



Rapport Scientifique sur le projet Bibliothèque Historique de l'Éducation Janvier 2021

Le financement Collex-Persée a permis, durant 18 mois (2019-2020), la numérisation, le traitement documentaire et la mise en ligne de trois revues, pour une volumétrie de 155 902 pages, venant compléter les collections déjà disponibles sur la Perséide de la Bibliothèque historique de l'Éducation.

Le projet de Bibliothèque historique de l'Éducation (BHE), initié en 2014, s'inscrit dans le cadre d'une collaboration entre les chercheurs des UMR de l'ENS de Lyon travaillant sur l'histoire de l'éducation (UMR Triangle, UMR LARHRA et UMS Laboratoire de l'Éducation), Persée, la Bibliothèque Diderot de Lyon et intégrant également les éditions de l'ENS et une collaboration avec les Archives nationales. Son objectif est de proposer un portail de ressources historiques en histoire de l'éducation, en particulier autour de l'histoire des politiques éducatives, de leurs institutions et de leurs acteurs, de la profession enseignante et des pratiques pédagogiques. Organisée dans un premier temps autour de la numérisation d'une série de monographies historiques et de dictionnaires publiés par les anciennes éditions de l'INRP et de la numérisation/traitement/mise en ligne d'une partie du *Bulletin administratif de l'Instruction publique* (1850-1890), la première étape a abouti en 2017 à la création d'une perséide : <https://education.persee.fr/>.

Doté d'un conseil scientifique, le projet BHE réfléchit aux corpus les plus pertinents à traiter, en termes de cohérence et d'intérêt scientifique. Il s'agit de mettre à disposition des ensembles de sources et de publications qui s'articulent les uns aux autres. C'est ainsi que le choix de répondre à l'appel à projet du GIS Collex-Persée s'est traduit par la proposition du traitement de trois revues complémentaires aux corpus déjà mis à disposition, autour de deux dimensions principales. Le premier aspect de cette complémentarité concerne la séquence chronologique retenue. Si le projet de la BHE n'est pas défini par des bornes temporelles strictes, une grande partie des documents déjà accessibles concernent une période d'intense activité en matière de politique éducative, à savoir la III^e République. Par ailleurs, les revues retenues constituent des acteurs essentiels dans les débats qui traversent alors les communautés éducatives et les administrations qui en ont la charge. Elles permettent par exemple d'éclairer les choix retraduits dans les décisions réglementaires et législatives, qui sont recensées dans le *Bulletin administratif de l'Instruction publique* (<https://education.persee.fr/collection/baip>). En miroir, la numérisation déjà effectuée de dictionnaires biographiques permet d'accéder facilement à de nombreuses informations concernant les auteurs qui publient dans leurs pages (<https://education.persee.fr/collection/inrp>).

La première de ces revues est la *Revue internationale de l'enseignement* (<https://education.persee.fr/collection/revin>), créée en 1881 par la Société pour l'étude des questions

d'enseignement supérieur et qui prend la suite d'un Bulletin édité entre 1878 et 1881. Cette Société, qui vise à promouvoir la réforme et le développement de l'enseignement dans les premières années de la III^e République, propose dans ses livraisons mensuelles de nombreuses analyses sur la situation française de l'enseignement secondaire et supérieur et sur celle des autres pays occidentaux. Elle rassemble professeurs des facultés parisiennes et provinciales et membres de l'administration de l'Instruction publique. La revue disparaît en 1940 avec la dissolution de la Société. Revue de référence en ce qui concerne les débats et les recherches du système éducatif dans son ensemble, elle a été numérisée intégralement.

Vient en second lieu la *Revue Pédagogique*, qui avec ses 125 volumes parus de 1878 à 1942, (<https://education.persee.fr/collection/revpe>) éditée par Ch. Delagrave, a connu une exceptionnelle longévité. Une première formule de cette publication mensuelle d'initiative privée est publiée de janvier 1878 à juin 1882. À partir de juillet 1882 est lancée la « nouvelle série » qui marque le rattachement de la *Revue pédagogique* au Musée pédagogique comme en témoigne alors son sous-titre : « Musée pédagogique et Bibliothèque centrale de l'enseignement primaire ». Le choix s'est porté sur la séquence 1878-1926, en cohérence avec les autres corpus.

