiple dont la succession des phases est imprévisible (donc insoluble par des prestations externalisées).

Avec l'article pour CIDE, cet article, nous testons un modèle éditorial à trois versions : une version papier pour l'évaluation et l'insertion dans les actes, une version numérique permettant à celui qui veut approfondir d'entrer dans les détails, et, évidemment, les diapositives de présentation.

4 Des pistes pour une généralisation

Nous manquons de recul pour apprécier les retombées éditoriales de notre approche et la capacité de mobilisation vers d'autres communautés. Cela dit, l'ensemble « base encyclopédique, rééditions diplomatiques, articles scientifiques ou de vulgarisation, serveurs d'exploration » nous paraît, à l'expérience, très puissant.

Cette expérience est-elle généralisable ?

Le démonstrateur Wicri/Musique concentre des outils et pratiques, testées préalablement, et avec satisfaction, sur d'autres wikis. Par rapport à des « bibliothèques de lecture », l'intention éditoriale amène à pousser très loin les manipulations. Elle exige un pilotage par un spécialiste du domaine possédant une compétence en technologies numériques de la connaissance (ou par un binôme).

Sur Wicri/Musique, les interventions ont été en très grande majorité effectuées par une seule personne mais nous avons suffisamment d'expérience sur d'autres wikis pour penser que la dynamique collective peut fonctionner (Wikipédia le montre en permanence).

Nous rencontrons une limite liée aux droits d'auteurs qui freine considérablement la mise en rayon d'articles contemporains. Comme le montre l'expérience du NCBI avec PubMed Central, cette limite a été dépassée dans les domaines de la santé et des sciences de la vie. Avec les pressions politiques sur l'open Science on peut espérer une évolution comparable dans les humanités numériques.

Le défi pour Wicri/Musique est donc d'attirer suffisamment de spécialistes pour lancer une revue ouverte.

Concernant la faisabilité technique, le réseau Wicri repose sur une technologie qui demande de l'expertise mais relativement peu de développements. La mise en

place du réseau de 150 wikis, le développement de la boîte à outils XML, ont été réalisés par une seule personne à temps partiel. Sur cette base l'ensemble des outils d'application tels que modèles sémantiques ont été réalisés par des spécialistes de disciplines d'applications formés au numérique.

Concernant la formation au numérique, nous avons constaté un progrès considérable. Dans les années 2000, il était quasiment indispensable de démarrer un parcours de formation à l'exploration de corpus par une formation Unix/langage C pour manipuler des documents XML et paramétrer les serveurs d'exploration. On peut maintenant imaginer un parcours qui démarre par des actions purement rédactionnelles pour évoluer progressivement vers l'expertise numérique. De plus chaque étape de la formation demande un apprentissage dans lequel les manipulations vont enrichir le contenu ou la performance de la bibliothèque.

5 Conclusion

Nous venons de décrire, sur la musicologie, un modèle de bibliothèque où des chercheurs et des praticiens peuvent mener presque simultanément et sur le même support toutes les actions en liaison avec le numérique. Nous avons montré l'aspect écologique de cette approche qui ne demande pas de ressources humaines spécifiques. Avec un gain de temps, ils peuvent en plus contribuer à construire leur grande bibliothèque scientifique qui s'enrichit par le travail d'investigation du chercheur ou de l'ingénieur.

A la différence des bibliothèques gérées par des robots moissonneurs, toute la construction de cet univers de connaissance est d'abord réalisée par des êtres humains, qui savent éventuellement apprivoiser des robots.

Pour les difficultés de mutation, avons évoqué les changements de paradigmes de la recherche, du document, des pratiques associées et des bibliothèques. Trop directement impliqués nous n'avons pas cherché à analyser les obstacles au changement à tous les niveaux. Voici une piste qui nous paraît intéressante à creuser. Les bibliothèques des paradigmes précédents étaient régies par des procédures bornées dans le temps, et dans l'espace, calibrées par les kilomètres de rayonnement. Dans une bibliothèque hypertexte l'espace devient quasiment infini. De même, dans la communication scientifique. Les chercheurs ont appris à faire des articles de 9 pages. Il faut maintenant rédiger dans un espace infini.

Comment gérer sa position dans l'infini sans crainte ?