

Warszawa, 17 sierpnia 2026

prof. dr hab. inż. Tomasz Trzeciński,
Instytut Informatyki Politechniki Warszawskiej

RECENZJA

**osiągnięcia naukowego i dorobku naukowego w ramach postępowania
o nadanie Panu dr. Markowi Kubisowi stopnia doktora habilitowanego
w dziedzinie nauk inżyneryjno-technicznych,
w dyscyplinie informatyka techniczna i telekomunikacja**

Recenzja została przygotowana dla Rady Naukowej Instytutu Podstaw Informatyki Polskiej Akademii Nauk na podstawie dokumentacji przedstawionej w postępowaniu habilitacyjnym dr. Marka Kubisa.

1. Wprowadzenie

Niniejsza recenzja dotyczy osiągnięcia naukowego oraz całokształtu aktywności naukowej dr. Marka Kubisa, przedstawionych w postępowaniu o nadanie stopnia doktora habilitowanego. Podstawą oceny są kryteria określone w art. 219 ust. 1 ustawy z dnia 20 lipca 2018 r. *Prawo o szkolnictwie wyższym i nauce*, w szczególności wymóg posiadania stopnia doktora, posiadania w dorobku osiągnięć naukowych stanowiących znaczny wkład w rozwój określonej dyscypliny oraz wykazywania się istotną aktywnością naukową realizowaną w więcej niż jednej uczelni lub instytucji naukowej.

Dr Marek Kubis uzyskał w 2013 roku stopień doktora nauk matematycznych w zakresie informatyki, w specjalności lingwistyka komputerowa, na Wydziale Matematyki i Informatyki Uniwersytetu im. Adama Mickiewicza w Poznaniu. Rozprawa doktorska nosiła tytuł *Język zapytań dla leksykalnej bazy danych typu WordNet*. Od 1 października 2013 roku jest zatrudniony jako adiunkt na Wydziale Matematyki i Informatyki UAM.

Jako osiągnięcie, o którym mowa w art. 219 ust. 1 pkt 2 ustawy, Habilitant przedstawił **cykl sześciu powiązanych tematycznie artykułów naukowych z zakresu przetwarzania języka naturalnego, poświęcony metodom automatycznej analizy konwersacji**. Cykl obejmuje dwa uzupełniające się nurty: analizę sieci konwersacji w tekstach literackich oraz analizę wypowiedzi w systemach konwersacyjnych. Taki podział dobrze oddaje ewolucję zainteresowań badawczych Habilitanta: od analizy struktury dialogów na poziomie całych dzieł i korpusów, poprzez identyfikację uczestników konwersacji, aż po modelowanie pojedynczych wypowiedzi w systemach mowy i systemach multimodalnych.

2. Ocena głównego osiągnięcia naukowego

Przedłożony cykl obejmuje następujące publikacje:

- M. Kubis, *Quantitative analysis of character networks in Polish 19th- and 20th-century novels*, *Digital Scholarship in the Humanities*, 2021;
- M. Kubis, *Geometric Deep Learning Models for Linking Character Names in Novels*, Joint SI-GHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature, 2020;
- M. Kubis, M. Méloux, P. Skórzewski, M. Lewandowski, G. Jho, H. Park, *Part-of-Speech Models Compression Methods for on-Device Grapheme-to-Phoneme Conversion*, ICASSP 2022;
- M. Kubis, P. Skórzewski, M. Sowański, T. Ziętkiewicz, *Back Transcription as a Method for Evaluating Robustness of Natural Language Understanding Models to Speech Recognition Errors*, EMNLP 2023;

- M. Kubis, P. Skórzewski, M. Sowański, T. Ziętkiewicz, *Center for Artificial Intelligence Challenge on Conversational AI Correctness*, FedCSIS 2023;
- I. Christop, M. Czyżnikiewicz, P. Skórzewski, Ł. Bondaruk, J. Kubiak, M. Lewandowski, M. Kubis, *A Benchmark for Audio Reasoning Capabilities of Multimodal Large Language Models*, EACL 2026.