Fondé sous le nom de *Journal de l'Instruction primaire* (1830-1832), par l'éditeur Louis Hachette, ancien élève de l'École normale, le *Manuel général de l'Instruction publique* (<https://education.persee.fr/collection/magen>), troisième titre de cette opération de numérisation, est un périodique mensuel qui avait pour fonction, dans un contexte de développement de l'enseignement primaire, d'apporter aux instituteurs et aux institutrices les informations nécessaires au bon exercice de leur profession. La publication fait office, à certaines périodes de journal officiel. Cela explique sa structure pérenne qui comprend de nombreuses parties réglementaires, au côté des articles à dimension pédagogique. Notre choix s'est porté sur la période 1882-1909, pour les raisons de cohérence préalablement évoquées.

Le travail déjà réalisé les années précédentes avec Persée a permis un traitement direct et efficace des trois corpus numérisés, en particulier concernant leur documentation. Cette dernière a cependant fait l'objet de débats plus approfondis au sujet de la granularité la plus pertinente à retenir lors de cette opération. Il fallait en effet arbitrer entre un choix très ambitieux, qui avait été celui retenu pour le *Bulletin administratif*, et qui avait produit finalement une structuration plus proche de celle d'une base de données que du traitement documentaire généralement choisi par Persée pour ses collections de revue. L'arbitrage final a résulté d'une réflexion approfondie sur la nature des documents mis en ligne, leur unité logique et les usages des chercheurs.

Le découpage a été traité de façon différente pour les trois supports, en fonction de leurs caractéristiques. Pour la *Revue internationale de l'enseignement* et pour la *Revue pédagogique*, organisées selon des logiques traditionnelles aux revues (articles et rubriques générales en fin de volume), le choix a été fait de conserver le découpage le plus proche de la présentation initiale. Dans le cas du *Manuel général*, nous avons décidé en revanche de conserver le découpage en grandes parties thématiques qui scandent la publication, au sein desquelles les différents éléments constituent des entrées de sommaire et non des documents autonomes (ayant leur propres métadonnées), voire ne sont pas détaillés. L'arbitrage a reposé sur le fait que l'interrogation « plein texte », possible grâce à l'OCR, permettait de retrouver ces différents éléments sans difficulté, alors que leur découpage en entités de plus petite taille aurait imposé un allongement considérable du temps de traitement, sans bénéfice évident. On peut prendre pour exemple la « partie scolaire » qui rassemble documents et exercices,

discipline par discipline. Organiser un découpage qui prenne en compte chaque élément (entre 5 et 15 lignes) désigné par sa discipline de rattachement aurait entraîné la création de 5 à 10 entités par page, qu'il aurait fallu documenter. Les reproduire sous forme d'entrées de sommaire aurait été à peine moins long. Or l'usage systématique des termes pour désigner des disciplines scolaires stabilisées permet de retrouver facilement, par le biais du moteur de recherche général, toutes les occurrences des exercices de mathématiques par exemple.

Ces discussions et ces arbitrages ont amené les porteurs du projet à choisir de travailler à partir de 2021 les questions relatives au développement d'outils de fouille, d'indexation, de classement et de mise en relation des corpus déjà présent sur la Perséide. Ce faisant, il sera possible d'améliorer et de diversifier les modes d'interrogation en vue d'optimiser les opérations de traitement documentaires. Un second chantier est envisagé autour des questions d'extraction et de visualisation des résultats.

L'expérience de ces opérations de numérisation et traitement de collections supplémentaires a également permis la poursuite et l'approfondissement du travail collaboratif entamé depuis 6 ans. Les interactions nombreuses entre les différents acteurs investis dans le projet et les habitudes d'une discussion suivie, à la fois scientifique, documentaire et technique ont permis que les décisions soient instruites collectivement, en prenant en compte les contraintes spécifiques de chacun des univers professionnels impliqués.

Emmanuelle Picard

Directrice du département de sciences sociales de l'ENS de Lyon
Maîtresse de conférences HDR en histoire contemporaine
Triangle - UMR 5206 /Le Laboratoire de l'éducation – UMS

Hélène Bégnis

Chargée des partenariats recherche (UMS 3602 Persée)
Coordination du projet pour Persée

Solenn Huitric

Porteuse scientifique du projet
Maîtresse de conférences, Université Lumière Lyon 2

Régine Matéossian

Bibliothécaire à la Bibliothèque Diderot de Lyon

Clément Pieyre

Directeur de la Bibliothèque Diderot de Lyon