

Patrimoine scientifique : droits d'auteur et de lecteur.

Laurent GUILLOPÉ¹

Résumé : Des programmes ambitieux de numérisation rétrospective de fonds scientifiques anciens ont débuté récemment, avec diverses approches pour les rapports juridiques entre les acteurs concernés. Les choix adoptés dans le domaine des mathématiques, en particulier par le programme français NUMDAM, sont discutés.

1. Communication scientifique et archives numériques

Un accès aisé à la littérature spécialisée, tel est le souhait du scientifique : en tant que lecteur pour alimenter son travail de création, en tant qu'auteur pour diffuser ses travaux. Cette volonté est dans le droit fil de l'une des premières lois sur le copyright, le Statute of Anne édicté par le parlement anglais en 1709 qui est désigné comme *An act for the encouragement of learning*. L'histoire de la propriété intellectuelle indique combien l'auteur a été invoqué plus d'une fois pour mieux établir des monopoles : les tensions actuelles ne sont pas nouvelles. Comment le droit de la propriété intellectuelle et ses multiples facettes (droit d'auteur, droit moral, droit patrimonial, copyright, droit public à l'information,...) permet-il des rapports équilibrés dans la communication scientifique ? Établit-il un juste équilibre entre tous les acteurs concernés (scientifiques, éditeurs, bibliothécaires) ? S'adapte-t-il aux évolutions technologiques et aux changements de mode de travail ?

Un bon exemple où se jouent ces interrogations est donné par la *bibliothèque mathématique numérique mondiale* [WDML]². Initié au début des années 2000, ce projet vise à numériser les fonds anciens mathématiques afin de construire une bibliothèque distribuée sur le réseau, obéissant à des normes d'interopérabilité précises et tressée de liens hypertextes permettant une navigation facile. Plus que pour beaucoup d'autres sciences, la littérature mathématique est pérenne (un résultat mathématique garde sa validité à travers les âges) et le recours à des textes anciens est fréquent : c'est dire tout l'intérêt scientifique de cette entreprise et l'enthousiasme exprimé par la communauté mathématique. L'ensemble de la littérature imprimée mathématique est estimé à 50 millions de pages. Les coûts de constitution de collections numériques, tels qu'ils

¹Laboratoire de mathématiques Jean Leray, Université de Nantes, CNRS & École centrale de Nantes.

²WDML : world digital mathematics library, www.wdml.org.

ont été constatés dans les opérations en cours, attestent du caractère raisonnable de l'ambition de la WDML : ce sont les problèmes juridiques liés aux droits sur les documents qui représentent les plus grandes difficultés.

Les standards techniques sont désormais bien établis, décrits précisément dans un ensemble de *bonne pratiques* adoptées par l'Union mathématique internationale : description bibliographique des documents par des métadonnées normalisées et librement disponibles, numérisation en mode image à haute résolution des documents permettant lecture à l'écran et reproduction à l'identique de l'original sur papier, restitution du texte par des techniques de reconnaissance de caractères donnant la possibilité de recherche textuelles, apposition de liens entre les documents et les bases de données bibliographiques, entre les références bibliographiques du document et les unités documentaires en rapport dans la bibliothèque numérique (recension, articles, commentaires). Les réalisations, suivant partiellement ou totalement ces recommandations, témoignent de leur cohérence, de leur qualité et de leur coût raisonnable : en France (à Grenoble avec le programme NUMDAM), en Allemagne (à la bibliothèque de Göttingen), aux États-Unis (à l'Université Cornell), en Pologne (à l'Université de Varsovie), ... D'autres bibliothèques numériques incluant des fonds mathématiques significatifs existent :

- la bibliothèque Gallica de la bibliothèque nationale de France, avec quelques journaux et six cents monographies,
- l'archive payante JSTOR, reprenant entre autres une vingtaine de revues de mathématiques,
- des archives constituées par les éditeurs, par exemple les groupes Elsevier et Springer.

Dans le domaine des sciences humaines et sociales, où les initiatives se multiplient tout autant, le portail PERSEE propose l'accès public et gratuit à plusieurs séries numérisées³ : ce programme est une initiative du Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche et est basé sur des choix techniques, économiques et juridiques analogues à ceux à la base du programme NUMDAM, avec la même ambition de soutenir l'édition françaises et donner un outil de qualité à la recherche contemporaine, en France et à l'étranger.

2. Le programme NUMDAM

³La numérisation rétrospective de sept revues est en cours : Annales (Éditions de l'ÉHÉSS), L'Homme (Éditions de l'ÉHÉSS), Revue économique (Presses de sciences politiques), Revue française de science politique (Presses de sciences politiques), Bibliothèque de l'École des chartes (Société de l'École des chartes), Revue de l'art (CNRS-périodiques), Matériaux pour l'histoire de notre temps (Association des amis de la BDIC et du Musée). Inauguré en janvier 2005, le portail www.persee.fr donne actuellement accès à une vingtaine d'années de publication environ pour chaque revue.

