

Graphes typés pour l'exploration d'actualités

Rémi Bois*, Guillaume Gravier*, Emmanuel Morin**, Pascale Sébillot***

*CNRS, IRISA & INRIA Rennes, 263 Avenue Général Leclerc, 35042 Rennes
remi.bois@irisa.fr, guillaume.gravier@irisa.fr

**LS2N, Université de Nantes, 2 Chemin de la Houssinière, 44300 Nantes
emmanuel.morin@univ-nantes.fr

***INSA, IRISA & INRIA Rennes, 263 Avenue Général Leclerc, 35042 Rennes
pascale.sebillot@irisa.fr

1 Navigation dans des collections d'actualités

L'accès à l'information n'a jamais été aussi simple. Des dizaines d'articles sont créés chaque minute et rendus accessibles instantanément via les plateformes de médias en ligne. Néanmoins, confrontés à une telle masse d'informations disponibles, les professionnels peinent à rassembler efficacement l'ensemble des faits liés à un évènement. C'est le cas des journalistes, qui doivent mettre en contexte une information qu'ils s'appêtent à publier, ou des attachés de presse qui sont chargés de rassembler les informations disponibles à propos d'un sujet de société. Dans ces deux tâches, l'approche privilégiée consiste à utiliser un moteur de recherche afin de trouver les informations importantes liées à une thématique. Cette approche est cependant imparfaite, une recherche menant à des documents souvent très similaires, des pans entiers d'une information pouvant être introuvables sans l'utilisation du mot-clé pertinent.

D'autres outils permettent l'exploration de grandes collections d'actualités, tels que les frises chronologiques, l'organisation en fils d'actualités, ou le clustering d'informations. Ces outils sont néanmoins inadaptés à grande échelle, générant des frises chronologiques trop longues ou un nombre de clusters trop important. Dans le cadre du projet LIMAH, et en nous fondant sur des entretiens initiaux avec des professionnels afin de déterminer leurs besoins réels en termes de fonctionnalités (Gravier et al., 2016), nous proposons une représentation sous forme de graphe de l'ensemble des documents, avec des liens créés automatiquement entre des paires de documents d'actualité reliés sémantiquement¹. Ces liens sont typés via des heuristiques, permettant d'indiquer à l'utilisateur la nature de la relation entre deux documents.

2 Graphes typés : construction et évaluation par l'usage

L'interface réalisée consiste en un moteur de recherche et un outil de navigation. Après une première requête adressée à un moteur classique, l'utilisateur peut visualiser l'article de son

1. Les entretiens et tests utilisateurs ont été conduits avec l'aide de Éric Jamet et Maxime Robert, chercheurs au CRPCC de Rennes. L'interface a été réalisée avec Arnaud Touboulic, ingénieur.