Opinions Corpora Acquisition Software - demo

Zygmunt Vetulani (1), Suleyman Menken (2) and Marta Witkowska (1)
(1) Adam Mickiewicz University in Poznań, Poland
(2) University of Kocaeli, Turkey

Abstract

Opinion processing is, since recently, in focus of interest for computational linguists, public relation experts, marketing companies and politicians. Studies of natural language expression of opinions, desires, emotions and related phenomena require tools and methodologies. Within the empirical approach to these studies, we propose Opinions Corpora Acquisition Software (OCAS), a tool for collection of the empirical data in form of a corpus. The tool is used to collect corpora of the customers' written opinions concerning widely used on-line booking services in the area of hotel reservation (via Booking.com).

Keywords: text language resources, opinion processing, corpus-based methods, multilingual corpora, customers' opinions, OCAS

Figure 1. An example of the Booking.com opinion record completed by the guest

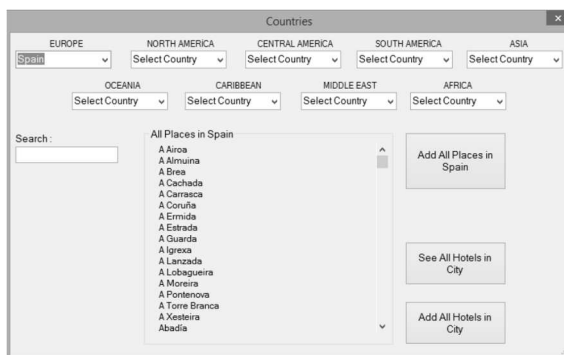Below (Figure 2. a-f) we present the screenshots of Opinion Corpora Acquisition Software (OCAS)
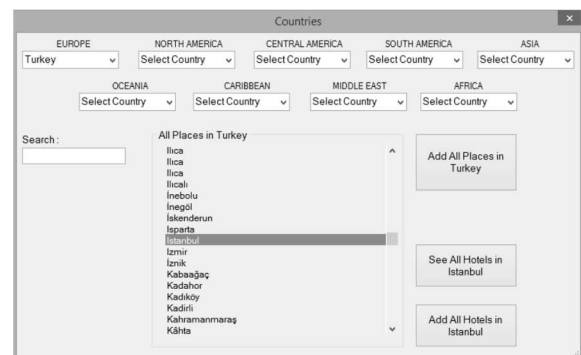

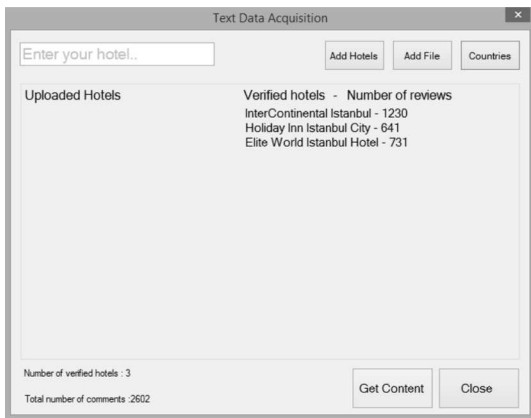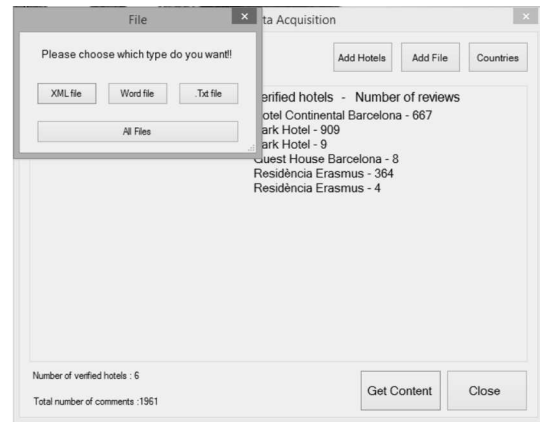Figure 2 a).   b).

c).

d).

e).
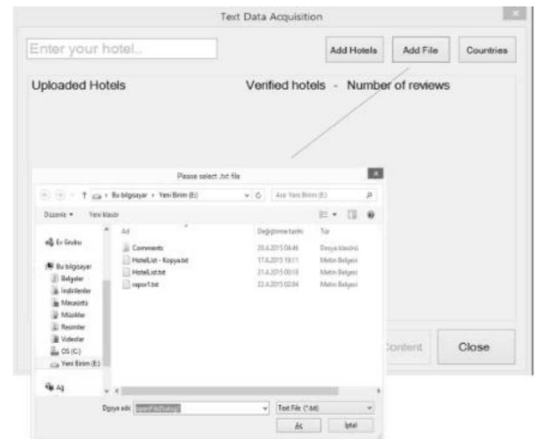
f).

Format of final data present Figure 3 to 6.

Figure 3. A fragment of the XML formatted corpus presenting three different opinions about hotels in Istanbul.
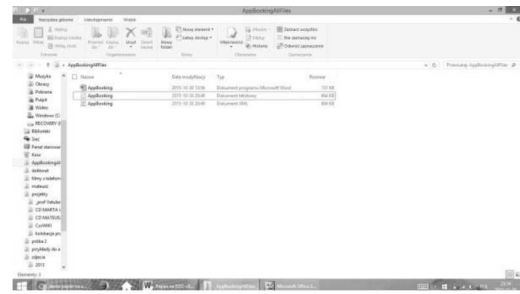
Figure 4.

Figure 5.
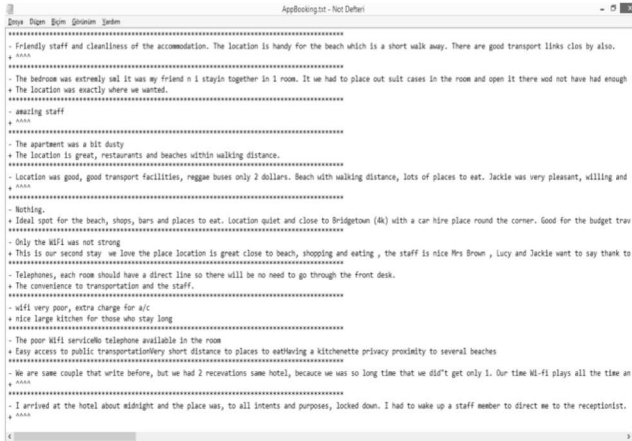

Figure 6.

Bibliography:

Vetulani, Z., Witkowska. M., Menken, S.(2015): Corpus Based Studies on Language Expression of Opinions, Proceedings of LTC 2015, Poznań.