

Comprendre et interpréter les données : enjeux et implantations d'un système de codage dans des gisements de données historiques

Stéphane Lamassé* et Julien Alerini**

*Université Paris 1 Panthéon-Sorbonne - Lamop (laboratoire de médiévistique occidentale)

Stephane.Lamasse@univ-paris1.fr

**PIREH-Lamop

julien.alerini@wanadoo.fr

1 Résumé

L'accès croissant à une information pléthorique et le développement de gisements de données ambitieux posent aujourd'hui deux grands types de difficultés aux historiens.

Le premier consiste à mettre en relation des gisements qui ont été développés de manière indépendante. C'est par exemple le cas pour l'intégration d'un ensemble de bases de données prosopographiques développées entre 1980 et 2010 au Lamop, ou même dans le cadre d'un projet dont le seul lien est une problématique spatiale et temporelle (projet ANR-DFG, *Euroscientia*).

Le deuxième tient en la nature des données introduites dans ces différents systèmes : elles sont souvent hétérogènes, ambiguës, floues. Pour que le chercheur puisse se les approprier, les données doivent faire l'objet d'un véritable travail, afin de comprendre comment elles ont été obtenues, structurées. L'historien doit donc les évaluer et les valider s'il souhaite les mettre en relation. Cette évaluation nécessitant, elle-même de pouvoir être commentée, partagée et critiquée par d'autres chercheurs.

Dans les deux cas, il est nécessaire de développer des outils d'appropriation, qui permettent d'entrer dans le réel historique contenu dans les stocks de données. C'est là la fonction du projet Histobase, un système permettant d'entrer dans la structuration des gisements, d'en évaluer l'information, d'ajouter des couches d'interprétation (qualification de l'information historique) de les évaluer et de partager les données « obtenues ». Chacune des analyses individuelles et collectives fait l'objet d'une mémorisation. Il faut pour cela laisser une place importante aux historiens en tant qu'expert en prêtant une attention particulière aux processus métiers qu'ils mettent en œuvre.

2 Biographie

Stéphane Lamassé et Julien Alerini sont docteurs en histoire médiévale et moderne et enseignent à l'Université de Paris 1.

Comprendre et interpréter les données.

Ils sont, avec Alain Dallo, Benjamin Deruelle et Léo Dumont, membres de l'équipe de recherche de Jean-Philippe Genet sur les usages de l'informatique en histoire. Participant à plusieurs projets ils contribuent à mettre en relation des problématiques de recherches en informatique contemporaines avec des sujets de sciences humaines et sociales.