HAL
open science

Revitalisation numérique du patrimoine littéraire

Jean Jon, J. Casenave

To cite this version:

Jean Jon, J. Casenave. Revitalisation numérique du patrimoine littéraire. . artxibo-00000044

HAL Id: artxibo-00000044

https://artxiker.ccsd.cnrs.fr/artxibo-00000044

Submitted on 9 Dec 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

REVITALISATION NUMERIQUE DU PATRIMOINE LITTERAIRE TERRITORIALISE

Dans tous les pays de l'Europe de l'ouest, les institutions nationales[1] prennent des initiatives pour mettre leur patrimoine culturel d'intérêt national ou universel en valeur par le biais des technologies numériques et du web. En France, le Musée du Louvre, la Bibliothèque Nationale de France et son serveur électronique Gallica, l'Institut National de l'Audiovisuel et l'ensemble des grands établissements offrent des services variés qui vont de la visite virtuelle à la mise en ligne de contenus téléchargeables ou consultables, en passant par une gamme étendue d'activités adaptées aux divers publics potentiels.

En revanche, la valorisation des fonds de notoriété moindre est beaucoup plus ardue, qu'il s'agisse de corpus relevant de registres « exigeants » ou plus circonscrits en termes de rayonnement géographique. Pour s'en tenir à l'exemple français, on peut toutefois citer quelques tentatives intéressantes à l'image de celles que proposent la Bibliothèque électronique de Lisieux[2] et le service des Archives de la Région Savoie sur son serveur électronique *Saubadia*[3]. Cependant, le domaine demande à être développé car le patrimoine dont disposent ces espaces territoriaux est aussi riche que varié et son potentiel de valorisation tout à fait incontestable.

C'est pourquoi, dans le cadre de ce colloque consacré au cycle de vie du document, nous proposons une réflexion sur la mise en valeur des fonds territorialisés –locaux, régionaux- et tout particulièrement des ensembles documentaires constitués autour des corpus littéraires anciens. Nous étudierons les conditions de « revitalisation » culturelles et techniques d'un corpus littéraire « oublié ». Nous nous efforcerons de montrer que, dans ce nouveau « marché des biens culturels » ouvert à la concurrence, il y a bien une place pour l'expression des singularités régionales dès lors qu'il existe un projet de socialisation conçu au plus près des ressources numérisées, dans un esprit de coopération entre les instances en charge du développement des territoires et les équipes de recherche spécialisées dans les contenus et les technologies de la communication.

Nous traiterons d'abord de l'intérêt patrimonial des ces fonds documentaires pour aborder, ensuite la liaison entre la conversion numérique des fonds, la problématique de leur diffusion et celle des dispositifs techniques et conceptuels qui président à leur communication, Par leur ancrage dans les lieux qu'ils évoquent, ils livrent d'indispensables éléments de connaissance objective mais aussi de précieuses représentations d'ordre subjectif. Indubitablement, ils contribuent à enrichir ce nouvel imaginaire des territoires que les nouveaux récits pédagogiques, médiatiques ou touristiques s'efforcent de reconstituer, bien trop souvent dans l'ignorance des sources bibliographiques disponibles. A l'heure de l'élargissement de l'Europe, ce repositionnement de « l'archive territorialisée » dans le discours culturel représente un enjeu important. Par leur conversion au format électronique,

---

[1] Bibliothèque Nationale de France : http://gallica.bnf.fr
[1] Archives Littéraires Suisses : http://www.snl.ch/f
[1] Bibliothèque Nationale du Luxembourg, Centre National de Littérature : http://literaturarchiv.lu.fr
[1] Bibliothèque Royale de Belgique : http://www.aml.cfwb.be
Etc.
[2] Bibliothèque Municipale de Lisieux : http://bmlisieux.com
[3] Archives Départementales de Savoie : http://www.sabaudia.org

ces ressources deviennent l'objet d'un partage élargi de l'information et des savoirs qui va bien au-delà des communautés scientifique et éducative. Leur socialisation touche à l'espace public dès lors qu'elle entre dans le champ du développement économique régional par le biais des activités touristiques ou par son ouverture au grand public.

Ce sont donc là autant de nouveaux usages à l'adresse de nouveaux publics et les technologies informatique et multimédia permettent d'assurer ce polymorphisme au document traité. Pour éclairer ce point technique dans la deuxième partie de notre exposé, nous aborderons la question de la préparation conceptuelle et technique de ces documents.

Qu'il s'agisse de littérature populaire, savante ou encore de représentations iconographiques, l'omniprésence des lieux caractérise le fonds local ou territorialisé et les documents fourmillent de notations et de repères spatiaux (onomastique, séquences descriptives, etc.). Compte tenu de la spécificité de l'archive territorialisée, nous avons choisi d'expérimenter une didactisation à partir de la sémiotique du territoire. Pour être véritablement opératoire, la préparation informatique du document (pose de balises ; appareillage conceptuel, technique et ergonomique) doit s'opérer dès la numérisation. En fonction de paramètres tels que le genre du « document-source » ou celui des publics ciblés, la préparation de l'application numérique devrait pouvoir se faire selon des protocoles reproductibles et extensibles à un large éventail de ressources relevant des corpus territorialisés.

L'équipe de recherche interdisciplinaire[4] que nous avons constituée pour mener à bien cette étude de la didactisation des corpus territorialisés (textes littéraires, ressources iconographiques et phonographiques) se compose de chercheurs spécialisés dans les domaines de la langue et de la littérature, de l'informatique et du multimédia et elle pourrait prochainement recevoir l'appui de spécialistes de l'aménagement du territoire et de la gestion de projet.

Première partie : Conditions de resocialisation numérique d'un corpus restreint.

1. La question des « corpus restreints » :

1.1. Une problématique spécifique, des contraintes économiques :

Compte tenu des coûts élevés que représente le traitement informatique des fonds « papier », les corpus restreints ou d'intérêt moindre au plan du rayonnement géographique, linguistique, littéraire ou culturel ne peuvent pas bénéficier des mêmes possibilités de numérisation que les œuvres perçues comme majeures : ils ne possèdent pas la même

---

[4] Cette équipe est constituée par des chercheurs appartenant aux laboratoires indiqués ci-dessous :
- IKER est une Unité mixte de recherche (5478) qui regroupe des membres du CNRS et des enseignants chercheurs de l'Université de Bordeaux 3 et de l'Université de Pau et des Pays de l'Adour. Centre de recherche sur la langue et les textes basques, ce laboratoire travaille sur le domaine basque à partir de directions de recherche variées.

