Politechnika Wrocławska

Wrocław, 8.12.2025 r.

Recenzja pracy doktorskiej mgr inż. Karoliny Gabor-Siatkowskiej zatytułowanej: „Improving Therapeutic Spoken Dialogue Systems with Eye Tracking (Ulepszanie terapeutycznych głosowych systemów dialogowych przy użyciu okulografii)" przygotowanej pod kierunkiem: dra hab. inż. Artura Janickiego, prof. uczelni w Politechnice Warszawskiej; w dyscyplinie Informatyka Techniczna i Telekomunikacja.

Recenzja została sporządzona w związku z powołaniem przez przewodniczącego Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacji pana prof. dra hab. inż. Jarosława Arabasa na posiedzeniu w dniu 23.09.2025 roku do pełnienia funkcji recenzentki w postępowaniu o nadanie stopnia naukowego doktora nauk technicznych w dyscyplinie Informatyka Techniczna i Telekomunikacja pani mgr inż. Karolinie Gabor-Siatkowskiej.

Niniejsza recenzja ma za zadanie zgodnie z art. 186 ust. 1 ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce, w celu oceny, czy rozprawa doktorska stanowi oryginalne rozwiązanie problemu naukowego i wnosi istotny wkład w rozwój dyscypliny naukowej Informatyka Techniczna i Telekomunikacja.

W ramach przeprowadzonej recenzji zostaną ocenione następujące punkty:

1. **Tematyka pracy doktorskiej i jej wkład w dyscyplinę Informatyka Techniczna i Telekomunikacja.**

Ad. 1. Temat pracy doktorskiej brzmi: „Improving Therapeutic Spoken Dialogue Systems with Eye Tracking".

W ramach realizacji niniejszej rozprawy doktorskiej opracowano system dialogu głosowego oparty na LLM (Terabot) z wykorzystaniem ruchu gałek ocznych, przeznaczony do pomocniczej terapii w przebiegu schizofrenii. Opracowany system opiera się na badaniach pilotażowych przeprowadzonych w IPIN Warsaw, w których zidentyfikowano problemy z płynnością dialogu za pomocą danych z eyetrackingu pacjentów. Przeprowadzona przez Doktorantkę analiza eyetrackingu ujawniła obszary zainteresowania pacjentów podczas mówienia, umożliwiając multimodalną aktualizację z wykorzystaniem modeli GPT, które dostosowują rozmowy do ruchu gałek ocznych, poprawiając naturalność i radząc sobie z pauzami w mowie podczas oceny zaangażowania. Opracowany system wspiera cyfrowo zdrowie psychiczne poprzez integrację sygnałów

niewerbalnych z LLM, przewyższając systemy oparte wyłącznie na mowie w realistycznej interakcji z pacjentem.

W mojej ocenie – wybrany temat rozprawy o charakterze wybitnie interdyscyplinarnym, jest aktualny i stanowi istotny wkład dla dyscypliny naukowej: **Informatyka Techniczna i Telekomunikacja**.

2. Zagadnienia naukowe rozprawy – teza.

W pracy została postawiona następująca teza główna badawcza:

- „*The use of eye tracking can improve a therapeutic spoken dialogue system (Wykorzystanie śledzenia ruchu gałek ocznych może poprawić terapeutyczny system dialogu mówionego)*".

Oraz **4** cele badawcze (Research Objectives - RO):

- **RO1**: Investigating whether an eye tracker needs to be calibrated when used in a spoken dialogue system (Badanie, czy śledzenie wzroku wymaga kalibracji w przypadku korzystania z systemu dialogów mówionych).
- **RO2**: Analysis of areas of interest in the graphical interface of the dialogue system (Analiza obszarów zainteresowania w interfejsie graficznym systemu dialogowego).
- **RO3**: Enhancing the fluency of conversations with a dialogue system by the use of an eye tracker (Poprawa płynności rozmów dzięki systemowi dialogowemu wykorzystującemu śledzenie ruchu gałek ocznych).
- **RO4**: Enabling automatic user engagement assessment while interacting with the dialogue system (Włączanie automatycznej oceny zaangażowania użytkownika podczas interakcji z systemem dialogowym).

Biorąc pod uwagę osiągnięte przez Doktorantkę wyniki - postawiona w pracy teza oraz cele badawcze zostały udowodnione.

3. Struktura pracy

Ad. 3. Niniejsza rozprawa doktorska została w całości napisana w języku angielskim i składa się ze **120** stron (wliczając stronę tytułową, podziękowania, streszczenia - w języku polskim oraz angielskim, spis treści, bibliografię, załączniki, spis rysunków oraz tabel) i została podzielona na **8** ponumerowanych rozdziałów, załączniki (**A** i **B**), spis rysunków oraz tabel, nieponumerowaną i niewymienioną w spisie treści bibliografię. W pracy zawarto **11** tabel (wliczając **2** tabele w załącznikach) oraz **46** rysunków. Załączniki zawierają przykład dialogu - wybrany temat: złość (Załącznik **A**) oraz scenariusze dla aktorek i aktorów "Teatru PW" (Załącznik **B**). Bibliografia składa się ze **138** pozycji, w skład w których wchodzi **5** publikacji Autorki rozprawy.

