

# Révélation numérique du patrimoine culturel et immatériel des sociétés savantes. Un exemple dans la Région Grand Est

*Digital revelation of the cultural and intangible heritage of learned societies. An example in the Grand Est region*

**Résumé.** Nous présentons une expérience (en cours) de valorisation de l'ensemble des collections de l'Académie lorraine des Sciences dans la Région Grand Est. Cette société a été créée à Strasbourg avec un déménagement à Nancy en 1970. Ce fonds représente 200 ans de publications scientifiques pluridisciplinaires (et pluriculturelles). Il était dispersé sur plusieurs sites (une partie conséquente aux États-Unis) et sous forme de fichiers en format image. Nous décrivons un premier niveau de valorisation « assez classique » où l'on donne accès aux articles proprement dits dans une vraie bibliothèque numérique. Mais surtout, nous montrerons qu'un traitement ultérieur dans un wiki sémantique permet de révéler, pour cette société, le patrimoine immatériel et sociologique porté par ces publications. De plus nous disposons d'un réseau de wikis sémantiques spécialisés (zoologie, botanique, informatique, etc) qui permet de révéler le patrimoine scientifique ou culturel porté par des ensembles d'articles sur un thème donné. Nous discuterons des enseignements de cette action sur un plan individuel (par exemple, acquisition de connaissances scientifiques en réalisant leur modélisation sémantiques). Nous évoquerons également les retombées institutionnelles potentielles de cette action pour un grand projet numérique adapté aux algorithmes de l'intelligence artificielle.

**Mots-clés.** Société savante, bibliothèque numérique, hypertexte.

**Abstract.** This paper describes an experiment (in progress) of valorization of the entire collections of the Lorraine Academy of Sciences in the French Grand Est region. This company was created in Strasbourg with a move to Nancy in 1970. This corpus represents 200 years of multidisciplinary (and multicultural) scientific publications. It was dispersed over several sites (a significant part in the United States) and in the form of image files. We will describe a first level of "fairly classic" valorization where we give access to the articles themselves in a real digital library. But above all, we will show that a subsequent treatment in a semantic wiki makes it possible to reveal, for this company, the intangible and sociological heritage carried by these publications. In addition, we have a network of specialized semantic wikis (zoology, botany, computer science, etc.) which makes it possible to reveal the scientific or cultural heritage carried by sets of articles on a given theme. We will discuss the lessons learned from this action on an individual level (for example, the acquisition of scientific knowledge by carrying out their semantic modeling). We will also discuss the potential institutional repercussions of this action for a large digital project adapted to artificial intelligence algorithms..

**Keywords.** Learned society, digital library, hypertext.

## 1. Introduction

Cet article présente l'intégration des publications d'une société savante dans un wiki sémantique qui est intégré dans un réseau de sites scientifiques ou culturels.

Cette société savante est l'Académie lorraine des sciences (ALS), actuellement basée à Nancy mais qui a été créée à Strasbourg en 1828. Elle a été déménagée à la suite de la guerre de 1870.

Nous présenterons, dans un premier temps, le contexte culturel et historique de cette action au sein d'une région qui a subi de nombreuses mutations au cours du temps. Nous présenterons également cette société.

Les publications de l'ALS, pour la période nancéienne sont disponibles sous forme de fichiers PDF, copie des bulletins trimestriels. Au départ de cette action, nous avons repéré les mémoires de la période strasbourgeoise numérisés dans des institutions américaines. Nous avons alors entrepris de réunir ces collections patrimoniales dans une vraie bibliothèque numérique (avec un accès par sommaires aux articles).

Les articles les plus intéressants sont alors enrichis avec des relations hypertextes et sémantiques. Le lecteur peut alors explorer le corpus et se faire une idée du patrimoine ainsi révélé. Pour les articles les plus pertinents sur un plan scientifique nous appliquons ce mécanisme au sein d'espaces numériques spécialisés (Botanique, Histoire de la Lorraine, histoire naturelle des animaux).

Dans une dernière partie, nous parlerons des retombées de ces pratiques dans l'acquisition des connaissances et nous proposerons des réflexions d'un caractère institutionnel.