- LIUPPA est un laboratoire d'informatique… *à compléter*.

Au cours de l'année 2003, ces deux équipes installées sur le campus universitaire de Bayonne ont décidé de s'associer autour d'un thème de recherche commun : *la valorisation du patrimoine local dans la culture numérique*. L'étude proposée à l'occasion de ce colloque a été menée dans le cadre de ces activités communes.

« légitimité culturelle » auprès du public et la même reconnaissance institutionnelle. Dans l'esprit du plus grand nombre comme pour beaucoup de spécialistes –institutionnels et professionnels confondus-, cette généralisation s'applique essentiellement aux corpus consacrés par les institutions éducatives et éditoriales.

En effet, la numérisation des « classiques » entre dans les usages contemporains sans autre justification que leur appartenance au « canon » littéraire national voire universel La transposition électronique de l'ouvrage d'un auteur considéré comme secondaire ou celui d'un romancier en voie de consécration sera difficile à justifier auprès d'un grand établissement de diffusion.

Il en est de même pour des genres moins prisés aujourd'hui comme la poésie (à l'exception des valeurs sûres) ou encore la production poétique, théâtrale ou journalistique de certains écrivains consacrés des XVIIIe et XIXe siècles. En effet, ce sont toujours les mêmes titres (romans, essais, recueils de poèmes, d'articles ou pièces de théâtre en fonction de l'image que notre époque se construit de cet auteur) qui profiteront des possibilités de publication et, par voie de conséquence, auront quelque chance d'accéder à la numérisation. A ces cas particuliers, il faut ajouter d'autres corpus restreints qui sont écrits dans des langues rares et peu répandues ou qui présentent une spécificité géographique. A titre d'exemple, nous développerons ici la problématique qui est attachée aux fonds locaux et territorialisés.

1.2. Un exemple de corpus restreint : le corpus « local » ou territorialisé :

Toute bibliothèque, en ville comme dans les bourgs ruraux réserve une partie de sa surface d'exposition aux livres qui évoquent la ville, la région, le terroir, bref, un territoire plus ou moins bien identifié, sujet de prédilection d'un ensemble de publications que les spécialistes s'accordent à désigner par l'expression « fonds local ». Nous l'appellerons plutôt « territorialisé » pour éviter le caractère restrictif et parfois condescendant qui est attaché au terme « local » par opposition à « national » ou « universel ». Le qualificatif « territorialisé » indique aussi que ce corpus s'inscrit explicitement dans cet espace déterminé, il s'en inspire et s'en imprègne, il en explore les dimensions physiques et oniriques, déploie son imaginaire et le met en scène avec bien souvent un luxe de détails qui fait la spécificité de ce type de document.

Aux ouvrages présentés au public, il faut généralement ajouter le nombre considérable des livres qui figurent dans les réserves et qui, n'apparaissant que sous la forme d'une mention dans le catalogue de l'établissement, ont bien peu de chances d'attirer l'attention du lecteur ordinaire. Au côté d'ouvrages très divers (histoire locale, géographie locale, économie, traditions, « beaux livres » essentiellement photographiques, etc.), la littérature (contes et littérature populaire écrite et orale, récits de voyages, romans ruraux et régionalistes, iconographie du territoire) occupe une place de choix. Prisés par un public toujours renouvelé mais relativement restreint de curieux, d'érudits, d'universitaires ou d'étudiants, les ouvrages littéraires qui garnissent ces rayons « régionalistes » ou « territorialisés » ont a priori peu de chances d'accéder au nombre des rééditions. Comme nous le verrons plus loin, ces divers points lui confèrent un caractère exemplaire et constituent un bon potentiel de « revitalisation ».

1.1.3. Un champ d'expérimentation interdisciplinaire :

Cette réflexion à laquelle nous avons souhaité donner le caractère le plus général possible s'appuie sur une expérience interdisciplinaire bien précise. En effet, la collaboration entre les laboratoires cités plus haut s'est concrétisée –à l'occasion d'un appel à projets lancé par les

institutions locales (département, région)- auquel nous avons répondu en proposant le thème de la valorisation numérique du patrimoine local. Nous avons engagé des recherches dans diverses directions et celle de la resocialisation numérique des fonds bibliographiques, iconographiques et phonographiques du domaine basque relatif à la première moitié du XXe siècle nous est apparue à la fois riche au plan documentaire et très prometteuse en termes d'interdisciplinarité. Cela nous a conduit dans un premier temps à définir les bases d'un traitement interdisciplinaire des documents puis, en fonction de ces critères à la fois théoriques et techniques (corpus « oublié », redoublé par une documentation iconographique et phonographique, etc.) , à proposer la numérisation d'un « corpus aux contours bien définis » afin d'étudier les possibilités de valorisation informatisée qu'il recelait.

En conséquence, dans le cadre d'une collaboration étroite entre spécialistes des documents (textes et des contenus iconographiques et phonographiques) et informaticiens, nous avons sélectionné un ensemble de textes issus de la littérature populaire du XIXe siècle (contes, proverbes, chansons) et un petit nombre de romans en langues basque et française[5] qui décrivent tous abondamment le Pays basque . Ces derniers ont en commun d'avoir tous été publiés entre 1893 et 1946 et appartiennent donc au courant littéraire régionaliste qui a connu son heure de gloire dans la première moitié du XXe siècle. Ils présentent tous la caractéristiques d'avoir connu une certaine notoriété à l'époque de leur parution mais n'ont pas été republiés depuis la fin de la seconde guerre mondiale.

La remise en circulation de ces fonds documentaires représente un défi technique et économique intéressant ainsi qu'un enjeu culturel important. Les propositions qui résultent de cette réflexion interdisciplinaire sont autant d'hypothèses de travail, sans doute partagées par la plupart des équipes de chercheurs qui étudient la valorisation numérique d'un « corpus restreint ». C'est ce qui nous a amené à penser que les problématiques auxquelles nous nous sommes heurtés à propos de notre corpus territorialisé étaient, dans une large mesure, transposables à la plupart des cas de figure présentés par les corpus restreints.

