

Bydgoszcz, 02.01.2024

dr hab. inż. Rafał Kozik, profesor uczelni  
Politechnika Bydgoska  
im. Jana i Jędrzeja Śniadeckich w Bydgoszczy,  
Wydział Telekomunikacji, Informatyki i  
Elektrotechniki,  
Al. prof. S. Kaliskiego 7,  
85-796 Bydgoszcz

Rada Naukowa Dyscypliny  
INFORMATYKA TECHNICZNA  
I TELEKOMUNIKACJA  
Sekretariat  
Data wpływu.....  
Numer.....

**Recenzja rozprawy doktorskiej pt.:**

**„Multilingual Machine Translation System for Dialogue Agents  
(Wielojęzyczny system tłumaczenia maszynowego dla agentów  
dialogowych)”,**

**której Autorem jest Pan mgr inż. Marcin Sowański**

**1) Jaki jest problem naukowy (teza) rozprawy? Czy został on trafnie i jasno sformułowany?**

Rozprawa, której Autorem jest Pan mgr inż. Marcin Sowański, dotyczy złożonych zagadnień tłumaczenia maszynowego danych treningowych, walidacyjnych i testowych przy opracowywaniu modeli rozumienia języka naturalnego. W szczególności, przedstawione i opisane w rozprawie zagadnienia badawcze skupione są na inteligentnych wirtualnych asystentach (IVA).

Przedstawione badania i wyniki eksperymentów mają na celu udowodnić, że tłumaczenie maszynowe jest efektywnym sposobem rozszerzania możliwości inteligentnych wirtualnych asystentów o nowe języki. W szczególności Autor rozprawy porusza aspekty związane z usprawnieniem i automatyzacją procesu tworzenia asystentów.

Pan mgr inż. Marcin Sowański definiuje trzy hipotezy, które (podobnie jak problem naukowy rozprawy) są jasno zdefiniowane. Pierwsza teza zakłada, iż tłumaczenie maszynowe stanowi efektywne narzędzie do tzw. lokalizacji modeli rozumienia języka naturalnego. Druga teza stanowi, iż tłumaczenie maszynowe musi zachować i odpowiednio przetłumaczyć (oraz wskazać położenie) nazwanych jednostek (ang. NERs), aby przetłumaczyć dane treningowe na potrzeby zrozumienia języka naturalnego. Ostatnia teza zakłada, że tworzenie wielu wariantów danych treningowych dla wirtualnych asystentów wirtualnych poprawia dokładność zrozumienia naturalnego języka.

Autor niniejszej rozprawy, również jasno podkreśla istotność i wagę podjętego problemu naukowego. W szczególności Pan mgr inż. Marcin Sowański podkreśla, iż aktualnie brakuje modeli tłumaczenia maszynowego, które byłyby dostosowane do inteligentnych wirtualnych asystentów. Ponadto, Autor rozprawy, przytacza argument, iż jakość istniejących modeli tłumaczenia maszynowego, które nie są dostosowane, jest niewystarczająca.

Podsumowując, w mojej ocenie, problem naukowy rozprawy jest trafnie i jasno sformułowany przez Pan mgr inż. Marcina Sowańskiego.

## **2) Czy autor rozwiązał postawiony problem? Czy użył do tego właściwych metod dowodząc, że posiadał umiejętności związane z metodyką i metodologią prowadzenia badań naukowych?**

W mojej ocenie, Pan mgr inż. Marcin Sowański w sposób odpowiedni rozwiązał problem, którego dotyczy rozprawa oraz użył do tego właściwych metod. W szczególności zagadnienie badawcze zostało podzielone na trzy hipotezy, których słuszność oraz sposób ich dowodzenia został szczegółowo opisany w rozdziałach 3, 4 oraz 5.

W przypadku pierwszej hipotezy (T1), autor słusznie podkreślił istotność posiadania wielojęzycznych zbiorów danych dla trenowania modeli tłumaczenia maszynowego. Autor rozprawy słusznie wskazał, iż istniejące zbiory danych NLU są angielskojęzyczne i mają ograniczone zastosowanie. Dlatego prace na zbiorze danych o nazwie Leyzer wydają się być jak najbardziej celowe. Dodatkowo jest to uzasadnione przez Autora rozprawy tym, że opracowany korpus pozwala na bardziej precyzyjną ocenę agentów dialogowych. Dowodząc drugiej hipotezy (T2), Autor rozprawy badał szczegóły transferu slotów w kontekście

tłumaczenia maszynowego dla inteligentnych wirtualnych asystentów (IVAs). Używając odpowiednio przygotowanych eksperymentów, Pan mgr inż. Marcin Sowański wykazał, iż proponowany mechanizm anotowania slotów i ich transfer pozwala uzyskać tłumaczenia wysokiej jakości. W celu zweryfikowania ostatniej hipotezy (T3), Autor rozprawy zbadał istotność generowania wielu poprawnych wariantów tłumaczeń. Poprzez zastosowanie tej techniki zaobserwowana została poprawa klasyfikacji intencji, zwłaszcza podczas tłumaczenia angielskiego na język polski.

