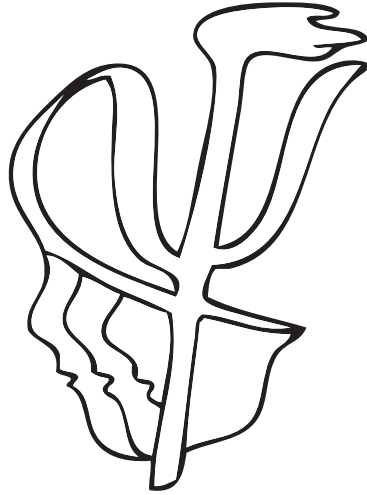




**Words, Vectors, and Feelings: Advancing Psychological Emotion Research
Through Natural Language Processing**

MSc Hubert Wojciech Plisiecki
Graduate School for Social Research
Institute of Psychology
Polish Academy of Sciences

Doctoral Thesis prepared under the supervision of
dr hab. Artur Pokropek
and auxilliary supervision of dr. Grzegorz Pochwatko

Warsaw, Poland, 2025

Funding information: The articles published as a part of this thesis were supported by the SONATA BIS grant from the Polish National Science Center "Research Lab for the Digital Social Sciences" (2020/38/E/HS6/00302) led by dr hab. Artur Pokropek

Streszczenie

Dotychczasowe badania psychologiczne wykorzystujące przetwarzanie języka naturalnego (NLP) w dużej mierze zapożyczały gotowe narzędzia z informatyki, często pomijając psychologiczne niuanse, które są kluczowe dla pełnego wykorzystania ich potencjału w tej dziedzinie. Niniejsza dysertacja ma na celu skorygowanie tego niedopatrzenia, tworząc nowe metody NLP dostosowane do potrzeb psychologii i uwzględniające specyficzne dla niej kwestie, takie jak trafność ekologiczna, stronniczość konstruktu i wierność pomiaru, otwierając tym samym nowe możliwości teoretyczne, praktyczne i pomiarowe w psychologii emocji.

Pierwsze badanie pokazuje, że podstawową strukturę emocji można odtworzyć bezpośrednio z codziennego języka. Korzystając z 58 000 postów z Reddita oznaczonych 28 emocjami, stworzono nienadzorowane, numeryczne reprezentacje emocji i poddano je analizie głównych komponentów składowych. Dwa pierwsze komponenty odtwarzają klasyczne wymiary walencji i pobudzenia, co dowodzi, że kołowy model emocji Russella wyłania się organicznie z naturalnego dyskursu; badanie to dotyka problemu czy struktury wywiedzione z laboratorium generalizują się na język „in-the-wild”, dostarczając jednocześnie metody do wydobywania ukrytych wymiarów psychologicznych z dużych korpusów i poszerzając zarówno potencjał teoretyczny, jak i eksploracyjny psychologicznych badań z użyciem tekstu.

Drugie badanie rozszerza bazy norm afektywnych przy użyciu modeli transformerów dostrojonych na podstawie istniejących ocen. Stworzone modele przewidują szereg wskaźników afektywnych na poziomie słów dla niewidzianych wcześniej wyrazów, osiągając korelacje do $r = 0,95$ z zestawem testowym w języku angielskim i porównywalnie wysokie wyniki w pięciu innych językach. Badanie to rozwiązuje praktyczny problem niedoboru danych normatywnych – częsty problem dziedziny – i wprowadza zautomatyzowane procedury ekstrapolacji oraz algorytm „stimuli-descent”, które pozwalają badaczom dobierać semantycznie dopasowane pary słów różniące się wyłącznie na docelowych wymiarach, ułatwiając projektowanie lepiej kontrolowanych eksperymentów.

Trzecie badanie audytuje uprzedzenia społeczne w nowym polskim modelu sentymentu. Analizy regresyjne ujawniają, że płeć polityka i przynależność partyjna wyjaśniają do 66% wariancji przewidywanej walencji - efekt zakorzeniony w osobistych uprzedzeniach anatorów. Empiryczne ukazanie, w jaki sposób nadzorowane modele NLP mogą propagować psychologicznie ważne uprzedzenia społeczne i polityczne, otwiera nowe ścieżki metodologiczne do ich diagnozowania i korygowania, ulepszając rzetelność pomiaru psychologicznego oraz wniosków naukowych.

Czwarte badanie przedstawia Semantic Propagation Graph Neural Network – wyjaśnialny model sentymentu zaprojektowany tak, aby zachować wysoką dokładność predykcji przy jednoczesnym ograniczeniu propagowanych uprzedzeń społecznych. „Oślepiając” sieć na tożsamość słów i pozwalając, by informacja emocjonalna pojedynczych wyrazów rozprzestrzeniała się wyłącznie przez powiązania syntaktyczne, model osiąga wydajność zbliżoną do transformerów, a jednocześnie znacząco redukuje uprzedzenia w predykcjach, dowodząc, że wysoka dokładność i bezstronność nie wykluczają się nawzajem. Architektura ta bezpośrednio odpowiada na problem równoważenia trafności i neutralności etycznej oraz oferuje narzędzie do śledzenia, jak znaczenie emocjonalne rozchodzi się w strukturze składniowej, poszerzając przyszłe możliwości zastosowania modelowania grafowego w psychologii.

Łącznie te cztery badania dostarczają zweryfikowanego zestawu metod eksploracyjnych, predykcyjnych, diagnostycznych i korekcyjnych, które umożliwiają psychologom badanie umysłu i zachowania poprzez tekst, jednocześnie unikając metodologicznych i etycznych pułapek gotowych narzędzi NLP. Razem odpowiadają na palące wyzwania specyficzne dla dyscypliny – trafność ekologiczna, pokrycie normatywne, ograniczanie stronniczości i interpretowalność – otwierając zarazem nowe ścieżki dla przyszłej teorii psychologicznej, zastosowań eksperymentalnych oraz szeroko zakrojonych, pomiarów bazujących na tekście.