1.2. Les caractéristiques du corpus territorialisé :

Chaque corpus restreint possède à la fois ses atouts et ses faiblesses. Le recensement de ces éléments et leur étude interdisciplinaire est fondamental dans la perspective de la valorisation numérique du corpus. Comme on le verra ci-dessous, les points forts du corpus, l'ancrage territorial en l'occurrence, donneront prise sur les documents et constitueront l'axe principal du travail interdisciplinaire.

1.2.1. Un fonds inscrit dans un ensemble historico-géographique :

Dans la tradition littéraire française, G. Sand pour les terres du centre de la France, J. Giono pour la Provence, E. Le Roy pour le Périgord, H. Pourrat pour l'Auvergne, P. Loti pour le Pays basque et la Bretagne viennent, immédiatement à l'esprit lorsqu'il s'agit d'associer une œuvre à un territoire. Et il y a les innombrables ouvrages d'auteurs oubliés qui ne demeurent vivants que pour quelques dizaines de lecteurs « spécialisés ». Plus ou moins longuement après leur parution, ces livres ont connu une notoriété nationale, régionale ou

[5] En matière de fiction, la plus grande part de la création de l'entre-deux-guerres est d'expression française.

seulement locale. Puis, non réédités et tombés dans l'oubli, ils dorment dans les réserves des bibliothèques et ne trouvent un délai de grâce que chez les bouquinistes où les habitués sont souvent davantage collectionneurs que lecteurs.

Cependant, ces fonds constituent des sources d'information précieuses à de nombreux titres. Inscrits dans un ensemble historico-géographique déterminé, ils font référence de façon dense et précise à des lieux qui sont nommés de manière explicite : toponymes bien sûr mais aussi noms de points d'eau ou de rivières, noms de villages ou de villes, noms de maisons aussi, dans le domaine basque très souvent en étroite relation avec les noms personnages qui sont également des domonymes. Ces documents abondent en descriptions particulièrement précises du territoire, des paysages mais évoquent aussi les mœurs et des activités laborieuses comme festives.

Le corpus des cartes postales, celui des images en général (photographiques, cinématographiques et vidéographiques), d'une manière souvent redondante par rapport à l'écrit, est à son tour venu « consacrer » jusqu'à en faire des « représentations stéréotypées » les mêmes paysages (le rivage, la montagne, la forêt, la ferme, etc.), les mêmes activités ludiques (la partie de pelote, la danse, la pastorale ou la foire, etc) ou les mêmes personnages (le chanteur, le berger, le bouvier, le grand-père, etc.). A l'intérieur comme pour l'extérieur, c'est donc l'image d'un « terroir » aujourd'hui « vécu »[6] comme territoire qui s'est constituée autour des représentations proposées par ce fonds « papier » et par ses adjuvants. Bien que remplacé par les nouveaux discours du territoire dont il sera question plus loin, on peut déjà avancer l'idée que l'imaginaire local contemporain est encore largement influencé par cette source, et ceci sans le savoir et sans y faire directement référence, simplement en s'appuyant sur les motifs stéréotypés qui ont perduré et qui sont invariablement repris avec peu de recul critique.

A un moment où la notion de « territoire»[7] relance l'ancien concept de « terroir » d'une façon largement renouvelée et ajustée à nos préoccupations actuelles par la géographie humaine et l'anthropologie contemporaines toutes les évocations localisées retrouvent un intérêt, qu'elles soient textuelles, iconographiques ou sonores. C'est sous cet éclairage qu'il est possible d'opérer une relecture et une réévaluation de ces fonds dans le champ culturel contemporain. De plus, l'informatisation des corpus documentaires renouvellent sérieusement les opportunités de resocialisation et de valorisation à leur endroit.

1.2.2. <u>Un fonds « oublié » et invisible</u> :

Cependant, avant d'envisager leur « renaissance », il faut savoir ce qu'il en est, à l'heure actuelle, de l'existence des corpus locaux ou territorialisés des périodes citées. Pour l'essentiel, ils sont devenus des documents d'archives et ne connaissent plus désormais que la fréquentation d'un public très réduit et spécialisé qui en fait une utilisation savante (recherche universitaire, érudition locale, source de références et dictionnaire de citations -souvent les mêmes- pour les journalistes).

Oublié et remisé dans les magasins des bibliothèques, ce fonds souffre donc d'un manque de visibilité. Il faut ajouter à cela, un indéniable déficit de lisibilité pour le non spécialiste. La difficulté, comme pour la plupart des textes littéraires ou documentaires du passé, tient bien

---

[6] Cette expression fait bien sûr référence au livre du géographe André Frémont, *La Région, espace vécu* (1976), ouvrage pionnier en la matière mais aussi à ses études sur la Normandie telle qu'elle apparaît dans les récits de Flaubert et de Maupassant.

[7] Dans le cadre de cet article, il n'est pas possible d'intégrer des développements nourris de l'apport effectué sur ces questions par les géographes (Brunet, Lévy, etc. cf. Bibliographie) les historiens et les anthropologues et les ethnologues (Certeau, Zonabend, etc. cf. Bibliographie).

sûr à l'obstacle linguistique et stylistique, c'est-à-dire à un état de langue et de style littéraires désormais très daté.

Mais elle provient également des contenus eux-mêmes car le discours tenu sur « l'espace local » peut sembler lui aussi très daté puisqu'il participe d'un mouvement de « singularisation » à outrance des « terroirs » bien caractéristique des courants folkloristes et régionalistes de la première moitié du XXe siècle mais difficile -voire impossible- à comprendre pour les lecteurs contemporains.

Autant dire que la remise à disposition du grand public de ce corpus nécessiterait un important travail d'historicisation, notamment à travers la mise au point d'un appareil critique. Un premier pas vers la revitalisation de ce corpus consisterait à le répertorier de façon systématique dans le catalogue électronique des établissements, ce qui n'est pas le cas pour l'instant, puisque, pour la plupart, les œuvres ne figurent que dans le catalogue « papier ». Cela leur permettrait d'accéder au statut « d'archive littéraire»[8] et les rendrait visibles et, par voie de conséquence, consultables ne serait-ce que sous sa forme « papier ». Cependant, il est indiscutable que ces fonds possèdent d'autres atouts de revitalisation.

