

**Ocena osiągnięcia będącego podstawą dla postępowania habilitacyjnego
pt.: „Ocena skuteczności otwartych modeli językowych i multimodalnych w
analizie deformacji kręgosłupa: integracja danych tekstowych i obrazowych na
przykładzie skolioz”
oraz dorobku naukowego, dydaktycznego i organizacyjnego dr. n. med.
Bartosza Polisa**

1. Podstawowe dane o Kandydacie

Dr n. med. Bartosz Polis ukończył studia na Wydziale Lekarskim Uniwersytetu Medycznego w Łodzi w 2005 roku. Stopień doktora nauk medycznych z wyróżnieniem uzyskał w 2008 roku w Instytucie Centrum Zdrowia Matki Polki w Łodzi na podstawie rozprawy pt. *„Wczesna wertebraktomia w połączeniu z zaopatrzeniem wady dysraficznej jako metoda zapobiegania skrzywieniom kręgosłupa”*. Od 2006 roku zawodowo związany jest z Kliniką Neurochirurgii ICZMP w Łodzi, a w 2013 roku uzyskał tytuł specjalisty neurochirurga. W 2010 r. ukończył również Kolegium Zarządzania i Finansów Szkoły Głównej Handlowej w Warszawie. W 2024 roku objął stanowisko adiunkta w macierzystej jednostce.

Dorobek kliniczny Kandydata uzupełniają staże i szkolenia, w tym dwumiesięczny staż szkoleniowy w Berlinie (2010), miesięczny staż w Paryżu (2011) oraz staże w Polsce i za granicą dotyczące leczenia i korekcji deformacji kręgosłupa. Aktywność ta pozostaje spójna z głównym kierunkiem jego zainteresowań klinicznych, obejmujących neurochirurgię dziecięcą oraz problematykę skolioz i stabilizacji kręgosłupa, a w ostatnich latach - zastosowania sztucznej inteligencji w diagnostyce obrazowej.

Analiza ścieżki zawodowej wskazuje na konsekwentny rozwój kliniczny Kandydata, jednakże w kontekście wymogów stawianych przed habilitantem, należy zwrócić uwagę na stosunkowo późne podjęcie intensywnej aktywności publikacyjnej ukierunkowanej na obecne postępowanie (lata 2023–2025), co zostanie omówione w dalszej części recenzji.

2. Ocena osiągnięć naukowo-badawczych Habilitanta

Dorobek naukowy kandydata charakteryzuje się wyraźną cezurą czasową oraz intensyfikacją aktywności publikacyjnej w latach 2023–2025 (okres powstawania cyklu habilitacyjnego). Wg. danych dostarczonych w dokumentacji dorobek naukowy Habilitanta obejmuje 24 artykuły oryginalne, 2 prace poglądowe i 6 opisów przypadków. Sumaryczny współczynnik oddziaływania publikacji Habilitanta wynosi 68,327 (punktacja KBN/MEiN=2205). Liczba cytowań (wg bazy ISI Web of Science) wynosi 155 (bez autocytowań), zaś indeks Hirscha = 6. Kwerenda bazy Scopus wykazała 137 cytowań (bez autocytowań) oraz indeks Hirscha = 6. Habilitant jest również współautorem 1 rozdziału w podręczniku krajowym. Ponadto Dr Polis brał udział w konferencjach krajowych i zagranicznych, jest współautorem 20 doniesień zjazdowych, z których 12 przedstawiono na konferencjach międzynarodowych.

Publikacje przed uzyskaniem stopnia doktora (do 2008 r.)

W tym okresie aktywność publikacyjna kandydata była niewielka i koncentrowała się na tematyce klinicznej związanej z neurochirurgią dziecięcą z jedną znaczącą pracą oryginalną o zasięgu międzynarodowym w tym okresie (*Child's Nerv Sys*, 2006). Głównym osiągnięciem badawczym była praca nad rozprawą doktorską pt. *„Wczesna wertebrektomia w połączeniu z zaopatrzeniem wady dysraficznej jako metoda zapobiegania skrzywieniom kręgosłupa”*, obroniona w 2008 roku. Brak również istotnych prac poglądowych w tym okresie.

