Ré-édition de Chrestien de Lihus dans l'hypertexte

Thierry DAUNOIS
NIT - Institut polytechnique de Lorraine, France

Résumé : Cet article traite de la ré-édition d'un ouvrage ancien traitant d'agriculture dans le réseau de wikis Wicri. Après une étude de différents projets de ré-édition numérique menés en France, il ouvre une réflexion plus large sur les possibilités offertes par la technologie wiki en matière d'édition numérique. On peut en effet imaginer viser la simple mise à disposition de textes non accessibles pour permettre leur réutilisation. Mais on peut également concevoir le développement d'outils spécifiques, l'exploitation de fonctionnalités sémantiques, dans une optique de recherche. Une alternative intermédiaire consiste à accompagner des projets de recherche, sur le volet de mise à disposition et de visibilité.

Mots-clés : Edition numérique, hypertexte, wiki, agriculture, annotation.

« Le livre, comme livre, appartient à l'auteur, mais comme pensée, il appartient, le mot n'est pas trop vaste, au genre humain. »

Victor Hugo
Discours d'introduction du Congrès littéraire international de 1878

1. Introduction

Certains documents (livres et manuscrits plus ou moins anciens, numérisés mais accessibles uniquement sous la forme de pdf, parfois avec un OCR[56] de qualité variable...) ne sont pas disponibles, en version exploitable (texte brut utilisable), sur internet. Dans le même temps, des chercheurs disposent, sur leur poste de travail, d'extraits, de chapitre, et même de livres entiers qu'ils ont entièrement retranscrits dans leur propre traitement de texte. Ce constat est à l'origine de l'idée initiale, qui

[56] OCR (optical character recognition) : acronyme qui désigne la reconnaissance optique de caractère. Avec l'utilisation de plus en plus fréquente de logiciels effectuant de la reconnaissance de caractères pour exploiter des pdf, "un OCR" désigne maintenant le fichier obtenu avec ces logiciels.

consistait à tester la possibilité de mettre ces textes à disposition de tous, pour que tout le travail de re-saisie ne soit pas perdu.
Même s'il n'est pas possible de quantifier cette "ressource", il n'est pas difficile d'imaginer qu'elle est importante. La production totale de l'humanité est estimée à quelques 130 millions d'ouvrages (évaluation effectuée par Google, dans le cadre de son vaste projet de numérisation)[57]. Les livres effectivement disponibles sur internet (sans même s'intéresser à leur "exploitabilité"), bien que l'on ait assisté à une véritable explosion en la matière depuis quelques années, se comptent plutôt en centaines de milliers. Le Projet Gutenberg[58] annonce 100.000 ouvrages traités, Gallica[59] 300.000. Même le méga-projet de Google Books portait sur 15 millions de livres. La marge reste donc colossale !
L'idée initiale a donc été de tester, sur le réseau de wikis Wicri, la mise en ligne de ressources éditoriales ayant déjà fait l'objet du travail ingrat de re-saisie en format de type word. Nous étions alors dans une simple optique de ré-édition directe.
De cet exercice - commencé comme une démarche d'information scientifique et technique (IST) et d'édition de "service public" -, est née rapidement l'idée qu'il était possible d'aller plus loin. Au-delà de ce cadre initial, pourquoi ne pas proposer un enrichissement des textes, avec des annotations collectives, mais, également, dans une optique de recherche ? Comment intégrer des outils permettant l'exploitation et la capitalisation de textes - enrichissement hypertexte [CLE 2007], analyse des données textuelles, paléographie[60], codicologie[61], philologie[62]... - ?
Cet article s'attache donc à présenter, dans un premier temps, le cadre technique et les choix initiaux qui constituent le contexte de cette expérimentation. Puis nous proposons une analyse des projets d'édition

---

[57] Article disponible sur le site américain mashable.com. (http://mashable.com/2010/08/06/number-of-books-in-the-world/?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+Mashable+%28Mashable%29&utm_content=Google+Reader Mashable).

[58] Données chiffrées sur le site du Project Gutenberg (http://www.gutenberg.org/wiki/Main_Page).

[59] Données chiffrées sur Gallica (http://blog.bnf.fr/gallica/?p=2991).

[60] Paléographie : "science qui traite des écritures anciennes, de leurs origines et de leurs modifications au cours des temps et plus particulièrement de leur déchiffrement - définition du Trésor de la langue française informatisé (http://atilf.atilf.fr/dendien/scripts/tlfiv5/advanced.exe?8;s=2840995800;).

[61] Codicologie : "science annexe, mais distincte, de la paléographie et ayant pour objet l'étude matérielle des manuscrits en tant qu'objets archéologiques (par l'étude des matériaux servant à la confection du livre manuscrit et leur mise en œuvre) - définition du Trésor de la langue française informatisé (http://atilf.atilf.fr/dendien/scripts/tlfiv5/advanced.exe?8;s=1988699865;).

[62] Philologie : "discipline qui vise à rechercher, à conserver et à interpréter les documents, généralement écrits et le plus souvent littéraires, rédigés dans une langue donnée, et dont la tâche essentielle est d'établir une édition critique du texte - définition issue du *Trésor de la langue française informatisé (http://atilf.atilf.fr/dendien/scripts/tlfiv5/advanced.exe?8;s=1988699865;)*.

hypertexte existants, avant d'effectuer un retour d'expérience sur notre expérimentation, depuis sa phase initiale jusque dans ses développements les plus récents. Enfin nous tentons de tracer les perspectives qu'ouvre le travail effectué dans le cadre du réseau Wicri - autant dans une optique d'IST que de culture scientifique et technique, et à destination de différents publics : grand public, enseignement, recherche -, et en quoi il pourrait constituer un apport pour les projets en cours.