1.2. 3. Le potentiel de revitalisation de ce fonds :

Comme indiqué plus haut, ce corpus porte en lui une connaissance approfondie des lieux et du territoire qu'il évoque. Ce point constitue le cœur de son potentiel de revitalisation. Malgré son caractère daté, il continue à faire sens en termes de données géographiques, historiques, ethnographiques et littéraires par exemple. Les nouveaux récits pédagogiques, médiatiques ou touristiques du territoire s'efforcent de reconstituer ces savoirs et, sans nul doute, ils gagneraient à se ressourcer dans les œuvres originales pour étoffer leurs descriptions et les appuyer de citations plus variées.

Paradoxalement, c'est même son caractère « daté » qui en fait le prix. En effet, il représente une mémoire de ces espaces et des visions dont ils ont fait l'objet dans un passé somme toute encore récent. Inscrits dans des représentations concrètes du territoire ou relevés comme les affleurements d'un imaginaire et d'une subjectivité plus intimistes, ces repères littéraires ou culturels situés au cœur des documents sont susceptibles d'alimenter ce « grand récit du territoire » qui se développe dans la littérature contemporaine ou dans la presse spécialisée (magazines consacrés à des territoires spécifiques ; ex : *Alpes magazine*, *Pyrénées Magazine Pays Basque Magazine*, *Bretagne Magazine*, etc), que dans les publications les plus variées (anthologies, albums photos, nombreuses publications liées à la découverte de l'environnement, etc), dans le matériel pédagogique préparé à l'intention des classes par les services éducatifs spécialisés et dans les manifestations culturelles et touristiques.

Bien entendu, la lisibilité du corpus doit être assurée par un appareil critique adéquat. Les documents cités participent donc pleinement d'une sémiotique du territoire qui, comme on le verra plus avant, peut devenir le dénominateur commun pour les utilisateurs comme le concept de traitement des données des concepteurs de la didactisation.

Les raisons a priori objectives de la remise en circulation de ce corpus paraissent donc bien établies ; encore faut-il voir maintenant à quelles conditions il est possible de réintroduire ces données dans le flux des savoirs et des informations qui circulent dans la société contemporaine. C'est aussi l'occasion d'observer comment elles peuvent s'intégrer dans la culture numérique en cours de constitution.

1.3. La resocialisation du corpus :

[8] Dans ce domaine de l'archive littéraire, il faut suivre l'évolution d'une initiative particulièrement intéressante portée par sept universités canadiennes : Projet IRMA : http:/www.unites.uqam.ca/irma/

Si l'on envisage de resocialiser un corpus restreint, il faut y trouver des entrées qui conviennent aux utilisateurs actuels « réels » c'est-à-dire déjà actifs comme potentiels (extension du nombre et des types d'utilisateurs). Il s'agit donc de trouver des dénominateurs communs qui rapprochent les diverses communautés.

1.3.1. Numérisation et « réalisme économique » :

La numérisation représente pour ces fonds documentaires (bibliographiques, iconographiques, phonographiques) oubliés la seule opportunité d'ensemble véritablement réaliste d'une resocialisation et d'une « revitalisation ». En effet, il n'est pas envisageable de faire republier un ensemble de textes désormais difficilement « lisible ». Certes, quelques maisons d'éditions ont tourné la difficulté en publiant des morceaux choisis ou en constituant des ouvrages thématiques ou encore en proposant des textes peu connus d'auteurs célèbres[9]. Toutefois, il est impensable de mettre en place une politique éditoriale « papier » qui permettrait une resocialisation d'un corpus étendu.

En revanche, la transposition numérique peut offrir de nouvelles perspectives. En effet, la conversion au format électronique modifie grandement les modes d'utilisation traditionnels puisque sa transformation en fichier informatisé scelle son polymorphisme et sa « banalisation » parmi bien d'autres ressources possibles (iconographiques et sonores : fonds photographiques, cartographiques, cinématographiques, etc.). Dès lors, ces documents numérisés deviennent utilisables pour eux-mêmes avec toutes les commodités que procure le format électronique : consultation à distance, téléchargement, annotations, etc. Ils peuvent aussi faire l'objet d'une valorisation en association avec les autres ressources évoquées plus haut, elles aussi informatisées.

Encore faut-il que les modalités de diffusion de ces ressources soient trouvées. Cela implique la nécessité de concevoir un projet de valorisation des corpus numérisés au-delà du seul souci de préservation d'un héritage du passé, sous la forme d'archives.

1.3.2. Le repositionnement du « corpus local » dans le champ culturel :

Il s'agit donc de rendre les contenus de ce corpus au public, à travers un usage élargi et actualisé dans le cadre de la culture numérique. En quelque sorte, il faut le repositionner parmi les ressources actuellement disponibles à propos du territoire évoqué et lui donner un statut nouveau. Ce processus de resocialisation a deux aspects complémentaires, un versant technique et un versant conceptuel et ils sont tous les deux étroitement imbriqués.

Le changement de statut de ce fonds est complexe. Il doit retrouver une pertinence qu'il a perdue au milieu des nouveaux récits du territoire évoqués. Comme indiqué plus haut, ceux-ci proviennent de milieux différents mais ces discours aussi bien journalistique, politique, économique, éducatif que touristique s'appuient tous sur la même entité territoriale et partagent avec le discours littéraire du passé la même volonté de mettre cet espace plus ou moins défini en valeur pour asseoir leur légitimité.

---

[9] On peut citer l'exemple d'une petite maison d'édition installée à Saint-Jean-de-Luz, les Editions *Pimientos*, qui publient des recueils d'extraits de textes « oubliés » d'auteurs célèbres. Consacré au Pays basque, *Voyage au Pays basque* (1999), le premier de la série, rassemble des pages prises chez Hugo, Stendhal, Gautier, Flaubert, Taine, Mérimée, Viollet-le-Duc.

