

PolEval 2021

Maciej Ogrodniczuk(Instytut Podstaw Informatyki PAN)Łukasz Kobyliński(Instytut Podstaw Informatyki PAN / Sages)



Started with a series of observations

- Polish is more difficult for automated processing (higher complexity due to inflection and free word order)
- Data for Polish is harder to get (fewer language users, fewer data collection and annotation projects)
- Fewer researchers working on Polish NLP (than on English)
- Lack of good exchange of knowledge between researchers as well as between researchers and business (fewer science-business projects, scientific community is not very large)

The main motivation

Since

the work on processing of Polish needs to be supported via promotional activities, creation of open resources and open methods that can be used in practice

why don't we

organise an open competition for teams working on Polish NLP?

Let's get concrete



And create a competition modelled on Kaggle but linked to tasks in the NLP area for Polish, which aims to:

- develop established evaluation procedures for NLP tasks
- producing annotated **datasets** that can be used for teaching and evaluating systems
- objectively **compare systems** performing relevant tasks
- bring together researchers from the scientific and business communities and to share knowledge between them
- to promote NLP for Polish

5 years of PolEval





Why taking part?

Prizes and accolade 🙂





Why taking part?

Wide recognition and publication!



CHALLENGES

New in 2021

Evaluation platform <u>https://beta.poleval.pl/</u>

- automated
- has leaderboard
- gonito.net-based (Filip Graliński AMU/Applica)

Discord server



Quiz Question Answering Challenge

The goal of the task is to develop solution capable of providing answers to generalknowledge questions typical for the popular TV quiz shows, such as Fifteen to One (PL: Jeden z dziesięciu).



Evaluation of translation quality assessment metrics (nonblind)

The task is to investigate metrics for automatic evaluation of machine translation results or other similar data types.



Evaluation of translation quality assessment metrics (blind)

The task is to investigate metrics for automatic evaluation of machine translation results or other similar data types.



Tasks of 2021



- 1. Punctuation restoration from read text Agnieszka Mikołajczyk, Piotr Pęzik, Adam Wawrzyński, Adam Kaczmarek, Wojciech Janowski, Michał Adamczyk (Voicelab and University of Lodz)
- 2. Evaluation of translation quality assessment metrics Krzysztof Wołk (Polish-Japanese Academy of Information Technology), Maciej Szymkowski (Bialystok University of Technology)
- 3. Post-correction of OCR results Szymon Rynkun (University of Warsaw), Łukasz Kobyliński, Witold Kieraś (IPI PAN)
- 4. Question answering challenge Maciej Ogrodniczuk, Piotr Przybyła (IPI PAN)