Rozdział pierwszy zatytułowany "Introduction (Wprowadzenie)" zawiera ogólne wprowadzenie do tematu, informacje na temat dorobku Doktorantki oraz jej osiągnięć

naukowych. **Rozdział drugi**, który stanowi przegląd literatury, omawia wyzwania i znaczenie opieki psychiatrycznej na świecie. Doktorantka skupia się w nim także na agentach konwersacyjnych, wyjaśniając główne komponenty systemów dialogowych i ich zastosowanie w cyfrowej opiece zdrowotnej, ze szczególnym uwzględnieniem rozwiązań opartych na dużych modelach językowych (LLM) wspierających dobrostan psychiczny. W rozdziale tym przedstawiono także technologię śledzenia wzroku, omawiając parametry techniczne urządzeń i ich wykorzystanie w interakcji człowiek-komputer. W **rozdziale trzecim** opisuje stworzony przez Autorkę system Terabot dla pacjentów z zaburzeniami psychicznymi w języku polskim, wraz z przeprowadzonymi eksperymentami klinicznymi w Instytucie Psychiatrii i Neurologii. W rozdziale tym omówiono także architekturę systemu, interfejs użytkownika oraz eye-tracker Gazepoint GP3 użyty podczas badań. Zidentyfikowano kluczowe problemy dialogu oraz oceniono satysfakcję pacjentów z rozmów z Terabotem. **Rozdział czwarty** analizuje potrzebę integracji eye-trackera z systemem dialogowym, opisując problemy dialogu zidentyfikowane w Terabocie. Przedstawiono w nim także metodologię eksperymentu, wyniki badań oraz dyskusję, kończąc wnioskami na temat konieczności wykorzystania danych wzrokowych do poprawy naturalności rozmów z pacjentami. W **rozdziale piątym** Doktorantka analizuje wzorce wzroku pacjentów podczas rozmów z Terabotem, opisując problem, metodologię badań oraz analizę danych z eye-trackera zsynchronizowanych z wypowiedziami. Przedstawiono w nim także otrzymane wyniki badań oraz dyskusję na temat klinicznego zastosowania proponowanego rozwiązania w terapii zaburzeń psychicznych. **Rozdział szósty** zawiera opisy dotyczące projektowania systemu dialogowego zależnego od wzroku, opartego na LLM (OpenAI Assistant API), dla celów terapeutycznych. Omówiono analizę czasów oczekiwania pacjentów, wykorzystanie danych eye-tracking w czasie rzeczywistym, testy offline na danych pacjentów oraz testy real-time z symulowanymi pacjentami i aktorami Teatru PW. Przedstawiono w nim wyniki analizy wydajności nowego systemu, w tym aktywację reakcji na eye-tracking, ocenę rozmów oraz ankietę satysfakcji uczestników. **Rozdział siódmy** przedstawia automatyczną ocenę zaangażowania pacjenta w dialogi terapeutyczne. Omówiono analizę notatek asystentów klinicznych dotyczących zaangażowania pacjentów, zaproponowano metodę automatycznej oceny wykorzystującą dane eye-tracking oraz przedstawiono wyniki badań i wnioski dotyczące tego rozdziału. **Rozdział ósmy** zawiera podsumowanie i dyskusję wyników badań. Autorka omówiła w nim ograniczenia pracy oraz kierunki dalszych badań. Ponadto praca zawiera załączniki A i B, spis tabel, spis rysunków oraz bibliografię składającą się ze **138** pozycji.

4. Uwagi redakcyjne, krytyczne oraz pytania do pracy.

Ad. 4.

- Uwagi krytyczne:
    - Str. 1.:
        - Z rozdziału do podrozdziału - bez wprowadzenia.
        - Rysunek 1 - lekko rozmazana czcionka, słaba jakość.
    - Str. 2.:
        - Rysunek 2 - j.w.
    - Str. 11.:

        - Z rozdziału do podrozdziału - bez wprowadzenia.
    - Str. 12.:
        - Rys. 3. - j.w.
    - Str. 18.:
        - Rys. 5. - j.w.
    - Str. 20.:
        - Rys. 6 - j.w.
    - Str. 27.:
        - Rys. 14. - j.w.
    - Str. 34.:
        - Z rozdziału do podrozdziału - bez wprowadzenia.
    - Str. 46.:
        - Rys. 38. - j.w.
    - Str. 47.:
        - Z rozdziału do podrozdziału - bez wprowadzenia.
    - Str. 48.:
        - Opis eksperymentu badawczego byłby bardziej czytelny w formie tabeli oraz/lub wykresu.
    - Str. 54.:
        - Z rozdziału do podrozdziału - bez wprowadzenia.
    - Str. 61.:
        - Z rozdziału do podrozdziału - bez wprowadzenia.
    - Mała próba badawcza: Analizy eye-tracking na danych zaledwie 9 pacjentów ze schizofrenią, co negatywnie wpływa na moc statystyczną.
    - Brak porównań ilościowych w pracy, gdzie nie pokazano przewagi multimodalnego Terabot nad innymi systemami dialogowymi (np. ELIZA, Woebot) w metrykach jak BLEU, ROUGE czy engagement score. Uzyskane wyniki offline nie znajdują statystycznego potwierdzenia w rozprawie.