2. Cadre technique - Choix initiaux

L'expérimentation se déroule dans le cadre de Wicri – WIkis pour les Communautés de la Recherche et de l'Innovation -, réseau de wikis sémantiques développé au sein de l'Institut national polytechnique de Lorraine (INPL). Initié en septembre 2008, le réseau Wicri compte aujourd'hui 101 wikis, tous développés à partir de la souche logicielle libre MediaWiki créée pour "l'encyclopédie libre" Wikipédia. Cette souche de wiki, si elle bénéficie (autant qu'elle en souffre) de l'aura du wiki le plus connu au monde, permet d'effectuer des choix différents de ceux qui animent l'encyclopédie en ligne. Ainsi, afin de prendre en compte les besoins spécifiques des communautés de la recherche, le réseau Wicri ne compte-t-il aucun wiki "libre" (sur lesquels la lecture et la contribution sont possibles sans être enregistré), mais uniquement des wikis publics (lecture libre, contribution uniquement pour les acteurs enregistrés et identifiés) et privés (lecture et contribution accessibles uniquement aux utilisateurs enregistrés et identifiés). Toute intervention est ainsi précisément rattachée à son auteur : chaque donnée est "traçable".

L'une des caractéristiques innovantes du projet Wicri est de constituer un réseau. Il est classique de trouver plusieurs wikis hébergés sur un même serveur, mais nous n'avons pas identifié d'initiatives proposant un véritable fonctionnement en réseau. Cela suscite des besoins particuliers, dont, pour assurer la cohérence des données d'un wiki à un autre, une réflexion approfondie sur la gestion des métadonnées. L'expérience menée par l'équipe Wicri en la matière a fait l'objet d'une publication au colloque DCMI 2010 à Pittsburgh [DUC 2010]. Autre différence de taille : Wikipédia exige de chaque contributeur qu'il appuie ses propos de références extérieures. À l'inverse, le réseau Wicri prévoit de s'appuyer sur l'expertise de comités scientifiques, fonctionnement adapté aux communautés de la recherche.

Au sein du réseau, on peut distinguer deux grands types de wikis. Les wikis "communs" (régionaux, Wicri/Lorraine et Wicri/Alsace ou thématiques, Wicri/Eau, Wicri/Bois...), d'une part, ont vocation à être animés par la communauté à laquelle ils se rattachent, tout en s'inscrivant dans les règles communes au réseau Wicri. Les wikis "institutionnels",

d'autre part, sont rattachés à une institution identifiée, qui en assure la direction éditoriale. Les choix éditoriaux, dans ce cas, peuvent être dérogatoires par rapports aux wikis communs du réseau : ouverture plus large, ou, au contraire, plus restrictive, du wiki aux contributeurs, par exemple.

Cette structuration offre deux intérêts. Elle favorise la construction collaborative de connaissances (en public ou en privé, au travers d'une application spécifique sur un wiki institutionnel, ou lors de la rédaction d'articles collectifs...). Puis elle assure la dissémination des informations ainsi générées, en leur assurant une bonne visibilité.

L'association de wikis "communs" et "institutionnels" permet de se positionner sur les différents niveaux de la connaissance. Les données brutes ont leur place sur des wikis institutionnel et de travail. Les wikis communs du réseau sont principalement destinés à la valorisation des résultats de la recherche et à s'intégrer dans les démarches de culture scientifique et technologique. Enfin, certains wikis institutionnels peuvent proposer de la vulgarisation scientifique grand public.

La question posée par Pierre Morlon, ingénieur au département Sciences pour l'action et le développement (SAD) de l'Institut national de la recherche agronomique (Inra), sur l'éventualité de mettre en ligne dans le réseau Wicri des ressources textuelles non accessibles par ailleurs - sous une forme facilement utilisable - sur internet a soulevé de premières interrogations. Ainsi, il fallait avant tout choisir où placer cette expérimentation dans le réseau, et quelle structure lui donner. L'ouvrage choisi pour cette expérimentation, les "Principes d'agriculture et d'économie", de Chrestien de Lihus (voir la documentation éditoriale et technique de l'expérimentation[63]), publié en 1804, a toute sa place sur le wiki thématique consacré à l'agronomie, Wicri/Agronomie[64].

Deuxième choix à effectuer : comment traduire, en pages wiki, un ouvrage de 336 pages ? Nous avions déjà eu l'occasion de travailler sur des articles, mais jamais sur des livres, ce qui demandait, de fait, une répartition de diverses sections sur différentes pages, avec un outil de navigation. Le travail préparatoire a donc consisté à étudier la table des matières de l'ouvrage. Cela a été l'occasion de la première constatation : la table des matières figurant dans l'ouvrage ne correspondait pas au découpage réel du texte, certaines sous-sections semblant être au même niveau dans le texte pouvant apparaître ou non dans la table. Seule la lecture du texte nous a permis de parvenir à une table des matières réelle qui semble satisfaisante, faisant apparaître quatre niveaux de titre. La

---

[63] La documentation éditoriale et technique est accessible sur Wicri/Agronomie (*http://ticri.inpl-nancy.fr/wicri-agronomie.fr/index.php/C_de_Lihus_1804_Principes_d%27agriculture_et_d%27%C3%A9conomie_-_Doc_%C3%A9ditoriale_et_technique*).