Publikacje po uzyskaniu stopnia doktora (od 2009 r.)

Jest to główny okres aktywności naukowej ze szczególnym nasileniem w ostatnich dwóch latach (cykl habilitacyjny). Poza pracami zawartymi w cyklu habilitacyjnym prace oryginalne obejmowały publikacje dotyczące neurochirurgii dziecięcej i patologii kręgosłupa, m.in. w *Child's Nervous System*, *Oncotarget* oraz *Advances in Clinical and Experimental Medicine*. W dorobku widoczny jest deficyt prac stricte przeglądowych (review papers) w wiodących czasopismach. Prace z cyklu habilitacyjnego, mimo że testują modele AI, mają strukturę prac oryginalnych (badawczych), a nie przeglądowych, choć ich materiałem badawczym są istniejące algorytmy.

3. Ocena osiągnięcia naukowego

Jako osiągnięcie naukowe, w rozumieniu art. 219 ust. 1 pkt 2 Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce, Kandydat przedstawił cykl publikacji powiązanych tematycznie pod tytułem: „Ocena skuteczności otwartych modeli językowych i multimodalnych w analizie deformacji kręgosłupa: integracja danych tekstowych i obrazowych na przykładzie skolioz”.

Na cykl składa się 6 prac oryginalnych opublikowanych w latach 2023–2025 w czasopismach takich jak *Diagnostics*, *Journal of Personalized Medicine*, *Journal of Clinical Medicine* oraz *Applied Sciences*. Sumaryczny Impact Factor cyklu wynosi 17.50, a punktacja MNiSW to 590 pkt.

Dokonując merytorycznej oceny przedstawionego osiągnięcia, z żalem stwierdzam szereg istotnych uchybień, które w moim przekonaniu dyskwalifikują ten dorobek jako podstawę do nadania stopnia doktora habilitowanego. Głównym zarzutem wobec przedłożonego cyklu jest jego odtwórczy charakter. Prace Habilitanta polegają w przeważającej mierze na testowaniu komercyjnie dostępnych, zamkniętych modeli (tzw. *black-box*), takich jak ChatGPT, Bing czy modele CLIP, w zadaniu klasyfikacji obrazów lub tekstów medycznych metodą *zero-shot* (bez dedykowanego treningu). Zastosowanie zamkniętych modeli komercyjnych (np. ChatGPT, Bing) metodą *zero-shot* do tych celów budzi poważne zastrzeżenia metodologiczne. Choć narzędzia te imponują wszechstronnością, opieranie na nich rygorystycznej ewaluacji badawczej jest obarczone krytycznymi wadami. Kluczowym błędem jest zakładanie prawdziwości trybu "zero-shot" w obliczu ryzyka zanieczyszczenia danych (tzw. data contamination). Zbiory treningowe modeli własnościowych są ściśle utajnione. Użytkownik nie ma możliwości weryfikacji, czy klasyfikowane obrazy testowe nie stanowiły już części ich ogromnych danych uczących. Istnieje wysokie prawdopodobieństwo, że model po prostu "zapamiętał" konkretne obrazy, co sztucznie zawyża jego skuteczność i czyni ocenę zerowej próby bezwartościową (vide Bubeck, S. et al. (2023). Sparks of Artificial General Intelligence: Early experiments with GPT-4. ArXiv.) Kolejną nieprawidłowością w stosowaniu w/w modeli w klasycznym wnioskowaniu naukowym jest całkowity brak odtwarzalności (reproducibility). Zamknięte interfejsy API podlegają ciągłym, niedokumentowanym aktualizacjom. Zjawisko tzw. model drift sprawia, że wyniki klasyfikacji tego samego obrazu mogą drastycznie zmienić się w ciągu kilku tygodni, całkowicie uniemożliwiając wiarygodne badania porównawcze (vide Chen, L. et al.

