
Encyclopédies médiévales en milieu numérique : les nouveaux enjeux de SourcEncyMe pour le traitement des "auctoritates"

Isabelle Draelants*¹

¹Institut de recherche et d'histoire des textes – Centre National de la Recherche Scientifique : UPR841, CNRS : UPR841 – France

Résumé

Cette communication propose un regard rétrospectif et réflexif sur les humanités numériques dans le domaine de l'histoire des textes ou des sciences de l'érudition, à partir de du cas du programme SourcEncyMe dédié à l'étude des sources des encyclopédies médiévales.

Les humanités numériques ne sont plus une discipline récente et le web est un cimetière à projets. L'obsolescence des outils et technologies mises en oeuvre en est une des causes, l'arrêt des financements obtenus par dépôt de projet en est une autre. Ce temps brevissime des financements et de la technologie va à l'encontre de la pérennisation requise par la recherche, et surtout de la nature même de l'érudition, qui nécessite expertise et temps long. L'investissement considérable en travail et en expérience se heurte au vieillissement rapide des outils, et le chercheur en "humanités" doit sortir de son domaine d'expertise pour devenir à la fois entrepreneur et ouvrier de l'édition électronique. Nombre de chercheurs en sciences de l'érudition ont le sentiment de perdre leur âme, car on dispose de moins en moins de temps pour approfondir sa discipline, alors qu'il est devenu inimaginable de ne pas recourir aux outils numériques. Dans le même temps, les exigences ont augmenté : il ne s'agit plus seulement de créer chacun son outil pour mettre en ligne des résultats vérifiés, mais d'utiliser des langages (ex. html), des balisages (ex. TEI) et des référentiels toujours plus nombreux. Il faut désormais aussi offrir en opensource le codage utilisé et la description des données (DMP), sans parler de la transformation du vocabulaire qui doit emprunter au jargon informatique : les contenus érudits de toute nature perdent leur nom pour être étiquetés en "données" ou "métadonnées". Enfin, il faut tendre à une interopérabilité croissante entre les réalisations en humanités numériques.

Tels sont les défis que doivent affronter les projets fondés sur un patrimoine textuel, cherchant à rassembler un corpus de textes vérifiés et à en enrichir le contenu scientifique au fil des ans. SourcEncyMe (Sources des Encyclopédies Médiévales) est un de ces programmes qui doit conjindre croissance des contenus spécialisés et évolution des outils techniques. Après une dizaine d'années d'existence, un déménagement institutionnel en 2014 et une mise à disposition du public en 2016, l'outillage du projet SourcEncyMe nécessite une complète rénovation, en cours, et des moyens humains pour introduire des contenus de la qualité attendue par les publications d'érudition traditionnelles.

SourcEncyMe rassemble les encyclopédies médiévales latines pour étudier leurs sources, c'est-à-dire tous les auteurs et les oeuvres utilisées par les encyclopédistes, en particulier au "siècle

*Intervenant

d'or des encyclopédies médiévales", le 13e siècle. Tissées de 75 à 95 % de citations de textes antérieurs ou contemporains, elles représentent un objet emblématique de l'intertextualité des œuvres médiévales. L'enjeu de leur étude est de donner accès virtuellement à la bibliothèque fréquentée par les " compilateurs " auteurs des encyclopédies, autrement dit à un large héritage littéraire, scientifique et philosophique.

La communication vise à montrer, avec un retour d'expérience sur SourcEncyMe, comment les défis ci-dessus peuvent être affrontés et quelles solutions mises en œuvre. Les objectifs scientifiques du projet et ses réalisations en matière de corpus d'encyclopédies médiévales seront exposés ; les questions relatives à la pseudépigraphie (attribution des *auctoritates*), à l'identification des citations, et aux annotations sur la tradition textuelle, seront abordées à partir d'exemples précis. On explorera également le volet technique du passage de bases de données reliées en PHP-MySQL à une base de données unifiée XML native (BaseX) et d'une interface d'administration collaborative à un outil de balisage " chercheur-friendly " en mode " stand-off embarqué " au format XML-TEI, ainsi que les liens à créer avec les référentiels d'autorités des institutions patrimoniales nationales et internationales, la potentialité de l'identification automatique des sources, et l'interopérabilité possible avec d'autres réalisations.

Mots-Clés: Encyclopédie, Moyen âge, auctoritates, autorités, corpus, sources, citations, humanités numériques, défis techniques, obsolescence