## 2. Les sociétés savantes du Grand Est

Le cadre général de nos actions est la valorisation du patrimoine historique ou culturel contenu dans les publications scientifiques de la Région Grand Est. Voici quelques repères historiques pour ce corpus potentiel :

- 1450, Création de l'imprimerie européenne par Gutenberg ;
- 1572, création de l'université de Pont-à-Mousson (ancêtre de l'université de Lorraine) ;
- 1665, création de la première revue scientifique, le *Journal des sçavants* ;
- 1666, création de l'Académie des sciences ;
- 1750, création de la première société savante en lorraine : l'Académie de Stanislas.

Plus modestement, dans ce projet, nous voulons traiter les publications de l'Académie Lorraine des sciences<sup>1</sup>. Son origine remonte à Strasbourg en 1828. Elle était portée le muséum d'histoire naturelle de Strasbourg (créé en 1804). Une grande partie des académiciens étaient professeurs à la faculté de médecine de Strasbourg. Celle-ci a des racines très anciennes avec Jean Sturm en 1538. Strasbourg était alors une ville libre du Saint-Empire germanique, qui a fondé une première université en

---

<sup>1</sup> <https://als.univ-lorraine.fr/>

1621. Celle-ci a été supprimée par la Convention en 1792. Napoléon Bonaparte a créé une des trois facultés de médecine (avec Paris et Montpellier) en 1808.

Cette académie s'est rapidement diversifiée pour intégrer la géologie, puis la chimie, etc.

En 1870, éclate la guerre entre l'Allemagne et la France. Strasbourg devient une ville allemande et un grand nombre de professeurs rejoignent Nancy. Les membres de la Société d'Histoire naturelle de Strasbourg modifient les statuts pour créer la Société des sciences de Nancy. Elle deviendra Académie lorraine des sciences en 1965 compte tenu de l'élargissement de ses activités.

D'un point de vue scientifique, l'ALS est maintenant organisée en 5 sections :

- Section 1 : Mathématiques, Physique, Chimie, Electronique, Informatique, Génie des Procédés.
- Section 2 : Biologie animale et végétale, Sciences de l'environnement.
- Section 3 : Médecine, Médecine vétérinaire, Pharmacie.
- Section 4 : Sciences du sol, de la terre et de l'univers.
- Section 5 : Sciences humaines.

Notre objectif est de faire ressortir le gigantesque patrimoine porté par les contributions des chercheurs et praticiens du Grand Est, pour 200 ans d'une histoire qui, en fait, a commencé il y a 500 ans.

### **3. Constitution d'un corpus numérique manipulable**

De façon prosaïque, la première étape a consisté à ressembler et homogénéiser les publications.

La partie la plus importante du corpus est constituée d'un ensemble de fichiers PDF, copies des bulletins produits par l'association depuis 1973 (à Nancy). Les ouvrages papiers ont été numérisés par l'INIST en 2005 pour créer un portail d'accès aux articles (par un service dénommé eRevues). Malheureusement l'INIST a démantelé ce service en 2014. Le portail a alors été supprimé et l'INIST a remis à l'ALS une simple copie des ouvrages papiers sous forme de fichiers PDF.

Cet ensemble est très hétérogène. Il est constitué de fichiers qui vont des quelques dizaines de pages à plusieurs centaines. La longueur des articles varie d'une simple page (procès-verbaux de séance) à des mémoires de plus de 150 pages.

En démarrant cette action, nous avons découvert que les mémoires de la société strasbourgeoise n'étaient pas disponibles en France. En effet, ils avaient été perdus dans le déménagement ! Heureusement, dans le cadre d'échanges entre sociétés savantes, des exemplaires étaient rangées dans des bibliothèques américaines qui les ont numérisées. Elles sont disponibles sur le site *Biodiversity Heritage Library* (BHL<sup>2</sup>), géré, à Washington par les « *Smithsonian Libraries and Archive* ». Compte-tenu des changements politiques aux États-Unis, il nous a paru urgent de rapatrier ce précieux patrimoine. Il est constitué de 6 volumes de 350 à 700 pages que l'on peut télécharger, page par page, en format jpeg. Nous avons également

---

<sup>2</sup> <https://www.biodiversitylibrary.org/>

trouvé quelques articles en PDF sur le portail numérique du Jardin botanique royal de Madrid.

