

Titre :

Un cadre générique pour l'extraction, le traitement, l'analyse et la valorisation de contenus issus de réseaux sociaux.

Résumé :

Les dernières décennies ont connu une croissance significative des sources de données disponibles couvrant de nombreux sujets, ainsi que l'essor de contenus en ligne générés par les utilisateurs, notamment grâce aux réseaux sociaux. Ces données peuvent être exploitées par les professionnels pour mieux comprendre les pratiques et les besoins des utilisateurs. Dans cette présentation, nous proposons un cadre de travail générique se décomposant en quatre modules pour aider les professionnels dans leur prise de décision. Le premier module (collecte) assure la collecte d'informations. Nous proposons une méthode générique qui permet d'extraire des jeux de données à partir d'un réseau social selon trois dimensions complémentaires à définir au préalable : (1) la thématique désirée, (2) la zone géographique à considérer et (3) la période dans le temps souhaitée. Le deuxième module (transformation) vise à extraire des informations fines et à enrichir les données précédemment collectées. Ce module fait notamment appel à des techniques de traitement automatique du langage naturel (TALN). Le troisième module (analyses proxémiques) est lié à l'analyse du jeu de données. Nous proposons une nouvelle approche de traitement également générique fondée sur la proxémique. La proxémique est la relation que les individus entretiennent avec la distance sous tous ses aspects (par exemple, distance entre localisations d'individus, distance entre localisations d'individus et de points d'intérêt, distance entre catégories d'activités menées par des individus, distance entre catégories sociales d'appartenance des individus, etc.). Nous avons adapté la théorie proxémique pour analyser les jeux de données issus des réseaux sociaux afin d'identifier de nouvelles interactions entre les données. Enfin, le quatrième module (valorisation) cible les utilisateurs finaux en proposant différentes manières de présenter les analyses produites (par exemple, *visualisation par tableau de bord, système de recommandation ou de mise en relation, etc.*). Nous avons évalué notre proposition sur des jeux de données provenant de Twitter dans le secteur du tourisme. Nous démontrons dans cette présentation que notre chaîne de traitement peut se généraliser dans d'autres secteurs d'activité, notamment celui de la mode.