

WIB : un navigateur intégré pour Wikipédia destiné à l'évaluation participative de modèles de pertinence

Christophe Brouard*, Jean-Pierre Chevallet**, Téo Orthlieb*, Habib Slim*

Univ. Grenoble Alpes, CNRS, LIG, équipes AMA* et MRIM**, Grenoble, France
Christophe.Brouard@univ-grenoble-alpes.fr
<http://ama.liglab.fr/brouard/>

Résumé. Nous présentons ici l'application WIB (pour Wikipedia Integrated Browser) qui permet de naviguer dans les documents Wikipédia en même temps que dans les termes contenus dans ces documents et les catégories auxquelles ils appartiennent. Selon le(s) type(s) d'item(s) considéré(s) en entrée et en sortie (termes, documents ou catégories), la tâche résolue par l'application varie (recherche d'information, extension de requête, extraction de mots clefs, classification automatique,...) mais il s'agit toujours de sélectionner les items pertinents vis-à-vis de la requête en s'appuyant sur un modèle de pertinence. Cette application est un moyen d'expérimenter en ligne différents modèles de pertinence. Toutes les actions des utilisateurs sont enregistrées et stockées dans une base de données en vue d'analyses comparatives ultérieures. Une première version de l'application est déjà en ligne (<http://echo.imag.fr/apps/echopedia/>).

1 Introduction

Au-delà des traditionnels systèmes de recherche d'information qui permettent de sélectionner des documents à partir de mots clefs, on peut s'intéresser plus globalement aux modèles de pertinence (Huang et Soergel, 2013), (Brouard et Nie, 2004) permettant de mettre en relation différentes entités. La recherche de tels modèles qui suppose l'élicitation de la relation liant la requête et l'information pertinente reste un sujet de recherche actuel et la définition de méthodes d'évaluation pour ces modèles est un prérequis. L'évaluation traditionnelle repose sur le paradigme de Cranfield (Harman, 2013) dans lequel un ensemble figé de documents et de jugements de pertinence réalisés par des experts pour des requêtes données servent de base de comparaison aux réponses des systèmes. L'avantage du paradigme de Cranfield est de permettre un certain contrôle. Les experts sont censés fournir des jugements fiables et le nombre de documents étant limité, il est possible d'établir des mesures de rappel (en considérant l'ensemble des réponses pertinentes existant dans le corpus). Néanmoins ce mode d'évaluation suppose la mise à disposition de ressources humaines importantes (beaucoup de jugements de pertinence à produire) et par conséquent ne peut être mis en oeuvre que sur des corpus de taille relativement limitée. Ces ensembles de documents, requêtes et jugements de pertinence associés sont par ailleurs fabriqués de façon artificielle. Considérant l'existence de banques de connaissances accessibles en ligne, la recherche de solutions s'appuyant sur une évaluation participative des systèmes de recherche d'information se développe (Kazaï, 2018). Une approche