Nous venons de rapatrier tous ces fichiers dans un site wiki (moteur Semantic MediaWiki). Nous avons commencé à créer un ensemble d'articles. Au départ chaque article est une « page MediaWiki » qui contient une en-tête bibliographique, suivie d'un ensemble de copies de pages PDG ou JPEG suivant le cas. Dans la description bibliographique les auteurs sont étiquetés avec un attribut « A pour auteur : : ». Ceci permet de générer par exemple des listes d'articles par auteur. De même, les articles peuvent être indexés par des catégories.

Cette opération est assez longue pour les PDF issus du site ALS. Dans une phase intermédiaire, nous constituons une liste des sommaires (à partir des tables des matières des fascicules PDF). Cette opération est souvent fastidieuse car les PDF ne sont pas toujours de qualité suffisante pour un copier/coller (ou pour une bonne reconnaissance OCR). Par chance, nous avons trouvé sur le site Internet Archive des copies du site de l'ALS. Nous y avons trouvé les sommaires de qualité qui avaient été réalisés par l'INIST. Là encore, une partie de notre patrimoine numérique perdu a pu être rapatrié... à partir des États-Unis.

Nous disposons donc d'un premier niveau de bibliothèque qui permet d'accéder aux articles par sommaire, index auteur etc. Mais ces articles sont simplement juxtaposés en mode image. L'étape suivante consiste à les transcrire en OCR pour en faire des articles en format texte, puis hypertexte.

Cette opération a démarré. Nous pouvons déjà présenter trois types de valorisation hypertexte pour faire émerger le patrimoine porté par ces publications : au niveau des articles, puis de la collection et enfin de l'insertion dans un espace scientifique spécialisé.

#### 4. Où un article devient déjà un hypertexte

Les articles sont, nous l'avons dit, de taille très variable. Un compte-rendu d'une ou deux pages sans intérêt majeur peut rester en mode image.<sup>3</sup>

Pour une publication classique (5 à 15 pages), la conversion en *wikitexte* (pour reprendre la terminologie Wikipédia) donne déjà de bons résultats avec un balisage de base (italique, gras), éventuellement enrichi par des formules en LaTeX. Le repérage des titres permet de générer automatiquement une table des contenus en début d'article.

En revanche, un article de plus de 20 pages est déjà rébarbatif. Au delà de 50 pages il devient pratiquement illisible. Il faut alors le distribuer, déjà en mode image, sur plusieurs « pages wiki », la première page wiki d'accueil donnant alors le sommaire. On peut réaliser des améliorations progressives avec par exemple : l'introduction transcrite, suivie du sommaire pointant les pages des chapitre, en enfin la conclusion.

---

<sup>3</sup> [https://wicri-demo.istex.fr/Wicri/Europe/France/GrandEst/SocGrandEst/fr/index.php/Bull.\\_Soc.\\_sci.\\_Nancy/1898\\_33/CR/1er\\_f%C3%A9vrier](https://wicri-demo.istex.fr/Wicri/Europe/France/GrandEst/SocGrandEst/fr/index.php/Bull._Soc._sci._Nancy/1898_33/CR/1er_f%C3%A9vrier)

Mais certains articles, de taille conséquente, contiennent parfois des petits dictionnaires. Ainsi, nous disposons d'un article de Philippe Voltz (daté de 1830) qui décrit la collection des bélemnites du Muséum de Strasbourg sur 72 pages (plus huit pages de planches)<sup>4</sup>. Il contient 17 descriptions d'espèces de ce fossile. Nous avons créé autant de pages wiki que d'espèces. Chacune est alors décrite sur une page-wiki où l'on a rapproché la (ou les) planches de figures relatives à l'article. Chaque fossile ainsi décrit est souvent plusieurs fois cité dans le texte avec des liens hypertexte. Le lecteur put alors « explorer la collection ».

