

# Utilisation des termes clés en recherche d'information - le cas du challenge TREC COVID

Adrian Chifu\*, Bernard Dousset\*\*,\*\*\*  
Josiane Mothe\*\*,\*\*\*\*

\*LIS UMR CNRS 7020,  
Aix-Marseille Université/Université de Toulon, France  
adrian.chifu@univ-amu.fr,  
<https://adrianchifu.com>

\*\* Institut de Recherche en Informatique de Toulouse  
UMR 5505 CNRS, Toulouse Prénom.Nom@irit.fr  
<http://www.irit.fr/~Prénom.Nom>

\*\*\* Université Paul Sabatier, Université de Toulouse

\*\*\*\* INSPE Université Toulouse Jean-Jaurès, Université de Toulouse

**Résumé.** En avril 2020, le NIST a proposé un premier cycle d'évaluation de recherche d'information comprenant un ensemble de sujets d'intérêt ou requêtes et une collection de documents scientifiques en lien avec la COVID-19. Cet article présente notre participation à ce challenge en considérant la première série de requêtes du challenge. Notre modèle s'appuie sur l'utilisation des termes clés plutôt que les usuels mots simples ou racines. Nous avons utilisé un modèle d'extraction de termes clés à base de n-grammes. Nous avons comparé les performances du moteur de recherche lorsque les documents sont représentés par des termes simples et par des termes clés. Nous montrons que certaines requêtes bénéficient de la représentation par des termes clés.

## 1 Introduction

Le NIST/TREC<sup>1</sup> a proposé un effort conjoint appelé TREC-COVID<sup>2</sup> au printemps 2020 (Voorhees et al., 2021). Comme les autres challenges proposés par TREC, TREC-COVID vise à rassembler des chercheurs du domaine de la recherche d'information pour évaluer les moteurs sur des tâches spécifiques. Ici, il s'agit de retrouver des documents scientifiques qui répondent à un sujet en lien avec la COVID-19. Cette campagne d'évaluation a lancé cinq cycles successifs d'évaluation. Chaque cycle propose des sujets d'intérêt pour la construction de requêtes et une collection de documents. Après la clôture d'un cycle, les jugements de pertinence correspondant aux documents pertinents de la collection pour chaque sujet d'intérêt sont également fournis.

Dans cet article, nous présentons les résultats officiels que nous avons obtenus lors de notre participation à cette campagne d'évaluation pour le premier cycle et les complétons par

---

1. <https://trec.nist.gov>

2. <https://ir.nist.gov/covidSubmit/>