La resocialisation passe nécessairement par deux opérations qui sont grandement facilitées par la transposition des œuvres en version électronique. En effet, grâce à la numérisation, l'œuvre devient visible. Par le simple fait de la faire apparaître dans un catalogue électronique, la ressource reprend vie. Dès lors, elle peut circuler à titre de référence. On peut aussi envisager sa mise à disposition totale auprès du public puisque la numérisation a transformé l'œuvre en fichier informatique. Il est donc possible d'afficher le texte, de le télécharger, de le manipuler au sens propre du terme, de l'annoter, de le fractionner pour l'utiliser à des fins particulières. Mieux, on peut aussi penser à le transformer avant de le proposer à nouveau à la lecture. Bref, après l'avoir rendu à nouveau visible, le processus de numérisation peut lui rendre une « existence » socialisée en le remettant en circulation à travers les voies largement ouvertes de l'Internet. Ainsi, dans le cadre de la culture numérique, la « ressource territorialisée » produite localement qui n'avait qu'un rayonnement limité se voit proposée à un public aléatoire et considérablement élargi, tout au moins potentiellement.

Enfin, cette visibilité renouvelée, cette remise en circulation décuplée par les moyens techniques contemporains propulse ce corpus « oublié » au rang d'archive littéraire répertoriée et utilisable par le plus grand nombre. De plus, elle transforme insensiblement ce fonds d'archives en patrimoine littéraire territorial. Ce nouveau statut d'objet patrimonial classe ces œuvres du passé au rang de « monument » à la fois source de témoignages dignes d'intérêt et de « lieu » (virtuel) qui mérite le détour pour une visite.

On voit sans peine que la numérisation a rendu le corpus considéré à sa vocation de référence tout en le remettant au cœur des activités culturelles ; en un mot, elle l'a redynamisé, revitalisé. Toutefois, dans le même temps, elle a modifié le rapport à ces ressources, leur utilisation et, par voie de conséquence, leur interprétation. On peut penser qu'au prix de ces transformations, ces documents ont été rendus à leur public potentiel à savoir, les universitaires et leurs étudiants, les érudits, les professionnels de la plume, écrivains et journalistes. Ceci, dans une forme actualisée qui facilitent grandement l'utilisation des documents. Pourtant, il ne fait guère de doute que les divers changements évoqués ouvrent de nouvelles perspectives aussi bien en matière d'usages nouveaux que d'élargissement du public susceptibles de fréquenter et d'utiliser ce corpus.

1.3.3. Nouveaux publics, nouveaux usages :

A l'occasion de nos recherches interdisciplinaires sur la numérisation du patrimoine littéraire attaché au domaine basque, nous avons travaillé sur la question de la resocialisation du corpus tant du point de vue des spécialistes des contenus que des concepteurs de l'équipement numérique des applications. Tout en sachant que la mise à disposition du corpus sous la forme d'application informatique en circuit contraint et surtout sur internet donne à l'éventail des publics et des utilisations possibles un caractère aléatoire et imprévisible, nous avons privilégié trois types de publics pour trois types d'usages.

Il s'agit de la communauté scientifique et « érudite »[10] pour un **usage savant** puis de la communauté scolaire pour un **usage pédagogique** et enfin, du grand public susceptible d'avoir un **usage informativo-documentaire**, voire **ludique**, du corpus, par exemple dans le cadre d'une exploitation touristique des ressources numérisées.

Il est indéniable que l'usage le plus attendu correspond à celui de la communauté scientifique puisque, qualifié ici de « savant », il se place dans le droit fil de l'utilisation traditionnelle de ce corpus territorialisé, à savoir, une étude approfondie du texte, de ces

[10] L'expression « communauté érudite » recouvre ici le nombre des lecteurs spécialisés qui sont habituellement qualifiés « d'érudits locaux ».

sources et du contexte qui l'a vu naître. Cette valorisation du document débouche souvent sur la réalisation d'une édition savante dotée d'un appareil critique qui en enrichit la lecture à défaut de toujours la faciliter. Conformément aux possibilités actuelles de l'outil informatique, ce corpus peut être apprêté pour une consultation savante sous la forme d'une édition critique hypertextuelle.

Dans la mesure où il existe au format électronique, ce corpus pourrait aussi donner lieu à une préparation pour une utilisation fragmentaire, éventuellement en association avec d'autres documents disponibles dans d'autres médias. Cette présentation peut intéresser de nouveaux publics qui ne feront pas du fonds une utilisation exhaustive à l'instar des utilisateurs traditionnels. Ainsi, c'est la perspective de toucher de nouveaux publics susceptibles d'expérimenter de nouveaux usages des documents qui découle de la transposition numérique.

Pour être en mesure d'aborder la question du partage des corpus avec de nouveaux publics, il faut auparavant évoquer les changements profonds que l'informatisation entraîne dans le rapport au texte et à l'œuvre littéraires. Dès lors qu'il est numérisé, le texte littéraire peut éventuellement être intégré dans une base de données extérieure à son environnement habituel et, dès lors, se pose la question de l'élargissement des publics susceptibles de s'intéresser à lui. En effet, il ne s'agit pas simplement de vouloir numériser des corpus parce que la possibilité technique existe. Il faut aussi savoir pourquoi et pour qui le faire. Il faut donc avoir un projet pour ce patrimoine, un projet de transmission, de partage. Il faut alors réfléchir aux conditions de cette ouverture, de ce partage.

Pour se limiter au patrimoine littéraire territorialisé et sans chercher le moins du monde l'exhaustivité, nous nous contenterons d'illustrer cet élargissement des publics et des usages à travers deux exemples situés aux deux extrêmes du spectre des activités rendues possibles par l'informatisation des ressources : la communauté éducative et les professionnels du tourisme.

Le texte peut fort bien être reçu de façon autonome comme celui de toute œuvre littéraire mais son ancrage territorial peut également lui permettre de constituer un précieux apport d'informations dans d'autres modalités de lecture. Qu'ils servent de repères, de descripteurs explicites ou même de caution culturelle sous forme de citations ou d'allusions dans un article de revue voire dans une brochure touristique, ces éléments prennent tout naturellement place au rang de références indispensables et parlantes.

Ils peuvent aussi de façon moins parcellaire être inclus dans la documentation pédagogique d'une classe-patrimoine par exemple parmi des ressources de types et d'origines très divers : collections photographiques, cinématographiques, vidéographiques, phonographiques, etc. En effet, ces repères linguistiques et littéraires doivent s'intégrer parmi d'autres sources de connaissance du territoire, qu'il s'agisse des données liées à la Nature (géologie, hydrologie, flore, faune,etc.), ou des études scientifiques au sujet des activités humaines traditionnelles ou contemporaines.