Un autre exemple est donné par un répertoire archéologique de la Moselle<sup>5</sup>. Il s'agit d'un « article » de 250 pages qui contient plusieurs centaines de description de communes réparties dans une cinquantaines de cantons, eux mêmes répartis dans 8 arrondissements. Nous avons créé une page par arrondissement, qui donne la liste des communes par canton. Il y a ensuite une page par canton avec des liens internes vers des communes. Certaines communes, qui contiennent plusieurs descriptions, disposent de leur propre page.

Un article d'un document peut donc devenir un hypertexte qui permet de naviguer dans des espaces patrimoniaux (les fossiles du Muséum de Strasbourg ou les descriptions archéologiques des communes de Moselle).

## **5. Un exemple au niveau de la collection : l'histoire de la société et de ses membres**

Il est également possible de concevoir un hypertexte sur l'ensemble de la collection. Voici un exemple avec la vie de la société et de ses membres.

Parfois, quelques académiciens donnent un historique ciblé de la vie de la société. Par exemple, Gilbert Percebois, professeur de parasitologie et mycologie, a écrit un article de synthèse sur « La biologie dans les publications de la Société Lorraine des Sciences (1828-1978) »<sup>6</sup>. Il y cite notamment des sociétaires et leurs articles les plus intéressants (et les sessions où ils ont été discuté). Un traitement hypertexte permet au lecteur d'accéder « un un clic » à ces éléments.

Un tel article est également très précieux pour donner au rééditeur un ordre de priorité pour traiter les milliers d'articles et les centaines d'auteurs de l'académie.

Un point important à ce niveau est la désambiguïisation des patronymes. Cet exercice est parfois difficile car il n'était pas d'usage de mentionner les prénoms. Par exemple, pour le patronyme « Duvernoy », célèbre zoologiste, Wikipédia donne 10 entrées. Il est alors facile d'identifier Louis Georges Duvernoy, notre ancien président. Nous donnons alors comme identifiant celui de Wikipédia (et donc de WikiData). C'est nettement plus compliqué avec un patronyme tel que : Gross, Maire ou Nicolas.

---

<sup>4</sup> [https://wicri-demo.istex.fr/Wicri/Europe/France/GrandEst/SocGrandEst/fr/index.php/Collection\\_ALS/M%C3%A9moires/Tome\\_1/L\\_1/B%C3%A9lemnites\\_\(Voltz\)](https://wicri-demo.istex.fr/Wicri/Europe/France/GrandEst/SocGrandEst/fr/index.php/Collection_ALS/M%C3%A9moires/Tome_1/L_1/B%C3%A9lemnites_(Voltz))

<sup>5</sup> [https://wicri-demo.istex.fr/Wicri/Europe/France/GrandEst/SocGrandEst/fr/index.php/Collection\\_ALS/S%C3%A9rie\\_6/Tome\\_9/N\\_1/Toussaint](https://wicri-demo.istex.fr/Wicri/Europe/France/GrandEst/SocGrandEst/fr/index.php/Collection_ALS/S%C3%A9rie_6/Tome_9/N_1/Toussaint)

<sup>6</sup> [https://wicri-demo.istex.fr/Wicri/Europe/France/GrandEst/SocGrandEst/fr/index.php/Collection\\_ALS/S%C3%A9rie\\_7/Tome\\_19/N\\_2/Percebois\\_biologie](https://wicri-demo.istex.fr/Wicri/Europe/France/GrandEst/SocGrandEst/fr/index.php/Collection_ALS/S%C3%A9rie_7/Tome_19/N_2/Percebois_biologie)

Pour des membres peu connus (et qui n'ont pas de pages sur Internet), cela demande une stratégie de recherche parfois infructueuse. Nous avons par exemple un « x' Thouvenin ».

Il est donc possible de naviguer à partir d'articles de synthèses ou sur les listes d'auteurs. Chaque auteur possède sa fiche où il est possible de citer les laboratoires où il a exercé, sa ville de naissance etc.

