

AFRILEX 2021: International Keynote

Automating Dictionary Production

Miloš JAKUBÍČEK

Masaryk University, Brno, Czech Republic

In the talk I will present recent efforts in automatic dictionary drafting and production. I will briefly summarize some historical aspects of the topic, its relatedness to corpus linguistics and corpus building, state of the art in terms of technology used and outstanding challenges. The main portion of the presentation focuses however on the impact of such a post-editing lexicography on the overall lexicographic workflow and editorial management.

This post-editing workflow will be exemplified on a set of three projects in bilingual lexicography that we recently carried out.

In these projects, we applied the post-editing workflow to create a Lao, Urdu and Tagalog dictionary, with translations into Korean and English, entirely from scratch, each comprising 50,000 entries.

I will detail the workflow used as well as bottlenecks and issues we discovered, and lessons learnt from the project.