

Analyses synchroniques et diachroniques des thématiques EGC- Défi ECG 2016

Sofiane Bouzid*, Adrian Tanasescu*

*Institut des Sciences de l'Homme, Lyon
sofiane.bouzid@ish-lyon.cnrs.fr
adrian.tanasescu@ish-lyon.cnrs.fr

Résumé. Les articles scientifiques publiés dans les actes des conférences EGC, qui se déroulent chaque année depuis 2001, constituent la richesse de ces événements mettant en avant le fer de lance de la recherche francophone portant sur la gestion et l'extraction de connaissances. Nous nous sommes penchés sur l'analyse de ces publications scientifiques afin d'en extraire l'essence en termes de thématiques de recherches abordées. Premièrement, nous avons analysé les points communs et les spécificités des publications dans les différentes éditions de la conférence ainsi que les principales différences entre les éditions consécutives. Puis nous nous sommes intéressés à la façon dont les publications s'articulent autour des thématiques extraites et sur lesquelles nous avons essayé de visualiser une approximation sémantique. Enfin nous nous sommes intéressés à l'évolution des thématiques depuis les débuts de cette conférence et jusqu'à l'édition 2015.

1 Introduction et préliminaires

La littérature scientifique se distingue par une structuration qui diffère selon la discipline de l'article. Des unités essentielles, appelées les clés de texte (titre, résumé, mots-clés), sont placées au début de chaque articles, elles permettent de décrire d'une façon brève et efficace tout son contenu (Bénichoux et al. (1985)).

Les analyses présentées dans cet article, ont porté sur le contenu textuel décrit dans les titres et les résumés des publications scientifiques présentées lors des 12 dernières éditions de la conférence EGC.

Premièrement, nous nous sommes concentrés sur l'analyse des occurrences des termes puis la visualisation des fréquences des termes les plus fréquents qui permettent de ressortir les orientations des articles d'une façon orientée *mots-clés*.

Puis, nous avons utilisé des algorithmes de détection de thématiques afin de détecter des typologies de publications.

Enfin, nous avons étudié l'évolution des thématiques découvertes au fur et à mesure des éditions successives de la conférence.