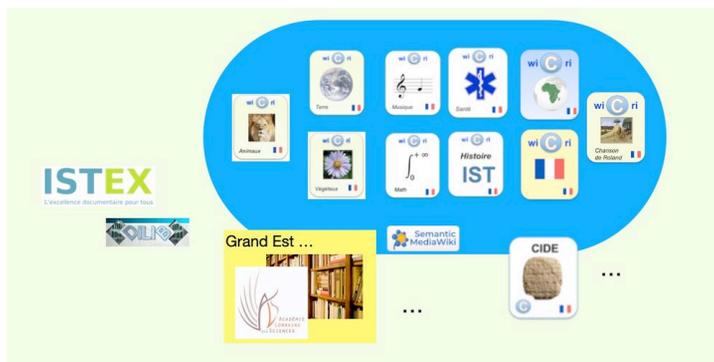
La collection des publications de l'ALS permet donc de faire émerger un ensemble de faits porteurs du patrimoine que représente la vie de la société.

Enfin, les collections de l'ALS sont gérées dans un « portail wiki » où elles vont cohabiter, sur le même espace numérique, avec d'autres sociétés du Grand Est. Au niveau du wiki, nous avons donc la possibilité de constituer un vaste réseau de connaissance sur l'histoire de la science en Lorraine (et dans la région Grand Est).

Il est possible d'y faire émerger les coopération avec d'autres régions en traitant les membres correspondants de ces sociétés.

## 6. Valorisation d'un article dans un espace spécialisé

Comme le montre la figure suivante, le portail de l'ALS fait partie du réseau Wicri, un ensemble de wikis.



Ces wikis sont souvent spécialisés par domaine scientifique (Animaux, végétaux, musique, santé etc). Un mécanisme nommé « lien interwiki » permet de réaliser des liens hypertextes entre les wikis.

Une des utilisations les plus simples de ce dispositif consiste à expliquer une notion citée dans le wiki des sociétés savantes avec un développement dans un autre wiki (sans « encombrer » le wiki de départ). Par exemple, un article de l'ALS traite des mammifères et de l'ornithologie à Madagascar<sup>7</sup>. L'auteur, Victor Sganzin, y cite une liste d'animaux très spécifiques comme « *le petit maki gris de Buffon* ». Ce sujet peut alors être développé sur le wiki « animaux ».

<sup>7</sup> [https://wicri-demo.istex.fr/Wicri/Europe/France/GrandEst/SocGrandEst/fr/index.php/Collection\\_ALS/M%C3%A9moires/Tome\\_3/L\\_1/Sganzin\\_Madagascar](https://wicri-demo.istex.fr/Wicri/Europe/France/GrandEst/SocGrandEst/fr/index.php/Collection_ALS/M%C3%A9moires/Tome_3/L_1/Sganzin_Madagascar)

Plus précisément, le wiki « animaux » contient une réédition de l'histoire naturelle de Buffon en cours de réalisation. Le ré-éditeur de l'article sur Sganzin (qui a aussi le droit de contribuer sur Wicri/Animaux). peut réaliser la réédition numérique de l'article de l'encyclopédie de Buffon.

L'Histoire naturelle de Buffon est un monument éditorial d'une cinquantaine de tomes de quelques centaines de pages chacun<sup>8</sup>. Le réseau de wiki associant des sites régionaux et d'autres plus spécialisés favorise les contributions collectives.

Un des mémoires strasbourgeois contient une monographie de 150 pages sur la girafe rédigée en 1858 par Nicolas Joly<sup>9</sup>. Elle offre par exemple une liste de plus d'une cinquantaine d'auteurs qui ont écrit sur la girafe, depuis Moïse dans le Deutéronome, jusque Cuvier ou Geoffroy Saint-Hilaire, en passant par Aristote ou Marco-Polo.

Ce type de traitement s'avère être d'une très grande richesse pour structurer une éditions hypertexte de l'Histoire de l'Histoire naturelle. De façon surprenante, il nous a aidé à identifier les références bibliographiques données par Buffon. Par exemple quand Buffon écrit « *Gefner, "hist. quad. pag. 147"* » Joly précise : « C. Gesner. "*Conradi Gesneri historia animalium. De quadrupedibus*", p. 147-150. ».