[64] Wicri/Agronomie *(http://ticri.inpl-nancy.fr/wicri-agronomie.fr/index.php/Accueil)*.

table des matières "corrigée" comporte ainsi, outre une préface et une conclusion, trois parties, constituées, pour la première, de deux chapitres, pour la deuxième, d'un chapitre unique, et, pour la troisième, de onze chapitres. Autrement dit, chaque chapitre fait en moyenne une vingtaine de pages (dans l'édition originale), les extrêmes étant de 2 pages (pour la conclusion), et de 43 pages (pour le chapitre "Août").
L'unité de travail retenue a donc été le chapitre : l'ouvrage de Chrestien de Lihus, dans le réseau Wicri, est donc publié sur 16 pages distinctes, préface, quatorze chapitres, et conclusion. Nous avons également créé plusieurs "modèles" (équivalent, dans MediaWiki, de macros, permettant de générer, sur plusieurs pages, un même texte), dont l'un destiné à faciliter la navigation d'un chapitre à un autre.
Il était intéressant également de conserver l'indication de la pagination initiale : ainsi, si l'on recherche un extrait dont on sait qu'il figure en page 228 de l'édition originale, on peut le retrouver rapidement. A cet effet, la pagination originale est indiquée (entre crochets et en caractères de couleur). En poussant cette démarche, nous avons ajouté une page consacrée à une "table des matières inverse", dans laquelle on peut retrouver directement, en fonction de la page que l'on recherche dans l'édition originale, à quel chapitre elle appartient.

3. Les projets d'édition hypertexte

Depuis l'un des premiers projets d'édition numérique français dont on retrouve la trace sur internet - l'expérience menée à l'Institut de recherche et d'histoire des textes (IRHT) en 2002-2003, et qui avait mobilisé un groupe de travail autour du manuscrit de "La lettre volée" [BUQ 2004] - bien du chemin a été parcouru. Un nombre croissant de projets d'édition numérique se sont organisés, qu'ils visent des textes isolés ou des corpus plus vastes.
Une analyse rapide des documents accessibles concernant ces projets montre qu'il ne semble pas y avoir eu de travaux menés sur l'idée de "ré-édition de service public", telle que nous la décrivons au démarrage de ce projet. Ainsi, on ne trouve pas trace de tentatives d'évaluation de la "ressource" disponible, ni d'expérimentation de mise en ligne de textes dans l'optique qu'ils deviennent simplement exploitables par d'autres.
Mais il semble également qu'il n'y ait eu que très peu de projets visant à mettre à disposition du plus grand nombre des textes d'intérêt scientifiques. En effet, on observe plutôt, aussi bien dans la littérature consacrée à l'édition numérique [LER 2008] que dans les projets qui semblent se rapprocher de l'expérimentation décrite ici, deux grandes orientations qui diffèrent sensiblement de notre démarche.
La première orientation consiste à travailler un corpus dans le sens des travaux de recherche d'un groupe de chercheurs identifiés, et, souvent,

de disciplines proches. Ainsi, la plate-forme Dinah revendique le fait de proposer un cadre de travail pour les philologues (voir encadré). C'est également le cas avec un projet comme Sourcencyme [DRA 2009], qui vise à créer une base de travail aux spécialistes de médiévistique sur les encyclopédies de l'époque.

La plateforme Dinah

La plateforme philologique Dinah [POR 2010] est destinée à annoter, transcrire et classer des documents manuscrits. Elle vise à "permettre l’expression conjointe de points de vues différents sous la forme de reclassements et d’annotations, [et en] la mise en œuvre des procédures nécessaires à la construction collaborative de vocabulaires d’annotations". Initiée dans le cadre du Cluster 13 (allocation de recherche 2007), la plateforme est accessible depuis avril 2010[65].

Cet outil, clairement destiné à une phase de travail, peut être utilisé quel que soit le contexte d'édition envisagé. Il peut donc être employé pour préparer une édition dans le cadre du réseau Wicri.

Fonctionnalités disponibles dans le cadre de la plateforme Dinah : annotation, travail collaboratif préparatoire.

Dans de tels contextes, la priorité est donnée au travail de recherche, à l'exploration des corpus dans l'optique des projets du laboratoire concerné, sans que, dans la plupart des cas, ces corpus soient du tout accessibles ne serait-ce qu'à d'autres équipes de recherche. Cette réalité est d'ailleurs à la base de la demande croissante de l'Agence nationale de la recherche (ANR) d'un véritable volet de "mise à disposition" des corpus ainsi traités.

La seconde orientation est celle que l'on pourrait désigner comme "technology-driven". En effet, qu'il s'agisse ou non d'un choix conscient, il apparaît clairement que de nombreux projets - et par exemple ceux menés dans le cadre du "cluster 13"[66] - s'appuient fortement sur une expertise en terme de traitement des images et de numérisation. Les techniques de numérisation bénéficient ainsi d'une abondante littérature [KAL 2000].

Le Cluster Culture, Patrimoine et Création (ou Cluster 13) [RAI 2008] porté par l'Université Lumière Lyon 2 vise à "coordonner les recherches pluridisciplinaires portant sur les productions, les objets et les usages sociaux qui engagent [...] une dimension et des enjeux d’ordre culturel et patrimonial, qu’il s’agisse du passé ou du contemporain le plus actuel.

