Katarzyna Frąckowiak

Niestandardowe CV

Językoznawczyni · Ewaluacja LLM · polski / angielski

kasia420.github.io

Poznań · katarzynaglowne@gmail.com · kasia420.github.io · github.com/Kasia420

PODSUMOWANIE

Studentka studiów magisterskich Language, Mind & Technology na Uniwersytecie im. Adama Mickiewicza, łącząca językoznawstwo z ewaluacją dużych modeli językowych (LLM). Projektuję polskojęzyczne zestawy testowe sprawdzające, jak modele radzą sobie z ironią, implikaturą i gramatyką, a w pracy magisterskiej badam, jak modele NLP przetwarzają angielski użytkowników nienatywnych. Polski jako język ojczysty, angielski na poziomie C1; praktyczne doświadczenie w pisaniu precyzyjnych pytań testowych, ocenie odpowiedzi modeli i analizie wyników w Pythonie i R.

PROJEKTY

Zestawy testowe do ewaluacji LLM w języku polskim — projekt własny · paź 2026 – obecnie

- Zaprojektowałam cztery zestawy po 50 pozycji badające rozumienie języka polskiego przez LLM: rozumienie ironii, implikatury, przypadki a szyk zdania („kto komu") oraz fałszywi przyjaciele angielsko-polscy
- Tworzę pary minimalne i krótkie dialogi z odpowiedziami zamkniętymi; testuję pytania na kilku modelach, aby usunąć przypadki niejednoznaczne lub zbyt łatwe
- Zestawy przekazane osobie ze społeczności SpeakLeash do ewaluacji modelu Bielik

Lost in Translation — projekt badawczy z kursu, rozszerzony · 2026 · kasia420.github.io/lost_in_translation

- Porównałam odpowiedzi modelu Claude na 25 par pytań (angielski/polski) o polską kulturę, mierząc podobieństwo międzyjęzykowe (LaBSE, mediana cosinusów .81) i wydźwięk (XLM-RoBERTa) na tle polskiej i angielskiej Wikipedii
- Polskie odpowiedzi okazały się istotnie mniej negatywne (test Wilcoxona, p = .0015, r = .70); powtórzenie badania z ustaloną wersją modelu, kontrolą długości i kalibracją klasyfikatora na 1000 parach zdań pokazało podobną różnicę w samej Wikipedii; wyniki zaprezentowane na interaktywnej stronie projektu

Praca magisterska (w toku) — Jak LLM przetwarzają angielski użytkowników z polskim jako językiem ojczystym

DOŚWIADCZENIE

Stażystka badawcza — Pracownia Psycholingwistyki i Neurolingwistyki, UAM · 2024 – 2025

- Prowadziłam sesje EEG/ERP zgodnie ze ścisłym protokołem eksperymentalnym, dbając o porównywalność danych między uczestnikami
- Pracowałam z paradygmatami primingu semantycznego i decyzji leksykalnej; przeszłam szkolenie z podstaw sygnału EEG (pasma częstotliwości, rodzaje artefaktów)

Wolontariuszka — 9. Międzynarodowa Konferencja CADAAD, UAM · lipiec 2024

- Koordynowałam rejestrację gości i wsparcie na miejscu dla zagranicznych uczestników konferencji poświęconej analizie dyskursu

Korepetytorka języka angielskiego — działalność własna · 2020 – 2023

- Zajęcia indywidualne dla polskojęzycznych uczniów; codzienna praca z błędami wynikającymi z transferu z języka polskiego (przedimki, przyimki, szyk zdania), czyli zjawiskiem, które dziś badam w LLM

WYKSZTAŁCENIE

Studia magisterskie, Language, Mind & Technology — Uniwersytet im. Adama Mickiewicza w Poznaniu · 2025 – obecnie

Studia licencjackie, Filologia angielska: literatura i kultura — Uniwersytet im. Adama Mickiewicza w Poznaniu · 2022 – 2025

UMIEJĘTNOŚCI

- **Programowanie:** Python (pandas, NumPy, SciPy, matplotlib), R (tidyverse, ggplot2), Git / GitHub, Docker, REST API
- **NLP / LLM:** HuggingFace transformers, sentence-transformers, wielojęzyczne embeddingi (LaBSE), ewaluacja LLM, projektowanie benchmarków, prompt engineering
- **Dane i statystyka:** testy nieparametryczne, wielkość efektu, modele mieszane, czyszczenie danych
- **Web i wizualizacja:** HTML, CSS, JavaScript, interaktywna wizualizacja danych, GitHub Pages
- **Automatyzacja:** n8n

JĘZYKI

Polski — ojczysty · Angielski — C1 (akademicki) · Hiszpański — A1 · Niemiecki — A1/A2

CERTYFIKATY

Generative AI in Business — University of Cambridge (Cambridge Advance Online) · 2025

DataCamp — Introduction to R, Introduction to the Tidyverse, Data Visualization with ggplot2, Exploratory Data Analysis in R · 2026

Wyrażam zgodę na przetwarzanie moich danych osobowych dla potrzeb niezbędnych do realizacji procesu rekrutacji zgodnie z Rozporządzeniem Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. (RODO).