- Pozytywne aspekty pracy:
    - Rozprawa wyróżnia się pięknym, poprawnym językiem angielskim oraz wyjątkową dbałością o szczegóły edytorskie.
    - Cała praca jest niezwykle precyzyjnie dopracowana, a nawet na rysunkach z dialogami w języku polskim konsekwentnie podano wyraźne tłumaczenia na język angielski, co świadczy o profesjonalizmie i spójności całej pracy.
    - Autorka opracowała innowacyjne połączenie eye-tracking z LLM (GPT-4o-mini + RAG) w terapii schizofrenii, gdzie zgodnie z moim stanem wiedzy, jest to pierwszy system reagujący na spojrzenie pacjenta w czasie rzeczywistym.
    - Autorski system Terabot był testowany klinicznie i wdrażany testowo w Instytucie Psychiatrii i Neurologii (IPiN) Warszawa, co znacząco wzmacnia wiarygodność proponowanych rozwiązań.

- Pytania do pracy:
    - Jakie są, Pani zdaniem, główne ograniczenia przeprowadzonych badań i, które z nich uważa Pani za najbardziej krytyczne?

- Jakie mechanizmy kontroli bezpieczeństwa wypowiedzi systemu opartego na LLM zostały zastosowane lub byłyby konieczne w warunkach rzeczywistej terapii?
- Czy uzyskane przez Panią wyniki na grupie pacjentów ze schizofrenią można przenieść na inne grupy kliniczne dotknięte zaburzeniami psychicznymi?

5. Podsumowanie.

Ad. 5. Niniejsza rozprawa została niezwykle starannie zredagowana i estetycznie opracowana, bez istotnych błędów redakcyjnych. Ogólna dbałość o detale znacząco podnosi wartość pracy. Recenzowanie tej pracy było prawdziwą przyjemnością, gdyż napisana ona była bardzo dobrym językiem angielskim, starannie, a tematyka jest niezwykle ciekawa i aktualna.

Krytyczne uwagi mają charakter głównie kosmetyczny, natomiast zadane pytania mają charakter pogłębiający i dociekliwy. Z trudem byłam w stanie znaleźć niedociągnięcia czy negatywy w tej pracy.

Na uwagę zasługuje także całokształt osiągnięć i działań naukowych Doktorantki. Cenne jest szczegółowe opisanie przez Autorkę swoich publikacji, osiągnięć oraz udziału w projektach krajowych i międzynarodowych. Rzadko spotyka się taki opis w pracach doktorskich, a przecież to właśnie taki pełny autoreferat pozwala na kompleksową ocenę wkładu w dyscyplinę i świadczy o aktywnym udziale w życiu naukowym Doktorantki.

Autorka jest autorką/współautką **13** publikacji naukowych (Google Scholar) i posiada następujący indeks Hirscha: **4** (Google Scholar), **47** cytowań. W **10** z **13** publikacji Doktorantka jest pierwszą (wiodącą) autorką.

Doktorantka opracowała rozwiązania problemów dialogu (przerywanie, pauzy) wykorzystując dane wzrokowe z GPT-4o-mini + RAG, potwierdzone testami real-time i automatyczną oceną zaangażowania. Posiada w swoim dorobku aż **13** publikacji naukowych, brała udział w projektach, otrzymywała nagrody Rektora, co świadczy o wpływie na rozwój dyscypliny jaką jest **Informatyka Techniczna i Telekomunikacja**.

W mojej ocenie rozprawa doktorska **mgr inż. Karoliny Gabor-Siatkowskiej** zawiera bardzo wartościowe wyniki badań, wyróżniając się potencjalną wartością praktyczną - potencjałem na wdrożenie klinicznie opracowanego systemem Terabot; a moja ocena jest jak najbardziej **pozytywna** - praca prezentuje cenne osiągnięcia naukowe w dyscyplinie **Informatyka Techniczna i Telekomunikacja**, spełniając wszystkie wymogi art. 186 ust. 1 ustawy z 20 lipca 2018 r. - Prawo o szkolnictwie wyższym i nauce, i nadaje się do publicznej obrony.

Ze względu na osiągnięte i zaprezentowane w pracy wyniki, a także na bardzo dobry dorobek naukowy Doktorantki, jej bardzo dużą (na tym etapie kariery naukowej) aktywność, wnioskuję o **wyróżnienie rozprawy mgr inż. Karoliny Gabor-Siatkowskiej.**

Dr hab. inż. Aleksandra Kawala-Sterniuk, prof. uczelni
Katedra Sztucznej Inteligencji
Wydział Informatyki i Telekomunikacji
Politechnika Wrocławska
aleksandra.kawala-sterniuk@pwr.edu.pl