

PolEval 2019 —
good to see you again!

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

PolEval 2019 Workshop @ AI & NLP Workshop Day
May 31, 2019

PolEval in a nutshell

What is PolEval?

- SemEval-inspired evaluation campaign for NLP tools for Polish

Rules for tools:

- tools **compete** against one another
- within certain **tasks** selected by organizers
- using available **data** prepared by task chairs
- are **evaluated** according to pre-established **procedures**

Why PolEval?

- we have plenty of LRT for Polish, cf. CLIP website
- but at the same time we lack(ed) meaningful comparisons of the tools (datasets, evaluation metrics, procedures...)

PolEval in the past

Two tasks in 2017:

- 1 part-of-speech tagging
 - A morphosyntactic disambiguation and guessing
 - B lemmatization
 - C complete system
- 2 sentiment analysis

Three tasks in 2018:

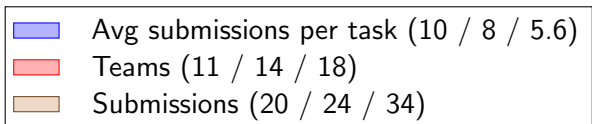
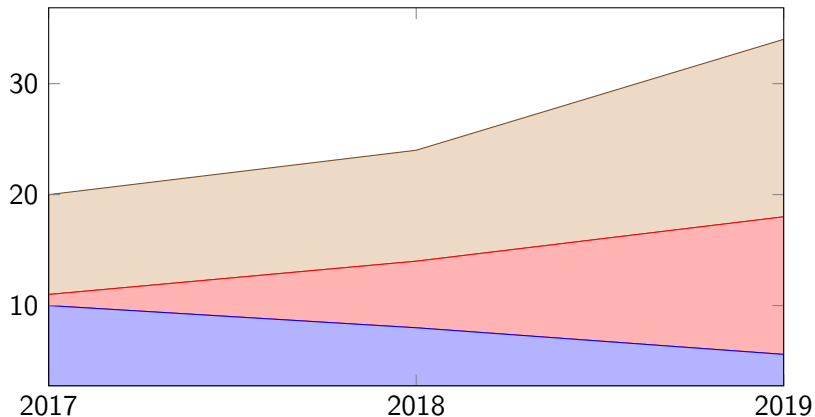
- 1 dependency parsing
 - A morphosyntactic prediction of dependency trees
 - B beyond the dependency tree
- 2 named entity recognition
- 3 language models

PolEval in 2019

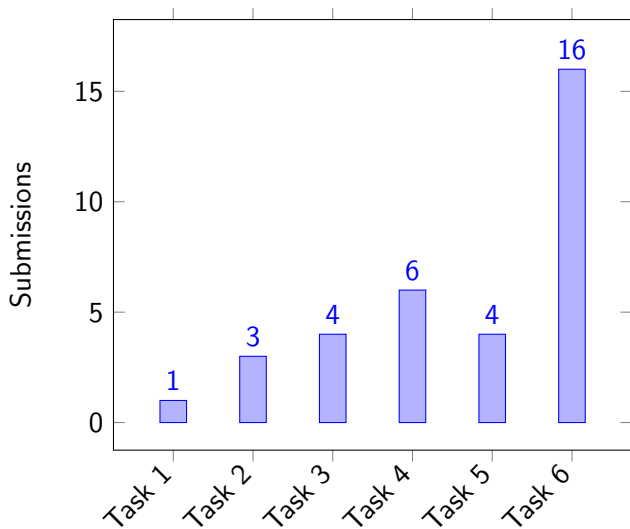
Six tasks in 2019:

- 1** recognition and normalization of temporal expressions
- 2** lemmatization of proper names and multi-word phrases
- 3** entity linking
- 4** machine translation
 - A** EN-PL
 - B** PL-RU
 - C** RU-PL
- 5** automatic speech recognition
- 6** automatic cyberbullying detection
 - A** harmful vs non-harmful
 - B** type of harmfulness

PolEval: Submissions and teams / year



PoEval 2019: Submissions / task



Was it worth it?

No doubt it was!

- many new resources created and made available
- new tools publicly released
- new tasks carried out
- not just accuracy but also tagging speed taken into account by participants
- accuracy improvements over previous state-of-the-art

PolEval 2020 and beyond

Call for tasks:

- fake news detection
- detection of political sentiment
- text normalization
- machine translation
- hate speech detection
- **your task!**

What's next?

To do:

- propose a common licensing scheme for all submissions
- make the rules more explicit
- automate the submission process
- give participants more time to prepare their submissions

The future of PolEval:

- an open platform for competitions?
- an open platform for sharing competition-related data?
- an open platform for annotating data?

Thank you!

And all our task chairs:

- 1 Recognition and normalization of temporal expressions**
Jan Kocoń
- 2 Lemmatization of proper names and multi-word phrases**
Michał Marcińczuk
- 3 Entity linking**
Aleksander Smywiński-Pohl
- 4 Machine translation**
Krzysztof Wołk
- 5 Automatic speech recognition**
Danijel Koržinek
- 6 Automatic cyberbullying detection**
Michał Ptaszyński, Agata Pieciukiewicz, Paweł Dybała