(2023). How is ChatGPT's behavior changing over time? ArXiv.) Wreszcie, architektura czarnej skrzynki (black-box) blokuje audyt procesu decyzyjnego. Badacz nie jest w stanie sprawdzić, czy model poprawnie rozpoznał cechy wizualne, czy też uległ halucynacji opartej na ukrytych uprzedzeniach (bias) (vide Zhou, K. et al. (2023). Don't Make Your LLM an Evaluation Benchmark: On the Data Contamination Problem. ArXiv.) Do miarodajnych badań zero-shot powinno się stosować wyłącznie przejrzyste modele z udokumentowanymi zbiorami treningowymi. Kolejnym problemem jaki dostrzegam jest fakt, na który Kandydat wskazuje w autoreferacie - celem prowadzonych badań było „określenie skuteczności nowoczesnych modeli AI”. Tak sformułowany cel badawczy sprowadza rolę naukowca do testera gotowych rozwiązań technologicznych, a nie twórcy nowej wiedzy medycznej czy algorytmicznej. W cyklu brakuje próby stworzenia własnych modeli lub ich douczania (*fine-tuning*) na danych klinicznych czy dogłębnej analizy przyczyn błędów (tzw. halucynacji) w kontekście patofizjologii skolioz. Otrzymane wyniki (np. stwierdzenie, że modele językowe słabo radzą sobie z oceną kąta Cobba na podstawie obrazu) są trywialne z punktu widzenia informatyki medycznej i nie wnoszą istotnej wartości do neurochirurgii. Co więcej, stosowanie modeli ogólnego przeznaczenia do diagnostyki radiologicznej bez odpowiedniego preprocessing'u i walidacji na dużych, zewnętrznych zbiorach danych, rodzi ryzyko wyciągnięcia fałszywych wniosków klinicznych. W pracach brakuje rygorystycznej analizy ryzyka klinicznego, co jest standardem w badaniach nad AI w medycynie. Ponad to, analizując strukturę cyklu, trudno oprzeć się wrażeniu sztucznego podziału wyników badawczych. Publikacje są do siebie zbliżone metodycznie – zmieniają się jedynie testowane modele (raz CLIP, raz BiomedCLIP, raz ChatGPT) lub badany parametr (kąta Cobba vs. implanty). Taka strategia publikacyjna, choć pozwala na uzyskanie wysokiej punktacji bibliometrycznej w krótkim czasie (2023–2025), nie świadczy o głębi problemu badawczego, lecz o powierzchownej eksploracji tematu.

4. Ocena pozostałej aktywności naukowej

Dorobek naukowy Kandydata spoza osiągnięcia habilitacyjnego jest obiecujący, lecz nie wybitny. Wyróżnia się wieloletnia praca kliniczna w ICZMP, jednak aktywność

badawcza nabrała tempa dopiero w ostatnich dwóch latach. Aktywność naukowa Habilitanta po uzyskaniu stopnia doktora, realizowana poza głównym cyklem habilitacyjnym obejmuje szerokie spektrum badań z zakresu neurochirurgii dziecięcej, neuroonkologii oraz zaawansowanej chirurgii kręgosłupa. Jednym z nurtów jego prac są badania nad patofizjologią i leczeniem wodogłowia. Habilitant analizował odległe efekty psychologiczne leczenia operacyjnego tego schorzenia metodami neuroendoskopowymi i zastawkowymi. Pełnił funkcję głównego badacza grantu NCN, w ramach którego dokonano oceny immunologicznej płynu mózgowo-rdzeniowego u pacjentów z powikłaniami infekcyjnymi po implantacji układów drenażujących. Badał również zastosowanie rezonansu magnetycznego z kontrastem fazowym do oceny przepływu płynu mózgowo-rdzeniowego. Ponadto angażował się w projekty dotyczące terapii wewnątrzmacicznej, oceniając skuteczność przecieków komorowo-owodniowych w leczeniu ciężkiej wentrykulomegalii płodu. W obszarze neuroonkologii brał udział w nowatorskich badaniach nad zastosowaniem spektroskopii Ramana do śródoperacyjnej diagnostyki guzów mózgu ze szczególnym uwzględnieniem rdzeniaka zarodkowego. Główny nurt zainteresowań naukowo-klinicznych Habilitanta obejmował metody operacyjnego leczenia ciężkich skolioz u dzieci obciążonych rzadkimi chorobami genetycznymi, takimi jak zespół Costello czy zespół Wolf-Hirschhorn. Zajmował się także wykorzystaniem nowoczesnych technik stabilizacji kolumny kręgosłupa, m.in. systemów DTRAX oraz implantów rozprężalnych Spine-Jack w złamaniach kompresyjnych. Habilitant uczestniczył również w eksperymentalnych badaniach z zakresu inżynierii tkankowej, weryfikując zastosowanie kolagenowych rusztowań zasiedlonych embrionalnymi komórkami nerwowymi do implantacji śródmózgowej. Brakuje w dorobku Kandydata kierowania prestiżowymi grantami badawczymi (NCN, NCBiR) oraz udokumentowanej, istotnej współpracy międzynarodowej wykraczającej poza standardowe współautorstwo.