Przy weryfikowaniu wszystkich hipotez użyte zostały właściwe narzędzia i metody. W szczególności Autor rozprawy użył popularnych miar jakości klasyfikacji czy oceny skuteczności systemów tłumaczenia maszynowego. Tym samym Autor dowiódł, że posiadał umiejętności związane z metodyką i metodologią prowadzenia badań. Niemniej, Autor rozprawy mógł porównać więcej (popularnych w literaturze [67]) metod/technik wykorzystywanych w procesie tłumaczenia maszynowego.

### **3) Czy tematyka rozprawy jest aktualna i dostatecznie ważna?**

Na aktualność i wagę tematyki niniejszej rozprawy wpływa kilka czynników. W szczególności, jak zauważył Autor rozprawy, pomysł wykorzystania modeli tłumaczenia maszynowego do adaptowania zestawów uczących dla agentów dialogowych jest opisany w literaturze, ale aktualnie można zauważyć brak takich modeli, które dostosowane byłyby do potrzeb budowania inteligentnych wirtualnych asystentów. Ponadto, niewystarczająca jest jakość tłumaczenia otrzymana w oparciu o modele, które nie są odpowiednio adaptowane. Jak wskazał Pan mgr inż. Marcin Sowański, tłumaczenia takie nie pozwalają przenieść anotacji używanych w zasobach NLU. W tym świetle, rozwiązania proponowane przez Autora rozprawy otwierają możliwość tańszego i łatwiejszego tworzenia agentów dialogowych dla nowych języków. Jednocześnie umożliwia to większej liczbie użytkowników na korzystanie z produktów opartych o sztuczną inteligencję obsługiwaną głosem.

### **4) Na czym polega oryginalny dorobek autora i jakie jest jego znaczenie poznawcze lub przydatność praktyczna dla nauki bądź techniki?**

Na oryginalność dorobku Pana mgr inż. Marcina Sowańskiego wpływa wiele elementów, które powstały w trakcie prac nad niniejszą rozprawą. W szczególności Autor zaproponował nowy zbiór danych o nazwie Leyzer, który rozwiązuje ograniczenia istniejących zbiorów z punktu widzenia różnorodności intencji i slotów. Na podkreślenie zasługuje fakt, iż Leyzer jest publicznie dostępny dla innych badaczy i może być wykorzystywany w niekomercyjnych



zastosowaniach. W mojej ocenie, istotne z punktu widzenia oryginalności dorobku autora, jest też technika adaptacji domenowej MT (przedstawiona w rozdziale drugim) oraz (przedstawiona w rozdziale trzecim) propozycja rozwiązania braku różnorodności w tłumaczeniach zwracanych przez modele MT. O przydatności praktycznej dla nauki i techniki świadczy wdrożenie przemysłowe, które zostało szeroko opisane w rozdziale 6 niniejszej rozprawy.

**5) Czy rozprawa świadczy o dostatecznej wiedzy autora, wiedzy o zaawansowanym poziomie, o charakterze podstawowym dla dziedziny nauk technicznych oraz o charakterze szczegółowym, odpowiadającej obszarowi prowadzonych badań naukowych?**

Nie mam wątpliwości, iż Autor posiada dużą wiedzę dot. zagadnień związanych z tematyką rozprawy. W szczególności, Pana mgr inż. Marcin Sowański wykazał się umiejętnościami projektowania rozwiązań informatycznych oraz techniczną wiedzą w zakresie budowania inteligentnych wirtualnych asystentów. Autor rozprawy wykazał się również umiejętnościami dotyczącymi prowadzenia badań naukowych. Wielokrotnie w przeprowadzonych eksperymentach przedstawione zostaje autorskie osiągnięcie na tle innych rozwiązań dostępnych w literaturze. W procesie tym wykorzystywane są znane i stosowane w środowisku wskaźniki jakości takie jak BLEU czy F1-score. Na podkreślenie zasługuje także istotny dorobek naukowy Pana mgr inż. Marcina Sowańskiego. W szczególności, w rozdziale 7 niniejszej rozprawy, Autor przedstawił listę powiązanych z tematem rozprawy publikacji oraz udzielony patent, który dotyczy systemów IVA.