Conçu en 2000, le programme NUMDAM (pour *NUMérisation de Documents Anciens Mathématiques*) a numérisé et mis en ligne à ce jour une douzaine de revues et de séminaires (soit 300 000 pages et 10 000 articles) édités scientifiquement en France : tous ces titres sont propriétés d'institutions publiques, d'associations ou de fondations, même si deux d'entre eux sont édités et diffusés jusqu'à maintenant par des grands groupes d'édition (Elsevier et Springer). Il est piloté par la Cellule MathDoc, une unité mixte de service du CNRS et l'Université J. Fourier localisée à Grenoble et constituée par une équipe permanente de deux mathématiciens à temps partiel, cinq ingénieurs spécialisés en informatique documentaire et bases de données, une gestionnaire et quelques vacataires : créée en 1995, la Cellule MathDoc a pour mission de soutenir des actions documentaires au service de la recherche mathématique française, par des projets avec le Réseau national des bibliothèques de mathématiques⁴ et plus largement à travers des collaborations européennes et internationales. Le programme NUMDAM est financé par des fonds publics (Ministère de la recherche et CNRS principalement) et représente un important soutien à l'édition française contemporaine en mathématiques.

Titre	Création	Prop.	Pages	Articles
Ann. Gergonne	1810-1831		000 8	935
Ann. École Normale Sup.	1864	ÉNS*	989 68	867 1
Bull. Soc. Math. France	1872	SMF	178 45	2 598
Ann. Scient. Faculté Toulouse	1887	Univ.	052 36	035 1
Ann. Institut Fourier	1949	AIF	054 51	798 1
Publ. Scient. IHÉS	1959	IHÉS	424 17	344
Sém. Proba. Strasbourg	1967	IRMA*	352 17	254 1
Journées É. D. P.	1974	CNRS	471 5	469

Extrait du catalogue des fonds numérisés par le programme NUMDAM. Dans la colonne *Propriétaire du titre*, les astérisques indiquent un contrat d'édition et de diffusion avec un éditeur privé.

⁴Le RNBM, groupement de service du CNRS, regroupe la quarantaine de bibliothèques de recherche en mathématiques en France. Elles sont hébergées pour la plupart dans des laboratoires universitaires, du CNRS ou de l'INRIA

Le programme NUMDAM repose sur quelques motivations et convictions politiques fortes :

- la littérature mathématique ne connaît pas l’obsolescence et son ensemble représente un outil essentiel dans le développement contemporain de la recherche, en mathématiques et pour d’autres sciences ;
- au moment où les concentrations se multiplient dans le monde de l’édition avec des effets monopolistiques qui mettent en péril les équilibres du système de communication scientifique, le soutien à l’édition académique est primordial ;
- une revue aujourd’hui doit être présente sur le web, archives comprises ; la connexion de l’ensemble de la publication par des liens hypertextuels provenant ou menant vers des bases bibliographiques ou de corpus numériques d’autres revues augmente son potentiel scientifique et son rayonnement ;
- les différents acteurs doivent être respectés : les revues par un dialogue contractuel approfondi, les auteurs en leur demandant leur accord pour toute représentation et diffusion de leur travail, les lecteurs en leur offrant un outil de qualité accessible à coût raisonnable ;
- le respect de standards internationaux (xml, unicode) et l’adoption de formats normalisés ouverts (pdf) permettent l’interopérabilité avec des projets similaires, assurent la pérennité de l’investissement et préparent aux extensions et migrations de format imposées par l’innovation technologique.

Concernant les modalités juridiques, les principes suivants ont été retenus :

- la version numérique reste sous le contrôle du propriétaire du titre ;
- chaque auteur ou ayant droit cède ses droits (exclusifs) de reproduction électronique (présentation, migration des formats de conservation, diffusion) au profit du propriétaire du titre ;
- un contrat, non nécessairement exclusif, est établi entre le programme NUMDAM et le propriétaire du titre pour installer la collection numérique sur le serveur www.numdam.org, serveur qui offre un environnement de qualité excellente pour l’affichage de la revue ;
- chaque document délivré aux lecteurs indique clairement le propriétaire du titre et les conditions de *bon usage* ;
- une *barrière mobile*, de 5 ans en général, limite l’accès aux abonnés pour les fascicules de l’édition courante, l’accès aux archives est sans restriction, contrepartie au soutien public ayant réalisé la numérisation et veillant à sa bonne diffusion.