Depuis quelques années, de nouveaux lieux réels ou virtuels sont aménagés pour exposer ou utiliser les matériaux et les objets qui permettent d'accéder la connaissance des territoires et ils sont souvent mis en œuvre par les institutions locales (département, région) : musées, centres culturels, services pédagogiques des musées, classes patrimoines, etc. A ces lieux bien réels qui utilisent les technologies numériques dans leurs activités, il faut joindre par exemple, les expositions itinérantes ou permanentes qui sous la forme de bornes interactives, mettent la culture locale à la disposition du grand public. A cela, il faudrait encore ajouter les visites virtuelles des musées, les bornes consultables dans les offices de tourisme ou les sites web et autres portails de la culture qui intègrent le domaine territorialisé et qui touchent aussi bien le grand public local que celui du tourisme.

A travers l'évocation de cet élargissement des publics susceptibles de s'intéresser aux fonds locaux, il apparaît clairement que les mentalités vont devoir évoluer de façon notable.

Les littéraires et autres spécialistes des contenus ne vont plus pouvoir se tenir fermement sur leurs principes quant à l'intégrité du texte littéraire ou de l'œuvre d'art. Les méthodes de travail des informaticiens et des experts en communication vont nécessairement bousculer les habitudes de leurs homologues linguistes ou géographes et vice-versa. Pour être en mesure de proposer la valorisation numérique d'un corpus, il faut lui appliquer un traitement que de nombreux littéraires jugeront… contre nature, à savoir une analyse des besoins des publics, une étude de leur horizon d'attente, exigences et attentes plus diffuses confondues, un bilan de leurs compétences techniques et culturelles. Cette analyse des besoins vise donc à déterminer pour qui et selon quel usage l'œuvre littéraire peut avoir du sens. Elle permet également de savoir comment il faut l'apprêter pour que ces nouveaux publics puissent utiliser les ressources selon leurs attentes ou pour faire naître de nouvelles idées d'utilisation.

La numérisation d'un corpus restreint comme celui qui est évoqué ici se conçoit donc de plus en plus autour de projets de diffusion et de valorisation du patrimoine auprès de publics spécialisés (communauté éducative) ou plus ouverts (tourisme). A l'inverse, des grands fonds bibliographiques reconnus et plébiscités par les institutions scolaire et éditoriale et consacrés par la communauté universitaire, dans le cas des corpus d'intérêt local, le critère scientifique n'est pas suffisant car la numérisation ne peut se plaider en termes d'économie et de financement que dans le cadre d'un projet culturel d'ensemble susceptible de fédérer un grand nombre de partenaires issus des horizons professionnels les plus variés.

1.4. Les conditions de « revitalisation » numérique du corpus :

La recherche dans ce domaine ne se pose donc pas seulement en termes de technique informatique ou de connaissances linguistiques ou littéraires. Elle implique aussi toute une réflexion sur la possibilité de mettre à disposition de nombreux utilisateurs aux objectifs et aux pratiques différents un même document de base, en l'occurrence un corpus littéraire et son iconographie annexe. Un tel travail d'équipement des corpus doit se faire en partenariat étroit entre les divers secteurs concernés : conservateurs et responsables techniques des établissements détenteurs des fonds, spécialistes des contenus linguistiques et littéraires, informaticiens et professionnels du multimédia.

Il comprendra nécessairement une phase d'équipement technique de la documentation essentiellement consacrée au marquage sémantique du corpus ainsi qu'une phase d'appareillage conceptuel plus particulièrement tournée vers les interactions et la conception de scénarios d'utilisation. Cependant, compte tenu de la spécificité du fonds évoqué, il faut d'abord s'arrêter sur la grille de lecture et d'interprétation du corpus qui sert de dénominateur commun aux divers préparateurs chargés de travailler à la préparation des « archives » en vue des usages traditionnels comme nouveaux.

1.4.1. La sémiotique du territoire comme « entrée polyvalente» dans le corpus :

Aujourd'hui, l'opération qui consiste à équiper un tel corpus associe spécialistes des contenus et de l'informatique. Elle se concentre essentiellement sur la mise au point d'appareil de notes, de renvois par des liens ancrés dans le texte vers des documents placés en annexes ou vers d'autres sites ; bref, c'est l'édition critique hypertextuelle destinée à une diffusion et à un usage savants qui constitue le principal horizon de ces nouvelles techniques appliqués au texte littéraire[11]. En effet, en dehors du catalogage ou de la mise à disposition

[11] Avec son projet de mise en ligne de la version électronique de *Madame Bovary*, le Centre Flaubert de l'Université de Rouen constitue un bon exemple de cette valorisation du texte numérisé : http//: www.univ-rouen.fr/flaubert/

d'une version numérique téléchargeable, on assiste déjà aux débuts de la valorisation commerciale de certaines œuvres phares (*Les Caves du Vatican* de A. Gide[12].) mais il s'agit essentiellement de la mise à disposition d'un public professionnel ou spécialisé d'éditions électroniques hypertextuelles à la maniabilité limitée. Au cours des prochaines années, compte tenu de l'élargissement des publics et des utilisations mais aussi du potentiel d'évolution technique que possède l'outil informatique, il sera indispensable de réfléchir à des formes de didactisation des documents beaucoup plus variées.

La préparation informatique du texte (pose de balises, outillage technique -conceptuel et ergonomique- etc…) dès sa numérisation par l'application de protocoles définis en fonction du genre des documents devra, pour des raisons de diffusion prendre en compte l'éventail le plus large possible des publics susceptibles de les utiliser. La didactisation de l'œuvre littéraire ou du document culturel ira donc de pair avec la mise au point d'une scénarisation des données comme des utilisations possibles.

