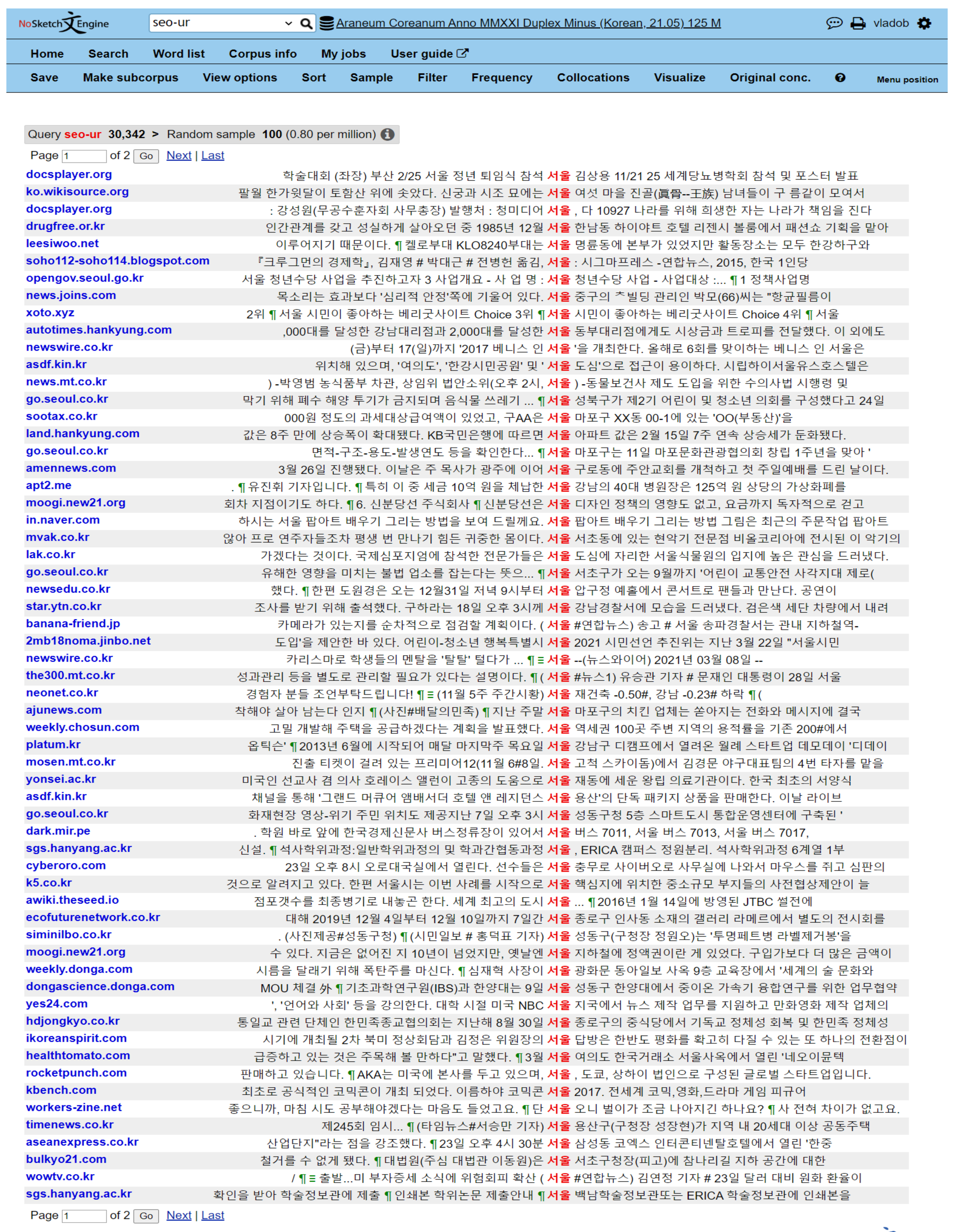
Aranea Web-Crawled Corpora: A Source of Diverse and Unified Language Data for NLP

Vladimír Benko | vladimir.benko@juls.savba.sk
Slovak Academy of Sciences, Ľ. Štúr Institute of Linguistics, Bratislava, Slovakia
Comenius University Science Park, UNESCO Chair in Plurilingual and Multicultural Communication, Bratislava, Slovakia
http://aranea.juls.savba.sk/guest | http://unesco.uniba.sk/guest



s