



PolEval 2020 — let's get it started!

Maciej Ogrodniczuk | Institute of Computer Science
Łukasz Kobyliński | Polish Academy of Sciences

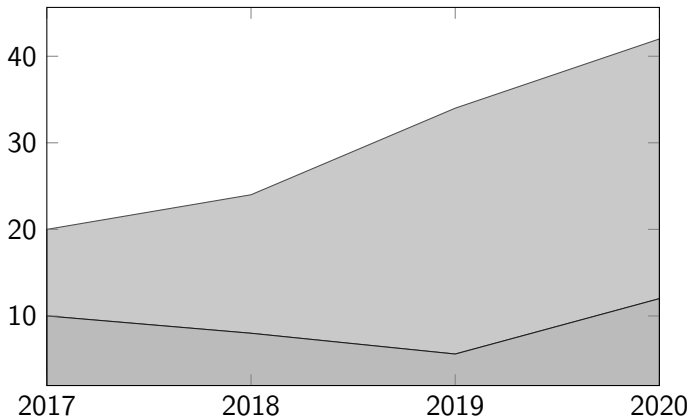
PolEval Workshop at the AI & NLP Conference
26 October 2020

PolEval in a nutshell

The evaluation campaign:

- *Eval-inspired
- for NLP tools for Polish
- promoting research on language and speech technologies
- creating objective evaluation procedures
- and (hopefully) improving state-of-the-art

PoIEval 2017–2020



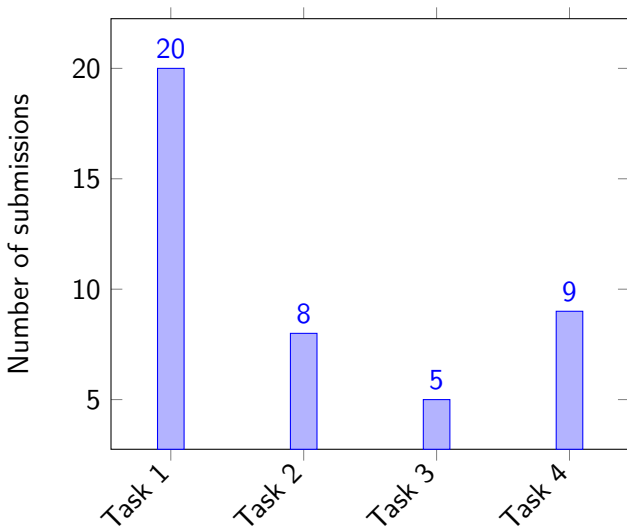
Submissions per task (avg)
Submissions

PolEval 2020

Four tasks:

- Task 1: Post-editing and rescoring of automatic speech recognition results
- Task 2: Morphosyntactic tagging of Middle, New and Modern Polish
- Task 3: Word Sense Disambiguation
- Task 4: Information extraction and entity typing from long documents with complex layouts

PolEval 2020: Submissions per task



PolEval 2020 workshop proceedings

Proceedings of the PolEval 2020 Workshop

Maciej Ogrodniczuk, Łukasz Kobyliński (eds.)

Agenda

Task 1: (11:10–12:45)

- **Danijel Koržinek**: Results of the PolEval 2020 Shared Task 1: Post-editing and rescoring of automatic speech recognition results
- **Tomasz Ziętkiewicz**: Post-editing and Rescoring of ASR Results with Edit Operations Tagging
- **Dominika Wnuk, Krzysztof Wołk**: Post-editing and Rescoring of Automatic Speech Recognition Results with OpenNMT-APE
- **Adam Kaczmarek, Tomasz Syposz, Michał Martusewicz, Jan Chorowski, Paweł Rychlikowski**: t-REx: The Rescorer-Extender approach to ASR improvement

Agenda

Task 2: (13:15–14:20)

- **Marcin Woliński, Witold Kieraś, Dorota Komosińska, Włodzimierz Gruszczyński**: Results of the PolEval 2020 Shared Task 2: Morphosyntactic Tagging of Middle, New and Modern Polish
- **Krzysztof Wróbel**: KFTT: Polish Full Neural Morphosyntactic Tagger
- **Piotr Rybak, Agnieszka Rybak**: Simple yet Effective Baseline for Morphosyntactic Tagging of Middle, New and Modern Polish
- **Wiktor Walentynowicz, Tomasz Kot**: CMC Tagger in the Task of Tagging Historical Texts

Agenda

Task 3: (14:25–15:00)

- **Arkadiusz Janz, Maciej Piasecki**: Results of the PolEval 2020 Shared Task 3: Word Sense Disambiguation
- **Dariusz Kłeczek**: Polbert: Attacking Polish NLP Tasks with Transformers

Agenda

Task 4: (15:05–16:00)

- **Filip Graliński, Anna Wróblewska, Dawid Lipiński:**
Results of the PoEval 2020 Shared Task 4: Information Extraction and Entity Typing from Long Documents with Complex Layouts
- **Michał Marcińczuk, Michał Olek, Marcin Oleksy, Jan Wiczorek:** CLEX — Information Extraction from Documents with Complex Layouts
- **Aleksandra Świetlicka, Adam Iwaniak, Mateusz Piwowarczyk:** BiLSTM Recurrent Neural Networks in Heterogeneous Ensemble Models for Named Entity Recognition Problem in Long Polish Unstructured Documents

Last but not least

Wrap-up: (16:00–16:20)

- **Łukasz Kobyliński:** PolEval: Future Perspectives