**6) Czy Autor wykazał umiejętności poprawnego i przekonującego przedstawienia uzyskanych przez siebie wyników?**

Zrealizowane przez Autora rozprawy eksperymenty są zgodne z dobrymi praktykami w zakresie pomiaru skuteczności systemów klasyfikacji i detekcji. Autor w tym zakresie użył znanych metryk (np. BLEU czy F1-score), co pozwala odnieść się do innych rozwiązań dostępnych w literaturze. W szczególności Autor rozprawy demonstrując wyniki jasno pokazał, iż proponowany model MT pozwala uzyskać (na zestawie testowym) więcej punktu BLEU niż linia bazowa. Ponadto, demonstrując wyniki, Autor rozprawy klarownie zademonstrował przydatność proponowanej techniki transferu jednostek oraz podejścia opartego wielowariantowe tłumaczenie.

## 7) Układ i poprawność redakcyjna rozprawy.

Rozprawa składa się z ośmiu rozdziałów. W pierwszej części Autor nakreśla tło, problem badawczy oraz motywację dla swoich badań. W tym świetle określony zostaje cel badawczy niniejszej rozprawy. W kolejnych trzech rozdziałach Pan mgr inż. Marcin Sowański przedstawia autorskie propozycje, które dowodzą, że specjalizowane modele tłumaczenia maszynowego mogą usprawnić proces tworzenia asystentów dialogowych dla nowych języków. W dalszej części rozprawy Autor przedstawia szczegóły na temat wdrożenia oraz swoje główne osiągnięcia w postaci wystąpień, publikacji naukowych oraz patentu.

Rozprawa napisana jest w sposób profesjonalny. Niemniej, w tekście można wskazać kilka niedociągnięć natury edytorskiej lub nieścisłości:

- W tekście nawiązującym i opisującym Figure 11 opisane są elementy ("creating English grammar, creating grammars for Spanish and Polish, slot expansion and splitting data into train-, dev- and test sets), których nie widać na obrazie.
- Fig. 12. Czarne kropki nie mają numeracji (cyfry wydają się być przycięte).
- Figure 6 (który podobny jest do Fig.7), może mylnie sugerować, iż m2m100 jest modelem opartym o sieć rekurencyjną. W moim odczuciu, zacytowanie "Beyond English-Centric Multilingual Machine Translation" (Angela Fan et al.) i przedstawienie ich architektury lepiej poparłoby stwierdzenie „sequence-to-sequence (Seq2Seq) is the most popular architecture in NLU”.
- Nazwy korpusów językowych, w szczególności różnych wersji WTM, mogłyby być lepiej opisane i opatrzone odpowiednią referencją do źródła.

## 8) Jakie są wady i słabe strony rozprawy?

Rolą recenzenta jest zauważenie ewentualnych niedociągnięć i mankamentów przedstawianej pracy oraz zgłoszenie uwag, które mogą być pomocne i przydatne w dalszych pracach. Dlatego w mojej ocenie:

- Wyniki otrzymane w rozdziale 4 mogłyby być bardziej rozbudowane, aby uwypuklić wpływ pewnych parametrów (np. chociażby wielkości modelu, czasu uczenia) na jakość otrzymywanych wyników (np. wartość BLEU).

- Eksperymenty w rozdziale 3 i 4 uwzględniają wyłącznie porównanie do modelu bazowego. Inne modele (poza m2m100) wykorzystywane w tłumaczeniu maszynowym wydają się być pominięte w eksperymentach i analizach (np. facebook/wmt21\*, czy facebook/mbart-large\*).
- W literaturze można wskazać metody, które wykorzystują różne podejścia do rozwiązywania problemów związanych z IC i SL. W szczególności wspomniana przez Autora rozprawa analiza [67], która dotyczy WMT22, zawiera całe spektrum różnych podejść związanych pośrednio lub bezpośrednio z tłumaczeniem maszynowym. Zastanawiające jest to, czy któreś z technik mogłyby być wykorzystane w przyszłych badaniach Autora rozprawy. Chociażby „back-translation” pojawia się wielokrotnie w [67].
- W rozdziale 1.1.4 Autor rozprawy mógł bardziej szczegółowo ocenić przewidywany wpływ rozwoju dużych modeli językowych na zagadnienia poruszane w pracy. We wrześniu 2023 roku OpenAI ogłosiło możliwość interakcji z modelem poprzez głos, co czyni go inteligentnym asystentem głosowym. Pytanie jak znaczący wpływ ma to na aktualną architekturę IVA oraz na rozwiązania proponowane w niniejszej rozprawie.

Pomimo przedstawionych powyżej uwag, rozprawa Pana mgra inż. Marcina Sowańskiego napisana jest w sposób profesjonalny i posiada wiele mocnych stron. Niewątpliwie recenzowana praca dowodzi dużej wiedzy Autora w tematach dotyczących zagadnień związanych z przetwarzaniem języka naturalnego.

## 9) Wniosek

Biorąc pod uwagę przedstawioną przez Doktoranta rozprawę stwierdzam, że spełnia ona wymagania stawiane rozprawom doktorskim przez obowiązujące przepisy i wnioskuję o dopuszczenie jej do publicznej obrony.