[65] Présentation de la plateforme Dinah sur le site de l’Institut Jean-Toussaint Desanti (*http://institutdesanti.ens-lyon.fr/spip.php?rubrique27*).

[66] Le Cluster 13 (*http://cluster13.ens-lyon.fr/*).

L'ensemble du dispositif concerne principalement les sciences humaines et sociales, tout en étant ouvert à des collaborations avec les sciences exactes, les sciences de la nature et, en particulier, les sciences et techniques de l'information et de la communication (STIC)."
Ainsi, le projet "Hyperdonat" [BUR 2009] ou celui consacré aux dossiers de *Bouvard et Pécuchet*, de Flaubert [DOR 2009] – sont fortement marqués par l'impact de la nécessaire numérisation des documents. L'enjeu devient alors l'exploitation la plus aboutie possible de la technologie de traitement des images, et non la mise à disposition du plus grand nombre des textes ainsi traités. Ce sont au total 15 projets d'éditions critiques qui sont menés dans le cadre du Cluster 13, dont les *Essais* de Montaigne, les *Pensées* de Pascal, les œuvres complètes de Spinoza et de Montesquieu, les *Éloges académiques* de D'Alembert, l'*Essai sur les mœurs et l'esprit des nations*, de Voltaire.
L'expérimentation menée à l'École des Chartes, dans le cadre de Theleme[67] (acronyme de "Techniques pour l'historien en ligne : études, manuels, exercices") [POU 2006], mérite également d'être citée, même si elle se distingue également sensiblement du travail mené sur le réseau Wicri. En effet, ce travail, qui a nécessité le développement d'un outil spécifique de diffusion, ne semble pas permettre de travail collaboratif, et porte (du moins en l'état actuel de ce qui est consultable librement) uniquement sur des extraits brefs de documents. 116 dossiers sont accessibles, portant chacun sur une page d'un texte plus vaste, donnant accès à différents niveaux d'annotation (paléographiques, linguistiques, diplomatiques[68] ou historiques). Theleme est conçu essentiellement comme un support d'enseignement et d'initiation aux sciences et méthodes de l'histoire.
Or il apparaît que ces orientations ne sont pas antinomiques, mais pourraient au contraire se retrouver, dans une démarche commune visant à faire bénéficier l'ensemble des communautés de la recherche d'outils complémentaires, à la fois en terme de traitement des images, d'outils spécifiques, et de mise à disposition.
Enfin, le projet le plus proche de notre expérimentation est consacré à l'*Essai sur le récit, ou Entretiens sur la manière de raconter*, édition électronique de l'ouvrage de François-Joseph Bérardier de Bataut (1776).

---

[67] Système mis en place par l'École nationale des chartes (*http://theleme.enc.sorbonne.fr/dossiers/*). Chaque document est présenté sur une page, associée à d'autres pages, sur lesquelles sont donnés des commentaires (paléographiques, linguistiques, diplomatiques ou historiques).
[68] La diplomatique est une discipline qui vise à mettre en œuvre une compréhension critique des actes écrits. La Commission internationale de diplomatique précise qu'elle englobe tout écrit utilisé ou utilisable comme titre, fondamentalement pour prouver un droit. Cette définition est extraite de la présentation figurant sur le site Theleme (*http://theleme.enc.sorbonne.fr/cours/diplomatique*).

L'*Essai sur le récit*, édition électronique de l'ouvrage de François-Joseph Bérardier de Bataut
L'édition électronique de l'*Essai sur le récit, ou Entretiens sur la manière de raconter* [SCH 2010] est un projet mené par Christof Schöch, de l'Institut de Romanistique de l'Université de Kassel (Allemagne). Idée née à l'occasion d'un travail de thèse, l'édition électronique[69] dont il est question ici reprend l'unique édition connue de l'ouvrage de Bérardier de Bataut, publiée en 1776 à Paris.
La représentation du texte proposée donne la priorité au récit au détriment de la matérialité du livre, avec un découpage en chapitres et non par pages. Deux vues alternatives du texte sont proposées : une transcription linéaire du texte de l'édition originale, et un texte de lecture modernisé. Des notes textuelles et explicatives sont ajoutées : elles sont présentées sur un seul niveau, mais des évolutions ont été annoncées, notamment sur le système de notes, par C. Schöch, pour les mois à venir.

L'équipe du projet, composée d'une dizaine de personnes (outre le responsable du projet, l'équipe est composée du responsable du département d'informatique pratique, de l'un de ses collaborateurs et d'un groupe d'étudiants), a travaillé de 2008 à 2010, et prépare actuellement une nouvelle version.
Fonctionnalités disponibles dans le cadre de cette édition électronique : deux versions alternatives, annotations.
Mise en ligne effectuée sur drupal (après une version initiale sur DokuWiki).

4. L'expérimentation sur les *Principes d'agriculture et d'économie* de Chrestien de Lihus

La préface débute avec un épisode historique, rapporté par Cicéron (note originale : "Cic. de Oratore."), mettant en scène Annibal, dont il est dit qu'il fut très mécontent à l'écoute d'un philosophe, Phormion, qui discourait des devoirs d'un bon général, sans avoir jamais été militaire de sa vie. Cette anecdote sert à Chrestien de Lihus pour indiquer qu'il ne prend la plume qu'en temps qu'agriculteur lui-même, et pour apporter son expérience.
Avec l'objectif d'établir un lien vers une ressource en ligne, une rapide recherche a permis d'effectuer plusieurs observations. La première constatation est qu'aucune traduction de Cicéron disponible en ligne ne comporte le texte exact cité par Chrestien de Lihus. Soit il s'agit d'une traduction qu'il a effectué lui-même (ce dont, sans disposer d'éléments

[69] Le site dédié à l'édition électronique de l'Essai sur le récit (*http://berardier.org/*).