Dans le cas de notre corpus, c'est l'objet « territoire » qui constitue à la fois le contenu sur lequel il s'agit de travailler et la forme qui va contraindre et tout à la fois faciliter la didactisation des documents. Le territoire est un *objet culturel* doté de sens documentés dans des contextes sociohistoriques donnés. Il est constitué d'un grand nombre d'informations qui vont des plus élémentaires comme les repères géographiques aux plus complexes qui pourraient être par exemple les « représentations » successives de ce territoire classées selon leur apparition chronologique. Mais, le territoire est aussi un *objet interactionnel* dans la mesure où sa signification est toujours le résultat d'échanges entre divers interlocuteurs ; en somme, il est un interdiscours permanent, toujours en construction, toujours en évolution, toujours en négociation. Dans le cas d'une exploitation numérique du corpus, ces dimensions donnent toutes deux prise à la didactisation, c'est-à-dire à la préparation de la documentation et de sa transmission à des publics ciblés. Le territoire « objet culturel » donnera lieu a une opération de marquage alors que le territoire en tant « qu'objet interactionnel » sera la matière de scénarios d'interactions entre l'application informatique et les usagers.

1.4.2. L'appareillage technique du corpus :

1.4.2.1. Marquage et éléments d'interprétation :

Cependant, spécialistes des contenus et informaticiens se rejoindront sur une porte d'entrée commune dans le corpus et la même grille de lecture et d'interprétation des documents. Cette dernière est fournie par la sémiotique du territoire ou plutôt par certains éléments empruntés à ce système sémiotique, à savoir les noms de lieux (toponymes, hydronymes, domonymes, etc.) référencés et documentés ainsi que les marques spatiales internes au récit et en relation avec le système énonciatif.

La sémiotique du territoire constitue donc le mode de lecture et d'interprétation commun des documents. En mettant l'accent sur les repères spatio-temporels, elle permet l'élaboration d'index sémantiques qui sont autant de points d'entrée ou d'accroche pour les différents concepteurs et préparateurs de la version informatisée du corpus. Ces repères représentent à la fois la richesse documentaire du corpus considéré et une partie essentielle de son potentiel de dynamisation. Sur le corpus considéré, nous avons donc appliqué une grille de balisage relativement fine et assez largement polyvalente puisqu'elle devrait pouvoir répondre aux besoins de plusieurs types d'utilisateurs pour des usages variés, depuis la modalité informative jusqu'à la recherche d'ordre linguistique.

---

[12] En CD-Rom : Edition génétique des Caves du Vatican d'André Gide, Alain Goulet, Université de Sheffield, Ed. Gallimard, 2000.

Ainsi, la conception du territoire comme *contenu sociohistorique constitué* ou *objet culturel* a tout naturellement trouvé un prolongement didactique dans la démarche d'informatisation. Toutefois, si l'on considère que le territoire est un *objet culturel* complexe, il est bien évident que la documentation relative aux noms de lieux constitue une dimension minimale voire minimaliste de l'ensemble des données que cet objet culturel comprend. Cependant, comme nous le verrons plus avant, ces repères territoriaux font signe dans le texte comme sur les autres types de documents utilisables (cartes géographiques, cartes postales, photographies, et autres documents iconographiques, etc.) ou sur le terrain pour l'ensemble des préparateurs-concepteurs de l'appareillage des documents comme pour tous les utilisateurs potentiels.

1.4.2.2. Les interactions et les divers scénarios d'utilisation des ressources :

Associé au linguiste ou au géographe, dans un premier temps, l'informaticien tirera de ces éléments territoriaux un index qui répertoriera les segments linguistiques simples comme complexes liés à l'expression du lieu. Il les classera en fonction des précisions qui lui seront fournies. Dans un deuxième temps, cette base de données sera susceptible de permettre toutes sortes d'opérations de « manipulation » et d'appropriation des documents par des utilisateurs experts comme plus occasionnels. Le linguiste, le littéraire -et la communauté scientifique dans son ensemble- y verront la possibilité d'en faire un usage savant et exhaustif. L'enseignant en proposera un usage pédagogique à l'intention de ses élèves. Quant aux professionnels de l'information, de l'économie locale ou du tourisme, ils en feront un usage partiel et parfois inattendu –parce que détourné ou inédit- pour apprêter un service à un public restreint (restaurateurs, hôteliers, touristes, etc.) ou en direction du grand public (office de tourisme, site web des grandes villes ou des institutions régionales).

Ainsi, compte tenu de l'ouverture potentielle à de nouveaux publics pour de nouveaux usages, le corpus doit être préparé ou « didactisé » de manière à le rendre préhensible aux plans conceptuels comme ergonomiques dans le cadre de pratiques étendues et renouvelées qui restent aujourd'hui, pour l'essentiel, à créer tant sur le plan conceptuel que technique.

Si dans la phase de didactisation et d'appareillage des documents, les points de rencontre entre les spécialistes des contenus et leurs homologues informaticiens étaient avant tout les toponymes et autres repères spatiaux, d'autres concepts deviennent progressivement des entrées particulièrement fructueuses dès lors qu'il s'agit d'aborder les usages. Ainsi, l'approche du concept de territoire comme *interdiscours* ou *objet interactionnel* ouvre des perspectives intéressantes en termes de scénarisation du corpus notamment quant à la mise au point de scénarios d'utilisation des ressources.

Il est toutefois certain qu'en associant étroitement la connaissance du domaine local et les concepts de la sémiotique du territoire par l'intermédiaire des multiples moyens de repérage et d'accès à l'information que fournissent les technologies informatiques et multimédias, ces nouveaux usages peuvent notablement modifier la relation aux corpus littéraires et culturels. Ils peuvent aussi susciter de nouvelles interprétations de leurs contenus à travers de nouvelles approches méthodologiques et un renouvellement des besoins documentaires.

Conclusion :

On le voit, la numérisation représente un tournant dans le cycle de vie du document. Du reste, il serait plus juste d'employer l'expression « cycle des vies » car c'est bien une nouvelle vie sous une nouvelle forme et selon des modalités d'utilisation et de socialisation élargies et amplement renouvelées qui se profile pour le corpus « oublié » que nous avons évoqué.

On peut penser que la numérisation du document « papier » recèle sur le plan technique un véritable potentiel de revitalisation pour les contenus convertis. Cependant, chaque corpus possède sa spécificité et le projet de numérisation sera lié aux perspectives ultérieures de valorisation de ce patrimoine. Les conditions de sa resocialisation seront nécessairement tributaires des points forts qui se dégagent d'une relecture actualisée et historicisée de l'archive.