Ces règles ont été appliquées à toutes les revues mis en ligne jusqu'à maintenant, qu'elles soient éditées par des sociétés savantes, des organismes publics, des fondations ou des grands éditeurs commerciaux. Cela a été obtenu parfois à la suite de négociations difficiles entre les éditeurs scientifiques et l'éditeur/diffuseur privé. Par ailleurs, après un travail important de recherches d'adresse, les auteurs des articles publiés entre 1980 et 2000 ont été contactés : des 2 200 courriers envoyés, il y a eu 30% de réponses en retour, toutes approuvant la cession des droits de représentation numérique et diffusion au propriétaire de la revue à l'exception d'une en attente⁵. La légitimité de la démarche pragmatique du traitement des problèmes juridiques se trouve fortement établie par l'ensemble de ces réponses : elle vaut probablement pareillement pour le programme PERSEE. La transposition de cette démarche pour des revues éditées hors de France et rétronumérisées dans le programme NUMDAM, telles *Compositio mathematicae*⁶ en Hollande ou les *Annali della scuola normale superiore di Pisa* en Italie, est en cours d'étude : le respect équilibré des intérêts des auteurs et des propriétaires du titre sera recherché, de manière compatible avec les règles juridiques en vigueur.

3. Droit d'auteur et héritage culturel

La Bibliothèque nationale de France s'en tient à l'interprétation stricte de la règle des 70 ans de la convention de Berne pour les revues : ainsi, par exemple, le *Journal de mathématiques pures et appliquées* (aujourd'hui propriété de Elsevier-France, après son rachat de l'éditeur Gauthier-Villars en 1996) et les *Comptes-rendus des séances de l'académie des sciences* (édité et diffusé par Elsevier-France dans le cadre d'un contrat de délégation de service public avec l'Académie) ne sont consultables dans Gallica que jusqu'en 1934, au-delà de la barrière mobile des 70 ans. Tant le *Journal de mathématiques pures et appliquées* que les *Comptes-rendus* sont l'archétype de la publication patrimoniale, constituant significatif de l'héritage culturel dont l'accès doit rester raisonnablement ouvert. Alors que ces collections ont été numérisées jusqu'en 1960, la règle des 70 ans impose une restriction majeure à la diffusion de ces fonds patrimoniaux et à l'accès à un corpus d'un grand intérêt scientifique.

Les investissements substantiels des éditeurs privés (par ex. les groupes Elsevier ou Springer)

⁵Cette position d'attente découle d'une attitude de défiance envers l'éditeur commercial actuel de la revue. Quelques auteurs ont fait part de leur souci qu'une telle cession n'empêche pas une réédition de l'article, tel que ou sous forme légèrement reprise.

⁶Fondé en 1934 par L. Brower, *Compositio mathematicae* est publié en coopération avec la London mathematical society depuis 2004, les éditeurs scientifiques ayant quitté l'éditeur Kluwer en raison des tarifs d'abonnement trop élevés.

dans la numérisation de leurs fonds anciens ont établi des archives d'accès (fortement) payant : ces éditeurs, entreprises mondialisées, se sont affranchis de la règle des 70 ans, profitant d'interprétations de systèmes juridiques divers ou les ignorant. Seules les publications des années récentes sont accompagnées de contrats où la diffusion numérique est explicitement prévue. Pragmatique, le programme NUMDAM s'est efforcé de tenir compte des règles juridiques : droit moral, droit de représentation, droit de reproduction, droit de diffusion. Sa consultation des auteurs, si incomplète soit-elle, affirme leurs droits et, en cas de contestation a posteriori d'un auteur, l'argument de bonne foi sera plaidé, s'appuyant sur la transparence des attendus du programme et de ses réalisations. Il ne peut en être de même pour les programmes de numérisation à l'initiative d'éditeurs, lancés dans une optique lucrative. D'autres programmes suivent un modèle économique à mi-chemin entre le service public et un modèle dicté par des stratégies purement financières : sans pouvoir profiter de la légitimité de la pure communication scientifique, la résolution du statut juridique y est moins aisée.

La prétention d'un éditeur à s'arroger l'exclusivité des fonds anciens s'oppose à la mission patrimoniale des bibliothèques : à ces dernières le devoir, et le droit, de prendre le relais afin de gérer la maintenance et la diffusion des archives numériques à conditions raisonnables. La coexistence des archives dans des systèmes de service public ou des organismes privés, à but lucratif ou non, n'est pas à exclure a priori : le scientifique ne pourra que gagner dans cette multiplicité de l'offre, avec des mises en valeur différentes. Le coût, direct ou indirect, doit néanmoins rester raisonnable.