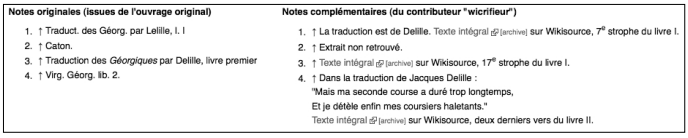
probants, on peut néanmoins douter), soit qu'il a repris quelque part, et qui diffère des versions que nous pouvons aujourd'hui trouver sur internet.

Dès lors, il a paru intéressant de donner des éléments plus complets sur cette citation, en apportant une note complémentaire à la note initiale : *De Oratore*, livre II, XVIII, Cicéron. Traduction consultée reprise des *Œuvres complètes de Cicéron*, publiées sous la direction de M. Nisard (1869). Texte intégral (lien) sur remacle.org". Cette première annotation du texte initial donne non seulement le lien vers une traduction en ligne mais vient également compléter la note originale, facilitant la recherche au lecteur.

Cette première note ouvrait la voie, renforcée dès la seconde : en effet, Chrestien de Lihus évoquait ensuite "L'auteur du *Préservatif contre l'Agromanie*", un ouvrage publié à Paris en 1762. Sans le citer nommément. Et pour cause, puisque ce livre était alors considéré comme anonyme, avant d'être attribué à Laurent-Benoît Desplaces. Figure ainsi la note complémentaire suivante : "Considéré un temps comme anonyme, le *Préservatif contre l'Agromanie* est attribué à Laurent-Benoît Desplaces. *Préservatif contre l'Agromanie, ou l'Agriculture réduite à ses vrais principes*, Paris : chez Jean-Thomas Hérissant, 1762, in-12, 197 p."

Il apparaissait dès lors qu'il y a un véritable intérêt à compléter, enrichir, et parfois apporter des éléments de correction aux notes originales (on parle de correction, par exemple, lorsqu'il est possible de constater qu'une citation, indiquée comme devant se trouver dans le tome II du *Voyages en France en 1787, 1788 et 1789* d'Arthur Young, se trouve en réalité dans le Tome I, page 452 (première traduction complète et critique par Henri Sée, édition Armand Colin, 1931).

Face à la constatation que la plupart des notes originales pouvaient ouvrir sur un ajout, il devenait utile d'opter pour une mise en page reprenant un double système de notes en bas de page, mettant en vis-à-vis la note originale et son commentaire (voir la figure I).



Notes originales (issues de l'ouvrage original)

1. ↑ Traduct. des Géorg. par Lelille, l. I
2. ↑ Caton.
3. ↑ Traduction des *Géorgiques* par Delille, livre premier
4. ↑ Virg. Géorg. lib. 2.

Notes complémentaires (du contributeur "wicrifieur")

1. ↑ La traduction est de Delille. Texte intégral [archive] sur Wikisource, 7e strophe du livre I.
2. ↑ Extrait non retrouvé.
3. ↑ Texte intégral [archive] sur Wikisource, 17e strophe du livre I.
4. ↑ Dans la traduction de Jacques Delille :
   "Mais ma seconde course a duré trop longtemps,
   Et je dételé enfin mes coursiers haletants."
   Texte intégral [archive] sur Wikisource, deux derniers vers du livre II.

Figure I : visualisation en vis-à-vis des notes originales et de leur commentaire.

On trouve ainsi, sur l'ensemble de l'ouvrage, matière à divers enrichissements. L'ajout de lien vers des ressources en ligne est le plus élémentaire. Parfois, il s'est avéré intéressant de comparer des sources diverses (traductions différentes, par exemple). Il a aussi été possible, parfois, d'identifier des erreurs dans des citations (sans pouvoir l'affirmer avec certitude, certaines de ces erreurs sont probablement directement

reprises des sources employées). Des sources imprécises ont également pu être éclairées : ainsi, une citation en latin, ""Delectant domi, non impediunt foris, pernoctant nobiscum, peregrinantur, rusticantur"", bénéficiait uniquement de la note suivante : "Cic. pro Archia, n°16." (figure II).

4. ↑ Plaidoirie de Cicéron lors du procès de Archia, p. 16. Texte intégral [archive] sur le site thelatinlibrary. Traduction juxtalinéaire [archive] sur le site Latin, Grec, Juxta :
[les études] *elles nous récréent dans nos foyers, ne nous embarrassent point au dehors ; elles veillent avec nous ; elles nous suivent en voyage, à la campagne.*

Figure II : les notes originelles peuvent être complétées, enrichies, rectifiées.