Dans le cadre de notre intervention, c'est la problématique de la revitalisation des corpus restreints qui a été envisagée et plus particulièrement celle du patrimoine local ou territorialisé. Nous avons circonscrit un corpus « oublié » et nous l'avons revisité en nous appuyant sur une grille élaborée à partir de la sémiotique du territoire qui a constitué le point d'entrée commun aux chercheurs issus d'horizon et de spécialités aussi différentes que la littérature et l'informatique. En nous appuyant sur la dimension « d'objet culturel » du territoire nous avons obtenu un ensemble de marqueurs territoriaux utiles pour de nombreux usagers. De même, la dimension « interactionnelle » du territoire a permis de proposer l'ébauche de scénarios d'interactions susceptibles de diversifier les usages traditionnels du corpus. Ainsi, nous avons pu envisager trois types d'usages, savant, pédagogique et touristique qui renouvellent l'approche du corpus converti en documentation électronique. C'est donc une indéniable évolution qui se produit à cette occasion puisque les modalités de lecture, d'appropriation et, par voie de conséquence d'interprétation du corpus sont largement modifiées.

Toutefois, les outils de traitement de la documentation territorialisée proposés ici n'ont pour l'instant été expérimentés que sur des textes qui s'inscrivaient au sein d'un territoire et de lieux explicitement référencés dans d'autres types de documents, à savoir des cartes géographiques ou des supports iconographiques (cartes postales, photographies, etc.). Cet ancrage référentiel constitue l'une des conditions de la resocialisation du corpus territorialisé. Il reste à aborder le redoutable problème posé par le texte de fiction. Cependant, les résultats obtenus laissent espérer qu'une transposition des procédures de marquage comme d'exploitation des documents à dimension référentielle est en partie possible à la condition de créer les ressources externes adaptées à l'univers abordé. Il faut notamment évoquer l'élaboration d'un panorama complet du système autoréférentiel élaboré dans le cadre du texte de fiction ou d'un corpus fictionnel consacré à un territoire[13] afin de traiter la dimension spatiale de cet univers imaginaire comme un espace tangible.

Désormais, la numérisation se conçoit de plus en plus autour de projets de diffusion et de valorisation du patrimoine auprès de publics spécialisés (communauté éducative) ou plus ouverts (tourisme). A l'inverse, des grands fonds bibliographiques reconnus et plébiscités par les institutions scolaire et éditoriale et consacrés par la communauté universitaire, dans le cas des corpus d'intérêt local, le critère scientifique n'est peut-être pas suffisant. C'est donc bien un projet culturel d'ensemble dans le cadre plus vaste d'un projet d'aménagement du territoire qu'il faut concevoir pour envisager de façon réaliste l'avenir des fonds numérisés et leur transmission. En effet, la mise en valeur numérique des ressources est liée à la création de lieux d'exposition (musées, écomusées, salles publiques, offices de tourisme, lieux privés, etc) ou de pratique (centres scolaires d'initiation au patrimoine, etc) réels ou virtuels (sites, bornes interactives, etc) et cela demande l'implication croissante des instances et des institutions les plus diverses en tant qu'initiateurs ou partenaires des projets.

La conversion numérique des corpus bibliographiques ou culturels n'est donc pas une procédure fermée sur elle-même. A cette opération technique initiale sont liées la question de la diffusion des fonds numérisés et la problématique complexe de leur communication. Elle ouvre de nouveaux horizons en matière de publics et d'usages et invite les spécialistes des

[13] L'exemple du *Seigneur de anneaux* de J.R. Tolkien vient immédiatement à l'esprit.

disciplines concernées à renouveler leur rapport aux documents qu'ils étudient dans le silence des bibliothèques.

BIBLIOGRAPHIE :

Sciences de l'information et de la communication :

Les technologies de l'intelligence, P. Lévy, Ed. La Découverte, 1990
Introduction à l'analyse de l'image, M. Joly, Ed. Nathan, 1993.
Précis de Sémiotique générale, J.M. Klinkerberg, Ed. De Boeck, 1996.
Le livre en révolution, R. Chartier, Ed. Textuel, 1997.
Banque de données et hypertextes pour l'étude du roman, N. Ferrand, Ed. PUF, 1997.
Introduction aux études littéraires assistées par ordinateur. M. Bernard, Ed. PUF, 1999.
Du papyrus à l'hypertexte, Ch. Vandendorpe, Ed. La Découverte, 1999.
Multimédia : les mutations du texte. Dossier coordonné par Th. Lancien, Ed. ENS, 2000.
Y-a-t-il vraiment des technologies de l'information ? Y. Jeanneret, Ed. Septentrion, 2000.
Introduction à la médiologie, R. Debray, Ed. PUF, 2000.
« Tout garder ? Les dilemmes de la mémoire à l'âge médiatique », E. Hoog, in *Le Débat* (n°125, 2003).

Géographie et anthropologie :

L'Ecriture de l'histoire, M..de Certeau, Ed. Gallimard, 1975.
La Région, espace vécu, A. Frémont, Ed. Flammarion, 1976-1999.
La Géographie culturelle, P. Claval, Ed. Nathan, 1995.
Géographie sociale et territoires, G. Di Méo, Ed. Nathan, 1998.
Anthropologie urbaine, A. Raulin, Ed. Armand Colin, 2001.
Mœurs normandes, F. Zonabend, Ed. Christian Bourgois, 2003.

Sites Web de référence :

Bibliothèque Nationale de France : http://gallica.bnf.fr
Archives Littéraires Suisses : http://www.snl.ch/f
Bibliothèque Nationale du Luxembourg, Centre National de Littérature : http://literaturarchiv.lu.fr
Bibliothèque Royale de Belgique : http://www.aml.cfwb.be
Bibliothèque nationale du Québec : http// : www.bnquebec.ca/
Archives Départementales de Savoie : http://www.sabaudia.org
Bibliothèque Municipale de Lisieux : http://bmlisieux.com
Centre Flaubert de l'Université de Rouen : http//: www.univ-rouen.fr/flaubert/
Projet IRMA : http:/www.unites.uqam.ca/irma/

CD-Rom : Edition génétique des Caves du Vatican d'André Gide, Alain Goulet, Université de Sheffield, Ed. Gallimard, 2000.

.