Nul doute qu'un éditeur, pour exercer son métier et ses talents, doive avoir des assurances pour rentabiliser ses investissements. C'est le rôle de la barrière mobile : la communauté mathématique juge qu'une période cinq ans est adaptée pour sa discipline, alors qu'en sciences de la vie les chercheurs préfèrent six mois. Le choix de la durée de cette barrière doit être fait afin de ne pas menacer l'équilibre financier de la revue par l'arrêt d'abonnements. Par ailleurs, la qualité de l'affichage des archives sur le réseau est une publicité efficace pour l'acquisition du volume imprimé, qui a un usage différent de celui offert par la navigation sur écran. L'exemple des 1 800 pages des *Éléments de géométrie algébrique* d'A. Grothendieck publiés aux Publications scientifiques de l'IHÉS est caractéristique : la version en ligne sur NUMDAM permet des recherches, alors que les huit volumes de la série imprimée sont acquis en permanence par les bibliothèques. Le premier volume du monumental *Séminaire de géométrie algébrique*, édité chez Springer dans les années 1970, a été ressaisi et est disponible dualement dans la base ArXiv en accès libre et dans une nouvelle édition imprimée qui est bien présente dans les bibliothèques.

La diversité des éditeurs (groupe éditorial à visée monopolistique, éditeur privé, éditeur universitaire, société savante, association), de l'organisation des centres de recherche rend

difficile la définition de règles qui puissent être adoptées de manière unanime. Par son caractère réparti, la viabilité de la WDML nécessite néanmoins des accords sur certains points :

- la notion de *accès libre* (*open access*) se réduit souvent à la substitution du lecteur-payeur par l'auteur-payeur (ce sont évidemment les institutions auxquelles sont rattachés les scientifiques qui payent), il faut insister sur l'établissement de barrières mobiles avec accès libre aux archives. La maintenance de ces collections numériques requiert des moyens : les bibliothèques devraient construire de manière concertée des systèmes d'archives réparties, avec un redondance de précaution mais sans doublons à l'infini.
- l'exception prévue à des fins de recherche ou d'étude privée par la directive européenne *sur l'harmonisation de certains aspects du droit d'auteur et les droits voisins dans la société de l'information* aurait dû être retenue automatiquement au moment de sa transposition dans les législations des différents états de l'Union européenne : cela ne semble pas être le cas en France en dépit de la volonté de la communauté scientifique, hélas de faible poids face aux enjeux économiques et financiers de l'industrie audiovisuelle et d'autres produits culturels de masse ; les droits de propriété littéraire et artistique sont préférés au droit d'accès au patrimoine culturel et scientifique...
- les clauses des contrats signés par les auteurs (en leur nom propre ou celui de leurs employeurs, tant les scientifiques aujourd'hui sont des salariés de la science) lors de la publication doivent être rééquilibrés en faveur des auteurs et *in fine* du public. Le dépôt de la version publiée d'un article sur la page personnelle de son auteur est aujourd'hui admise par des éditeurs qui y étaient totalement opposés il y a peu : c'est la *bonne pratique* dans certains domaines de recherche (physique, mathématiques, ...), alors que cette pratique reste quasi-inconnue d'autres champs scientifiques (chimie, sciences de la vie, ...) même si les changements y seront rapides. Ces usages modifieront radicalement les paradigmes de l'accès : quiconque connaîtra le signalement bibliographique d'un article pourra consulter son contenu s'il est proposé sur la page personnelle de l'auteur ou dans l'archive institutionnelle dont il dépend.

La protection des monographies par la norme juridique est plus stricte : on peut s'interroger sur la pertinence de ces règles à une époque où les éditions sont pilonnées de plus en plus rapidement, les ouvrages étant réduits précocement à l'état d'épuisement. La conservation de leurs matrices numériques, en même temps que des exemplaires imprimés, pourrait assurer leur préservation. Pour assurer une conservation véritable, et compte tenu de la fragilité des supports et de l'évolution des formats, un stockage de ces matrices numériques dans plusieurs archives, avec

accès plus ou moins libre, est nécessaire. Les modalités exactes sont encore à définir⁷, mais le souci de privilégier l'accès au patrimoine doit à nouveau être pris fortement en compte.

En particulier pour les articles de périodiques, la période des 70 ans de la convention de Berne est totalement inadaptée. La tendance actuelle qui est à l'allongement de la période réservant les droits, comme l'affaire de Mickey l'illustre, est à inverser. La durée précédant le passage au domaine public doit être considérablement raccourcie et revenir aux vingt-trente ans que fixaient les premiers textes sur le droit d'auteur.

⁷Les articles publiés dans des revues ne donnent pas lieu en général à rétribution directe, ce qui n'est pas le cas des monographies. Cette différence entre articles de revues et ouvrages peut justifier des règles de droit d'auteur différentes. Néanmoins, dans les deux cas, les délais actuels sont par leur étendue un frein à la communication scientifique.