RepLab: An Evaluation Campaign for Online Monitoring Systems

Jorge Carrillo-de-Albornoz, Julio Gonzalo, and Enrique Amigó

Abstract Over a period of 3 years, RepLab was a CLEF initiative where computer scientists and online reputation experts worked together to identify and formalize the computational challenges in the area of online reputation monitoring. Two main results emerged from RepLab: a community of researchers engaged in the problem, and an extensive Twitter test collection comprising more than half a million expert annotations, which cover many relevant tasks in the field of online reputation: named entity resolution, topic detection and tracking, reputational alerts identification, reputational polarity, author profiling, opinion makers identification and reputational dimension classification. It has probably been one of the CLEF labs with a larger set of expert annotations provided to participants in a single year, and one of the labs where the target user community has been more actively engaged in the evaluation campaign. Here we summarize the design and results of the Replab campaigns, and also report on research that has built on RepLab datasets after completion of the 3-year competition cycle.

1 Introduction

Corporate reputation has been an intense subject of study in the last 30 years. It has been shown to be one of the most valuable assets of companies and organizations (Doorley and Garcia 2011). Research confirms its great influence on the behavior of all the stakeholders. To begin with, companies with better reputations engender loyalty in consumers across several generations and countries (Alsop 2006). Second, a solid reputation adds value to the actual worth of a company and awakens the interest of investors (Kreps and Wilson 1982). Finally, having a good reputation is crucial to attract highly qualified employees and thereby become more efficient and productive (Chong and Tan 2010). It is only logical that companies and

J. Carrillo-de-Albornoz · J. Gonzalo (✉) · E. Amigó
NLP & IR Group at UNED, Madrid, Spain
e-mail: jcalbornoz@lsi.uned.es; julio@lsi.uned.es; enrique@lsi.uned.es

N. Ferro, C. Peters (eds.), *Information Retrieval Evaluation in a Changing World*, The Information Retrieval Series 41,
https://doi.org/10.1007/978-3-030-22948-1_20

organizations dedicate considerable resources to the management of such a key component of their business development.

Reputation management involves activities that aim at building and preserving a company's reputation. In the past, it was predominantly static, and mainly comprised building an attractive image via marketing campaigns and carefully planned corporate messages. Nowadays, social media have radically changed the traditional reputation management model, giving rise to new channels of communication between companies and their audience. Current technology applications provide users with a wide access to information, enabling them to share it instantly and 24 h a day due to constant connectivity. Information, including users' opinions about people, companies or products, is quickly spread over large communities. In this setting, every move of a company and every act of a public figure, are subject, at all times, to the scrutiny of a powerful global audience. The control of information about public figures and organizations has at least partly moved from them to users and consumers (Hoffman 2008; Jansen et al. 2009b; Glance et al. 2005). So that, for an effective Online Reputation Management (ORM), this constant flow of online opinions needs to be watched.

While traditional reputation analysis is mostly manual, online media make it possible to process, understand and aggregate large streams of facts and opinions about companies and individuals in an automatic manner. In this context, Natural Language Processing plays a key, enabling role and we are already witnessing an unprecedented demand for text mining software for ORM. Although opinion mining has made significant advances in the last few years, most work has been focused on products. However, mining and interpreting opinions about companies and individuals is, in general, a much harder and less understood problem since, unlike products or services, opinions about people and organizations cannot be structured around any fixed set of features or aspects, requiring a more complex modelling of these entities.

RepLab was an initiative promoted by the EU project LiMoSINe,[1] and aimed at structuring research on reputation management as a series of evaluation campaigns in which task design and evaluation methodologies are jointly developed by researchers and the target user communities (reputation management experts). The focus was on detecting challenges and opportunities for language technologies in online reputation monitoring problems, to define appropriate evaluation methodologies, build evaluation test collections with reference annotations provided by reputation experts, and run shared tasks on these collections with research labs from academia and industry.

Replab focused on Twitter data, and was designed to run in a 3-year cycle. The first evaluation campaign was held as a CLEF 2012 activity, and focused on a pilot task around the daily work of reputation experts. The monitoring task for analysts, as studied in RepLab, essentially consisted of searching the stream of tweets for potential mentions to the entity, filtering those that do refer to the entity,

[1] http://cordis.europa.eu/fp7/ict/language-technologies/project-limosine_en.html.