Au fur et à mesure de ce travail, il s'est également avéré utile d'apporter des notes sur le texte original, sur des éléments que Chrestien de Lihus n'avait pas annoté. Il reprend, par exemple, des citations latines sans les traduire. Il parle de "Rozier", sans préciser qu'il parle (probablement) de l'abbé Rozier, auteur d'un "Cours complet d'agriculture". De la même façon, lorsqu'il évoque le "chantre de Mantoue", il n'est pas forcément évident d'établir le lien (y compris en menant une recherche rapide sur internet) qu'il parle de Virgile. D'où une note : "Cette expression désigne Virgile. Voir à ce sujet les Études sur Virgile, tome III, page 132, de Pierre-François Tissot (1828, Paris). Texte intégral(lien) sur Gallica".
Enfin, pour des spécialistes de l'histoire des idées et de l'histoire de l'édition, on imagine facilement l'intérêt de ce type de démarche. Ainsi, dans le chapitre 1, partie I, la note originale [1] peut prêter à confusion, mais pourrait être intéressante dans cette optique. En effet, la note fait référence à un ouvrage employé comme source par Chrestien de Lihus, Histoire de l'Agriculture ancienne. Une première recherche fait apparaître qu'un ouvrage ainsi nommé est en effet paru, mais en 1830. Mais il n'est pas forcément totalement neutre d'observer également que ce même auteur a également publié, en 1804, justement, un autre livre, consacré aux Géorgiques, de Virgile, ce dernier étant abondamment cité par Chrestien de Lihus. Finalement, il apparaît (figure III), en se penchant plus en détail sur la question, que la note de Chrestien de Lihus fait plus probablement référence à Histoire de l'agriculture ancienne, extraite de l'Histoire naturelle de Pline, avec des éclaircissements et des remarques, livre XVIII, de Bernard-Laurent Desplaces (1765).

Notes originales (issues de l'ouvrage original)	Notes complémentaires (du contributeur "wicrifieur")
1. ↑ *Histoire de l'Agriculture ancienne*, dont j'ai tiré quelques morceaux dans ce chapître	1. ↑ Cette note fait probablement référence à *Histoire de l'agriculture ancienne, extraite de l'Histoire naturelle de Pline, avec des éclaircissements et des remarques, livre XVIII*, de Bernard-Laurent Desplaces, 1765 (Histoire de l'agriculture ancienne, extraite de l'Histoire naturelle de Pline, livre XVIII... [Texte imprimé]. - Paris : G. Desprez, 1765. - In-12, XLVIII-358 p.). Texte intégral [archive] du livre XVIII (*Traitant des céréales*) de l'Histoire naturelle de Pline, sur le site remacle.org. *Remarque* : dans un premier temps, cette note a soulevé un questionnement. En effet, Jean-Baptiste Rougier, Baron de la Bergerie, a publié *Histoire de l'agriculture ancienne des grecs, depuis Homère jusqu'à Théocrite*, mais seulement en 1830. Ce qui accentuait la confusion, c'est que le même Rougier a publié en 1804, l'année de parution des *Principes d'agronomie et d'économie*, un ouvrage consacré aux *Géorgiques* de Virgile, ce dernier faisant partie des références abondamment citées par Chrestien de Lihus.
2. ↑ Histoire de la Chine, par Martini	2. ↑ Martino Martini, jésuite italien, missionnaire et premier géographe et cartographe de la Chine, auteur de *Sinicae historiae dicas prima* (Munich, 1658). Texte intégral [archive] sur le site European Cultural Heritage Online (Max Planck Institute for the History of Science).
3. ↑ Vita rustica justiciae magistra est. Cic. pro Rosc. n° 39 et 75.	3. ↑ Ceci fait probablement référence à la plaidoirie de Cicéron lors du procès de Sextus Roscio Amerino, connu sous l'appelation *pro Roscio Amerino*. Texte intégral [archive] (en latin) sur le site thelatinlibrary.

Figure III : Une note complémentaire signale le questionnement soulevé initialement par la difficulté d'attribution de l'ouvrage cité dans la note originelle.

Ce travail sur le texte, ne nécessitant pas de compétences "disciplinaires" (bien que l'ouvrage traite d'agronomie, il n'est pas nécessaire d'être agronome pour apporter les éléments qui viennent d'être décrits), a mis en lumière l'intérêt de demander également à des spécialistes du(des) domaine(s) concerné(s) (ici, on peut imaginer faire appel à des agronomes, à des historiens, à des géographes...) de venir apporter leurs propres commentaires, afin d'enrichir encore la lecture du document. À titre d'exemple (figure IV), Pierre Morlon a accepté de se livrer à cet exercice, sur le thème de la jachère, d'une part (partie I, chapitre 2), et sur l'affouragement en vert des chevaux (partie III, chapitre Juin).

[138]

Chevaux au vert.

Lorsqu'on a d'abondans fourrages près de la maison, il est très-bon de mettre les chevaux au vert pendant trois semaines ; cela les purge et leur fait, pour ainsi dire, un corps neuf : on leur en donne à discrétion, c'est-à-dire au moins trois bottes chacun ; ce qui n'empêche pas de leur donner de l'avoine comme à l'ordinaire. Le trèfle est la meilleure nourriture en vert pour les chevaux, parce qu'elle est la plus purgative et la plus tendre à manger.

Bestiaux.

On continue toujours à nourrir les vaches au fourrage vert, et à les mener dans un herbage matin et soir. Lorsque la luzerne commence à sécher, il ne faut plus leur en donner, parce que ce fourrage n'a plus de jus et ne peut plus fournir de lait. Il faut donc leur donner du trèfle vert, qu'on gardera sur pied en assez grande quantité pour attendre la seconde coupe de luzerne, ou le mélange qu'on aura semé pour remplacer le trèfle.

