

DEFT'07 : une campagne d'évaluation en fouille d'opinion

Cyril Grouin*, Martine Hurault-Plantet*
Patrick Paroubek*, Jean-Baptiste Berthelin*

*LIMSI-CNRS

BP133 – F-91403 Orsay Cedex

{Cyril.Grouin, Martine.Hurault-Plantet, Patrick.Paroubek, Jean-Baptiste.Berthelin}@limsi.fr,
<http://deft.limsi.fr/>

Résumé. Depuis 2005, les campagnes nationales d'évaluation « DEFT » proposent des thématiques de recherche exploratoires axées sur la fouille de texte. L'édition 2007 a porté sur la classification de textes d'opinion : la tâche consistait à attribuer une classe d'opinion à chaque texte d'un corpus, parmi 2 ou 3 classes allant d'un jugement défavorable à un jugement favorable. Quatre corpus ont été mis à la disposition des participants : débats parlementaires sur un projet de loi, critiques de jeux vidéos, critiques de films et de livres, et relectures d'articles de conférences. Dans cet article, nous décrivons d'abord la phase préparatoire de la campagne, avec la collecte des corpus, la définition des mesures d'évaluation, et des tests humains de la tâche. Nous présentons ensuite une analyse des résultats des participants, et les remarques qui en découlent concernant les différents types de corpus. Enfin, nous faisons un bilan synthétique des méthodes proposées à l'évaluation.

Introduction

Dans le domaine du traitement automatique du langage, s'il existe depuis longtemps des campagnes d'évaluation anglophones récurrentes, c'est chose beaucoup plus rare en France. Celles qui ont eu lieu jusqu'ici, telles que GRACE (voir Adda et al., 1998) ou l'ensemble des campagnes TECHNOLANGUE (Mapelli et al., 2004), ont été organisées sous la forme de projets soumis à des appels d'offre, par essence limités dans le temps. Une exception notoire est constituée, au niveau de l'Europe, par la campagne d'évaluation CLEF (Peters, 2000; Braschler et Peters, 2004), qui est multilingue, subventionnée par la CE depuis sa première édition en 2000, et qui, à l'instar de TREC¹, gère plusieurs tâches.

Le défi fouille de textes (DEFT) a été créé en 2005 par un groupe de chercheurs (Prince et al., 2007), dans le but d'initier une série de campagnes d'évaluation francophones sur des thématiques relevant de la fouille de textes. Organisé d'abord par le LRI² puis par le LIMSI³, ce défi est proposé tous les ans sur une thématique différente. Chaque édition a rassemblé

¹<http://trec.nist.gov>

²Laboratoire de Recherche en Informatique, <http://www.lri.fr>

³Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur <http://www.limsi.fr>