5. Ocena aktywności dydaktycznej i organizacyjnej Habilitanta

Aktywność dydaktyczna i organizacyjna jest typowa dla pracownika instytutu badawczego. W zakresie dydaktyczno-popularyzatorskim jest ona realizowana przede

wszystkim w formie wykładów edukacyjnych oraz współorganizacji spotkań i konferencji. Habilitant od wielu lat jest związany ze środowiskiem Polskiego Towarzystwa Neurochirurgów (m.in. konferencje Sekcji Neurochirurgii Dziecięcej), gdzie pełnił role współorganizatora, moderatora i wykładowcy. Ponadto wygłaszał wykłady na krajowych i międzynarodowych konferencjach (m.in. Kopenhaga, Rzym, Queensland, Buenos Aires, Dubaj), dotyczące zagadnień wodogłowia, neurochirurgii dziecięcej oraz deformacji kręgosłupa.

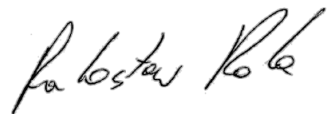
Na polu organizacyjnym Habilitant pełnił funkcję organizatora spotkania Okręgu Centralno-Wschodniego Polskiego Towarzystwa Chirurgii Kręgosłupa w 2014 r. Od wielu lat z powodzeniem pełni też funkcję współorganizatora i moderatora podczas cyklicznych spotkań Oddziału Łódzkiego PTNCh oraz podczas konferencji Sekcji Neurochirurgii Dziecięcej PTNCh. Ponadto dr Polis jest członkiem prestiżowych medycznych towarzystw naukowych – należy do Polskiego Towarzystwa Neurochirurgów, a od 2010 r. jest także zrzeszony w Polskim Towarzystwie Chirurgii Kręgosłupa, North American Spine Society (NASS) oraz AO Spine.

6. Podsumowanie i wniosek końcowy

Przedstawiony do recenzji dorobek naukowy jest obiecujący i wskazuje na stopniową ewolucję kandydata w kierunku samodzielnego pracownika naukowego, ale zgodnie z art. 219 ust. 1 pkt 2 Ustawy Prawo o szkolnictwie wyższym i nauce, osiągnięcie habilitacyjne musi stanowić znaczny wkład w rozwój określonej dyscypliny. W mojej ocenie przedłożony cykl publikacji pt. *„Ocena skuteczności otwartych modeli językowych i multimodalnych...”* wymogu tego nie spełnia. Prace te wnoszą jedynie częściowy wkład w istotnie szersze zagadnienie, a ich wartość poznawcza jest ograniczona przez zastosowanie metodologii opartej na nietrwałych narzędziach komercyjnych. Kandydat wykazał się sprawnością w obsłudze nowoczesnych narzędzi AI, jednak nie przełożyło się to na rozwiązanie istotnego problemu naukowego w neurochirurgii ani na stworzenie trwałego modelu diagnostycznego. W związku z powyższym, stwierdzam, że dr n. med. Bartosz Polis nie spełnia ustawowych wymagań stawianych kandydatom do stopnia doktora habilitowanego.

Przedkładam Wysokiej Radzie wniosek o odmowę nadania dr. n. med. Bartoszowi Polisowi stopnia doktora habilitowanego w dziedzinie nauk medycznych i nauk o zdrowiu.

Z wyrazami szacunku,

Handwritten signature in black ink, appearing to read "Fabian Kozłowski".