Commentaire - Agronomie

Thème : Affouragement en vert des chevaux [Enrouler]

La recommandation de Chrestien de Lihus dans ce paragraphe consacré aux chevaux consiste à appliquer le principe de l'affouragement en vert avec du fourage coupé.

Pierre Morlon - 15 juin 2011

Figure IV : Une note "disciplinaire", commentaire transmis par un agronome (la note est ici déroulée, lorsqu'elle est enroulée (position initiale), seuls la discipline, le thème, l'auteur et la date sont visibles).

Ainsi s'est effectué, progressivement, le glissement d'un test de simple réédition d'un ouvrage ancien destiné à donner à chacun la possibilité de travailler sur ce texte non récupérable par d'autres moyens, vers une expérimentation plus complète de réédition commentée et enrichie, qui se rapproche davantage d'une édition critique.

5. Perspectives : le réseau Wicri et l'édition hypertexte de ressources textuelles

La plupart des projets existants d'édition hypertexte semblent, comme on vient de le voir, intégrer une étape de numérisation des données, ce

qui induit assez logiquement de se concentrer sur la question du traitement des images.
Cependant, dans le cas de la démarche qui est l'objet de cet article, la question de l'acquisition des données ne se posait pas, et pouvait être considérée comme annexe, sinon négligeable. Il était sensiblement plus important de se concentrer sur la question de la mise à disposition du résultat "final" (dans le cas d'un travail collaboratif sur un wiki, la notion de résultat final ne recoupe pas celle d'un résultat qui serait "définitif").
Ainsi, on peut parfaitement imaginer appliquer cette démarche, que les données soient déjà disponibles sous la forme d'un texte exploitable (comme c'était le cas pour les Principes d'agriculture et d'économie), qu'il s'agisse d'une source déjà numérisée et pour laquelle il n'existe pas de version exploitable autrement que par du traitement des images, ou qu'il s'agisse d'une source pour laquelle il n'existe ni texte exploitable, ni numérisation.
De fait, la question de l'acquisition des données n'a pas d'influence sur le traitement ultérieur : elle modifie uniquement – même si c'est déjà important – les questions de timing et de moyens nécessaires.
Ainsi, on peut imaginer plusieurs modes de fonctionnement, en fonction du contexte. Pour des chercheurs ou amateurs éclairés soucieux de donner accès à la communauté à des ressources "rares" dont ils disposent – et qui sont donc dans la position qui était la nôtre au début de cette expérimentation, dans une perspective d'édition de "service public" -, il est possible de proposer un espace de mise en ligne, des outils d'enrichissement et un accès à une communauté d'experts. Pour des bibliothèques, des institutions, des sociétés savantes... elles peuvent bénéficier des mêmes éléments, associés à un soutien technique renforcé sur la phase d'édition, afin d'accompagner leurs projets d'édition hypertexte. Enfin, l'équipe Wicri peut s'associer à des projets de recherche qui nécessiteraient le développement de nouvelles fonctionnalités.
La majorité des projets présente également un volet de traduction (à partir du latin, du grec). Cet aspect mérite que l'on s'y arrête un instant. En effet, cette phase de travail est en général totalement invisible pour le lecteur, qui veut accéder à la consultation parallèle du texte initial et de sa traduction, mais peut être sensiblement enrichie si le travail de traduction s'effectue de façon collaborative. De plus, le décodage même des opérations de traduction (sous la forme d'une "trace") serait potentiellement riche d'enseignement pour des lecteurs experts.
Dans cette optique, nous suggérons un travail en deux étapes : préparation et traduction du texte sur un wiki privé, accessible aux experts identifiés et apparentés au projet, puis mise à disposition des textes sur un wiki public, sur lequel se ferait alors le travail d'annotation décrit dans notre expérimentation.

On constate enfin que la plupart de ces projets ont des débouchés en terme de recherche (philologie, analyse critique, mise en perspective d'une oeuvre...) et en terme d'enseignement, offrant aux pédagogues de diverses disciplines des moyens nouveaux – et inaccessibles jusqu'ici – d'exploiter des sources anciennes, que ce soit pour les mettre en avant, ou pour en critiquer les manques. La démarche décrite ici offre, pour un coût et dans des délais particulièrement raisonnables, la possibilité à ces deux communautés d'étendre encore le champ des possibles. Notamment dans le cas de textes pour lesquels l'acquisition des données est déjà effective, la solution existante est non seulement simple mais complète. On peut ainsi estimer, pour un ouvrage de taille moyenne (500 pages) le temps total de mise en ligne à 2 mois.
La mise en place progressive (un wiki est un espace de "chantier" autant que de "versions finales") permet en outre de stimuler l'action des divers spécialistes qui peuvent être sollicités pour travailler sur les commentaires critiques disciplinaires. Notre expérimentation n'a pas étudié la possibilité d'exploiter toutes les fonctionnalités - notamment sémantiques - déjà existantes des wikis. Il est néanmoins clair que celles-ci pourraient être mobilisées afin de développer des outils d'analyse pour la recherche (indexation des termes, traitement des auteurs...). Et cela sans parler, naturellement, d'éventuels développements susceptibles de générer des fonctionnalités spécifiques.

