

**Session « Patrimoine »**  
**22 mars – 13h40-14h05 – Amphithéâtre Cauchy**

**Une approche adaptative pour lier des données géoréférencées du Web.**

*A. Feliachi, N. Abadie, F. Hamdi, B. Bucher*

*Univ. Paris-Est, LASTIG COGIT, IGN, ENSG, Saint-Mandé*

Il est fréquemment nécessaire de mettre en correspondance des données représentant une même entité géographique du monde réel. L'automatisation de cette tâche est généralement fondée, pour chaque paire de ressources décrivant des entités du monde réel, sur la comparaison des valeurs de leurs propriétés considérées comme équivalentes, afin d'estimer leur degré de similarité. Parmi les propriétés dont les valeurs peuvent être comparées, les références spatiales directes (coordonnées ou géométries) occupent une place de choix. Ainsi, de nombreuses mesures ont été proposées afin d'estimer la similarité de ces références spatiales directes en comparant leurs localisations, leurs formes, leurs tailles ou encore leurs orientations. Le paramétrage de ces mesures requiert une expertise importante sur la généalogie et la qualité des références spatiales à comparer.

Or, de plus en plus de sources de données du Web sont produites par saisie participative ou par agrégation de sources hétérogènes et présentent, au sein d'une même source, des références spatiales avec des coordonnées d'origines diverses, des modélisations géométriques différentes ou encore des exactitudes planimétriques différentes et le plus souvent, ceci n'est pas documenté. Dans ce contexte, proposer un paramétrage unique pour estimer la similarité des références spatiales de deux sources de données du Web constitue un réel défi et conduit le plus souvent à créer des relations de correspondance erronées, ou au contraire à ne pas détecter des équivalences pourtant avérées. Dans ce travail, nous proposons une approche permettant d'adapter, pour chaque paire de références spatiales à comparer, le paramétrage de la mesure de similarité qui leur est appliquée en fonction de leurs caractéristiques. Nous avons testé cette approche sur deux jeux de données ouvertes publiés sur le Web : la base Mérimée qui recense les bâtiments protégés au titre des monuments historiques et les ressources de type "Monument historique de Paris" du chapitre francophone de DBpedia.

