



Biuletyn Polonistyczny

27.01.2022

Termin zgłoszeń: 28.02.2022

PolEval - konkurs dla narzędzi informatycznych do przetwarzania języka polskiego

Instytucje: Instytut Podstaw Informatyki Polskiej Akademii Nauk PAN

Celem szóstej już edycji konkursu jest poprawa jakości istniejących rozwiązań, przetestowanie nowych algorytmów oraz promocja inżynierii lingwistycznej.

Zgłoszenia od osób i organizacji przyjmowane są do końca lutego 2022 roku. W ramach zgłoszenia organizatorzy proszą o przesłanie:

- propozycji tematu zadania,
- krótkiego opisu przedstawiającego intencję zadania, uzasadniającego jego aktualność, a także potencjalnie wskazując inne konkursy, w których takie zadanie było organizowane (lub też wskazując na jego unikalność),
- stan przygotowania zadania, tzn. dostępność zasobów i gotowość procesu ewaluacji ew. zgłoszeń,
- czasu i innych zasobów potrzebnych do sfinalizowania koncepcji uruchomienia zadania w ramach PolEval 2022.

Wstępny harmonogram konkursu PolEval 2022 przedstawia się następująco:

- do końca lutego 2022 - ustalenie listy zadań,
- marzec/kwiecień 2022 - ogłoszenie zadań na stronie konkursu, w tym danych treningowych,
- lipiec/sierpień 2022 - ogłoszenie danych testowych,
- sierpień/wrzesień 2022 - zakończenie przyjmowania zgłoszeń od uczestników,
- wrzesień/październik 2022 - konferencja podsumowująca tę edycję konkursu.

Dotychczas zostały zorganizowane następujące zadania:

- Part of Speech Tagging
- Sentiment Analysis
- Dependency Parsing
- Named Entity Recognition
- Language Models
- Recognition and normalization of temporal expressions
- Lemmatization of proper names and multi-word phrases
- Entity linking
- Machine translation

Biuletyn Polonistyczny

- Automatic speech recognition
- Automatic cyberbullying detection
- Post-editing and rescoring of automatic speech recognition results
- Morphosyntactic tagging of Middle, New and Modern Polish
- Word sense disambiguation
- Information extraction and entity typing from long documents with complex layouts
- Punctuation restoration from read text
- Evaluation of translation quality assessment metrics
- Post-correction of OCR results
- Question answering challenge

W zadaniach tych rywalizowało dotąd łącznie ponad 120 systemów przygotowanych zarówno przez przedstawicieli środowisk naukowych, jak i firm prywatnych.

Informacje o konkursie dostępne są na stronie <http://poleval.pl>.

Platforma ewaluacyjna znajduje się pod adresem: <https://beta.poleval.pl/>

Uruchomiony został również serwer Discord, do prowadzenia dyskusji bieżących: <https://discord.gg/dpn94tUSyT>

Słowa kluczowe: inżynieria lingwistyczna, konkurs