Un autre article, daté de 1912, du Révérent Père Pionnier, missionnaire mariste, donne la liste des lichens de Nouvelle Calédonie. Il est naturellement très intéressant à exploiter du point de vue de la biodiversité. Il se pose alors un problème d'identification, la nomenclature de ces végétaux ayant évolué.

Là encore, l'architecture d'un réseau de wiki sémantique permet de décrire cette évolution de façon incrémentale et collaborative.

## **7. Retombées et perspectives**

Le réseau de wikis s'avère donc être un espace numérique sophistiqué pour exprimer la connaissance contenue dans un corpus bibliographique. Il permet de révéler un patrimoine intellectuel, la connaissance scientifique et culturelle sur plusieurs siècles.

La technologie utilisée est extrêmement souple pour gérer un mode d'action où l'éditeur est confronté à des difficultés inattendues.

Cet exercice présente deux types de retombées, au niveau intellectuel au niveau individuel d'une part et collectif d'autre part.

### **7.1. Au niveau individuel**

Rédacteur de cet article, je suis le ré-éditeur des collections de l'ALS. Deux points me semblent intéressants à mettre en évidence.

Depuis le début de l'aventure Wicri, j'ai souvent traité des sujets en fonction de mes préférences ou de mes connaissances. Ici avec un corpus totalement interdisciplinaire, j'ai été amené à aborder un grand nombre de terrains qui m'étaient inconnus. J'avoue avoir été parfois surpris par la rapidité avec laquelle on peut acquérir une première expertise sur un domaine donné. Ce type d'exercice me paraît donc très intéressant à exploiter dans une stratégie d'enseignement.

---

<sup>8</sup> [https://wicri-demo.istex.fr/Wicri/Biologie/Animaux/fr/index.php/Histoire\\_naturelle\\_\(Buffon\)](https://wicri-demo.istex.fr/Wicri/Biologie/Animaux/fr/index.php/Histoire_naturelle_(Buffon))

<sup>9</sup> [https://wicri-demo.istex.fr/Wicri/Biologie/Animaux/fr/index.php/Collection\\_ALS/1846/Joly\\_Girafe](https://wicri-demo.istex.fr/Wicri/Biologie/Animaux/fr/index.php/Collection_ALS/1846/Joly_Girafe)

Mais plus encore, j'ai trouvé cet exercice absolument passionnant !

## **7.2. Au niveau collectif**

Wikipédia est alimenté par des volontaires, rappelons-le. Ils contribuent parce qu'ils trouvent cet exercice absolument passionnant.

Dans les années 70 avec le Trésor de la langue française, puis jusqu'en 2000 avec les bases Pascal et Francis, la France jouissait d'une forte présence dans l'espace numérique de la connaissance.

Aujourd'hui, notre plus grand dictionnaire, notre plus grande encyclopédie, notre plus grande bibliothèque d'ouvrage en format texte sont gérées et administrées à San Francisco, par la Wikimedia Foundation.

L'anonymat des contributions avec des mécanismes d'IA générative est une menace très importante pour l'intégrité de... notre patrimoine numérique.

Le projet Wicri et cette expérience de valorisation permettent de donner des éléments de chiffrage pour notre indépendance relatif à notre patrimoine numérique. Pour des raisons diverses, j'ai mené seul les travaux dur l'ALS depuis le début de cette année. J'ai créé environ 2000 « pages wiki » en 9 mois, soit environ 1/1000 de Wikipédia depuis 20 ans. Autrement dit, avec les technologies « wiki sémantique » un ordre de grandeur de quelques (2 ou 3) centaines de personnes à temps plein peuvent rivaliser avec la galaxie Wikipédia.

Le ministère de la recherche affiche 91.000 enseignants chercheurs et 14.000 doctorants par an. Nous disposons donc d'un immense réservoir potentiel de contributions volontaires et qualifiées.

Bien entendu, cette approche peut s'intégrer dans un ensemble de coopérations francophones ou européennes.

Un vaste numérique vertueux est donc possible pour préserver et diffuser notre patrimoine humaniste.