

Introduction

In the last decade, several approaches to linguistic linked data (LLD) have proposed vocabularies for specifying linguistic resources, such as dictionaries, lexicons, and corpora (e.g. Ontolex Lemon [1]). Some of their benefits, such as seamless integration of analysis and resources, are especially relevant for sentiment analysis (SA) since multimodal approaches are increasingly popular. One of the most notable systemic attempts to converge research on SA and LLD was made at the 5th International Workshop on Emotion, Social Signals, Sentiment and Linked Open Data, held at LREC 2014 (ES*LOD 2014, [2]).

SA is defined in Wikipedia as referring “to the use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information” [3], while (open) LLD is “concerned with creating, sharing, and (re-)using language resources in accordance with Linked Data principles” [4]. While the former began in the 1960’s and has evolved substantially since the turn of the century, the latter emerged prominently over the last decade.

The workshop on Sentiment Analysis and Linguistic Linked Data (SALLD) was established with the aim of converging SA and LLD and offering a forum for discussion about the usage of LLD principles in the SA field, to explore relevant principles, methodologies, resources, tools, and applications, and to understand the primary approaches, their advantages and limitations, and the available resources and case studies. SALLD-1 [5] was initiated in the context and with the support of the COST Action ‘NexusLinguarum – European network for Web-centered linguistic linked data’ (CA18209 [6]) and was held in conjunction with the 3rd Conference on Language, Data and Knowledge – in Zaragoza, Spain on September 1, 2021 (LDK 2021 [7]).

The half-day event began with an invited talk by Carlos A. Iglesias, introducing the two fields and their convergence (**Sentiment Analysis meets Linguistic Linked Data: An overview of the state-of-the-art**). This was followed by one paper emphasizing SA (**Extending a general sentiment lexicon to specific domains in a (semi-)automatic manner**, by Pavel Brazdil, Purificação Silvano, Fátima Silva, Shamsuddeen Hassan, Fátima Oliveira, João Cordeiro and António Leal) and another emphasizing LLD (**Multilingual knowledge systems as Linguistic Linked Open Data for European Language Grid**, by Alena Vasilevich and Michael Wetzel). The second part of the workshop included three papers strongly combining the two fields (**LOD-connected offensive language ontology and tag set enrichment**, by Barbara Lewandowska-Tomaszczyk, Slavko Žitnik, Anna Bączkowska, Chaya Liebeskind, Jelena Mitrović and Giedre Valunaite Oleskeviciene; **Extending and using a sentiment lexicon for Latin in a Linked Data framework**, by Rachele Sprugnoli, Marco Passarotti, Marinella Testori and Giovanni Moretti; and **Aspect-based Sentiment Analysis of conference review forms with Linked Data-enabled review criteria**, by Sára Juranková, Vojtěch Svátek and Chiara Ghidini).

These six papers with this introduction form the proceedings of SALLD-1.

Ilan Kernerman

- [1] <https://www.w3.org/community/ontolex/>
- [2] <http://lrec2014.lrec-conf.org/en/conference-programme/workshops-and-tutorials/>
- [3] https://en.wikipedia.org/wiki/Sentiment_analysis
- [4] https://en.wikipedia.org/wiki/Linguistic_Linked_Open_Data
- [5] <https://salld.org>
- [6] <https://nexuslinguarum.eu/>
- [7] <http://2021.ldk-conf.org/>