

Graphes sémantiques et réseaux sociaux

Camille Roth

Centre Marc Bloch

Résumé La distribution sociale des informations et la structure des interactions sociales sont de plus en plus fréquemment étudiées de manière conjointe, notamment dans les travaux se réclamant des sciences sociales computationnelles. D'une part, l'analyse des contenus, diversement appelée "text mining", "automated text analysis" ou encore "text-as-data methods", s'y appuie sur un vaste éventail de techniques allant de simples statistiques numériques (similarité textuelle, termes saillants) à des approches d'apprentissage automatique s'appliquant au niveau d'ensembles de mots ou de phrases, en particulier en vue d'extraire divers types de graphes sémantiques – qu'il s'agisse simplement de liens de co-occurrence entre termes, de triplets "sujet-prédicat-objet", ou de structures plus élaborées au niveau d'une phrase entière. Ces données et, parfois, ces graphes sémantiques, sont d'autre part associés à des acteurs dont les diverses relations (interaction, collaboration, affiliation) sont également rassemblés fréquemment au sein de graphes sociaux. Cette présentation vise à proposer un tour d'horizon des approches mêlant contenus et interactions, où les espaces publics numériques et les communautés scientifiques représentent des terrains privilégiés en tant que systèmes sociaux où informations et savoirs sont produits et se propagent de manière décentralisée.

Bio Chercheur au CNRS en informatique depuis 2008, Camille Roth a également été enseignant-chercheur en sociologie (professeur à Sciences Po Paris et maître de conférences à Toulouse). Docteur de l'École Polytechnique (2005) et ingénieur des Ponts (2002), ainsi que titulaire d'un DEA de sciences cognitives (EHESS, 2002), il a un profil à la croisée entre sciences dures et sciences sociales. En 2012, il a fondé et dirige depuis lors l'équipe de sciences sociales computationnelles du Centre Marc Bloch à Berlin, où il encadre un groupe interdisciplinaire d'une dizaine de personnes mêlant sciences sociales et modélisation mathématique et informatique. Il y dirige notamment un ERC Consolidator (2018-23) sur le thème du confinement et des bulles dans les espaces publics numériques et y mène également des travaux sur le rôle des algorithmes dans l'accès aux contenus en ligne.