6. Conclusion

D'une expérimentation simple et qui s'inscrivait dans un contexte d'édition de service public, destinée à donner un accès à tous à des ressources textuelles non disponibles sur internet, nous sommes, comme on peut le voir, passés à un outil permettant de fonctionner collaborativement sur différents niveaux d'annotation, pouvant intégrer une phase de travail collectif (par exemple sur la traduction des œuvres), avec une traçabilité fine des actions des divers acteurs.
Sans prétendre a priori pouvoir répondre à tous les besoins, il nous apparaît, à tout le moins, que la technologie des wikis telle que nous l'exploitons dans le cadre du réseau Wicri offre une possibilité intéressante de mise à disposition de textes, autant du fait de sa souplesse que de sa simplicité de mise en œuvre.
Il deviendrait alors possible, pour reprendre Victor Hugo, de faire en sorte que le livre, comme livre, n'appartiennent plus uniquement à l'équipe de recherche qui travaille dessus, mais bien à la communauté élargie des chercheurs de toutes disciplines.

Bibliographie

[BUQ 2004] Thierry BUQUET, « Quelques réflexions autour de la chaîne éditoriale d’un document numérique : l’exemple de La Lettre volée », *Le Médiéviste et l’ordinateur*, 43, 2004 [http://lemo.irht.cnrs.fr/43/43-04.htm].

[BUR 2009] Hyperdonat, une édition électronique des commentaires de Donat aux comédies de Térence. Bruno BUREAU, Maud INGARAO, Christian NICOLAS, Emmanuelle RAYMOND (dir.), CEROR, Université Lyon III, ENS de Lyon, 2007-2011. Accédé en ligne le 24 juin 2011, [http://hyperdonat.ens-lyon.fr].

[CLE 2007] Jean CLEMENT, L’hypertexte, une technologie intellectuelle à l’ère de la complexité, in Brossaud Claire, Reber Bernard, *Humanités numériques 1., Nouvelles technologies cognitives et épistémologie*, Hermès Lavoisier, 2007.

[DOR 2009] Stéphanie DORD-CROUSLE et Emmanuelle MORLOCK-GERSTENKORN, *L’édition électronique des dossiers de Bouvard et Pécuchet de Flaubert : des fragments textuels en quête de mobilité*, publié dans « Le patrimoine à l'ère du numérique : structuration et balisage » organisé à Caen les 10 et 11 décembre 2009.

[DRA 2009] Communication à la journée d'études : Digital Edition of Sources in Europe: Achievements, (juridical and technical) Problems and Prospects, à l'occasion des 175 ans de la Commission Royale d'Histoire. Meeting Porta Historica. [http://www.crhistoire.be/fr/partenariat/intern/portaPres_fr.html].

[DUC 2010] Jacques DUCLOY, Thierry DAUNOIS, Muriel FOULONNEAU, Alice HERMANN, Jean-Charles LAMIREl, Stéphane SIRE, Jean-Pierre THOMESSE et Christine VANOIRBEEK, *Métadonnées pour WICRI, un réseau de wikis sémantiques pour les communautés de la recherche et de l'innovation*, rapport de projet présenté au colloque DC 2010 (Pittsburgh, Etats-Unis). Version française consultable sur le wiki Wicri/Ticri [http://ticri.inpl-nancy.fr/ticri.fr/index.php/M%C3%A9tadonn%C3%A9es_pour_WICRI,_un_r%C3%A9seau_de_wikis_s%C3%A9mantiques_pour_les_communaut%C3%A9s_de_la_recherche_et_de_l%27innovation].

[KAL 2000] Enriketa KALLDRËMXHIU, *Les logiciels de numérisation des livres anciens*, Technical report, Université Claude Bernard Lyon1, 2000. [www.letterpress.ch/APINET/IMMPDF/LIVRE/gedkall.pdf] (pdf).

[LER 2008] Françoise LERICHE et Cécile MEYNARD , « Introduction. De l’hypertexte au manuscrit : le manuscrit réapproprié », *Recherches & Travaux* , 72 | 2008, mis en ligne le 15 décembre 2009. [http://recherchestravaux.revues.org/index82.html]. Consulté le 29 juin 2011.

[POR 2010] Pierre-Édouard PORTIER et Sylvie CALABRETTO. *DINAH, a philological platform for the construction of multi-structured documents*, in The European Conference on Research and Advanced Technology for Digital Libraries (ECDL), Mounia Lalmas, Joemon Jose, Andreas Rauber, Fabrizio Sebastiani, Ingo Frommholz ed. ECDL 2010 September 6 - 10, 2010, Glasgow. pp. 364-375. Research and advanced technology for digital libraries LNCS. Springer. ISBN 978-3-642-15463-8. ISSN 0302-9743. 2010. [http://liris.cnrs.fr/membres/?idn=peportie&onglet=publis].

[POU 2006] Gautier POUPEAU, *Les apports des technologies Web à l'édition critique : l'expérience de l'Ecole des chartes*, présenté à Digital philology and medieval texts, 01/2006 (Arezzo, Italie). [http://halshs.archives-ouvertes.fr/view_by_stamp.php?&halsid=p27rt2a5en6gcskfqnrpughlu0&label=S

HS&langue=fr&action_todo=view&id=sic_00137229&version=1&view=extended_view].
[RAI 2008] Ludivine RAIMONDO, *Enjeux et représentations de la science, de la technologie et de leurs usages* - rapport ENS Lyon
[SCH 2010] François-Joseph BERARDIER DE BATAUT, *Essai sur le récit, ou Entretiens sur la manière de raconter* (Paris : Charles-Pierre BERTON, 1776). Édition électronique sous la direction de Christof Schöch, 2010. [http://www.berardier.org] (Version 0.6, 12/2010).