Podstawowym walorem cyklu jest jego spójność problemowa przy jednoczesnym zróżnicowaniu metod. Wspólnym mianownikiem wszystkich prac jest automatyczne wydobywanie, reprezentowanie i ocena informacji niezbędnej do analizy konwersacji. W pracach literaturoznawczych informacją tą są relacje między postaciami i odniesienia nazw własnych; w systemach konwersacyjnych są nią cechy językowe wypowiedzi, reprezentacje semantyczne, błędy rozpoznawania mowy oraz właściwości dźwięku. Z naukowego punktu widzenia cykl przechodzi zatem od analizy makrostruktury konwersacji do analizy jej elementarnych składników.

2.1. Metody analizy sieci konwersacji w tekstach literackich

Pierwsze dwa artykuły stanowią samodzielne prace Habilitanta i dotyczą automatycznej analizy struktury konwersacji w literaturze. W pracy opublikowanej w *Digital Scholarship in the Humanities* Habilitant zbudował kompletny potok przetwarzania obejmujący m.in. segmentację tekstu, tokenizację, lematyzację, oznaczanie części mowy, rozpoznawanie jednostek nazewniczych, wykrywanie granic dialogu, analizę aktów mowy i ujednoznacznianie postaci. Na tej podstawie powstał anotowany korpus 907 polskich powieści wraz z indukowanymi sieciami konwersacji. Korpus został podzielony na część dziewiętnastowieczną i dwudziestowieczną, obejmujące odpowiednio 383 i 524 utwory.

Wartość tej pracy widzę w połączeniu trzech elementów, które rzadko występują równocześnie: znacznej skali danych, pełnej automatyzacji przetwarzania oraz ilościowej weryfikacji hipotez literaturoznawczych. Habilitant sformułował i zweryfikował hipotezy odnoszące się m.in. do liczby postaci, liczby konwersacji, liczby interlokutorów, wielkości społeczności oraz modularności sieci. Uzyskane wyniki pozwoliły m.in. wykazać różnice pomiędzy prozą dziewiętnasto- i dwudziestowieczną w zakresie liczby dialogów i liczby postaci uczestniczących w konwersacji. Jest to dobry przykład pracy z pogranicza informatyki i humanistyki cyfrowej, w której metody informatyczne nie są wyłącznie narzędziem pomocniczym, lecz umożliwiają postawienie i zweryfikowanie nowych hipotez badawczych.

Druga praca rozwija kluczowy element tego potoku, tj. ujednoznacznianie nazw własnych postaci. Habilitant reprezentuje tekst powieści jako graf odniesień, którego wierzchołkami są występujące w tekście nazwy, a krawędzie opisują związki leksykalne i znaczeniowe. Następnie stosuje grafowe sieci neuronowe do zadań wierzchołkowych i krawędziowych. Przeprowadzone eksperymenty pokazują, że wykorzystanie informacji strukturalnej zapisanej w grafie może poprawić jakość ujednoznaczniania nazw, a w wariantcie krawędziowym pozwala osiągnąć porównywalne wyniki przy mniejszej liczbie parametrów. Uważam tę pracę za ważny składnik osiągnięcia, ponieważ rozwiązuje ona jeden z trudniejszych technicznie problemów niezbędnych do automatycznej rekonstrukcji sieci konwersacji.

2.2. Metody analizy wypowiedzi w systemach konwersacyjnych

Kolejne cztery publikacje przenoszą ciężar badań z tekstów literackich na systemy dialogowe i mowę. W pracy z ICASSP 2022 Habilitant współtworzył metody kompresji modeli oznaczania części mowy używanych w procesie konwersji grafemów na fonemy realizowanym bezpośrednio na urządzeniu. Szczególnie istotnym wynikiem praktycznym jest wskazanie schematów kompresji pozwalających na około 90% redukcję rozmiaru modelu przy zachowaniu dokładności oznaczania na poziomie nieprowadzącym do istotnego pogorszenia zadania docelowego. Osobisty wkład Habilitanta obejmował opracowanie planu badań oraz przygotowanie modeli opartych na architekturze Transformer poddawanych następnie eksperymentom kompresji.

Za najbardziej dojrzałą metodologicznie część osiągnięcia uważam pracę z EMNLP 2023 dotyczącą *back transcription*. Jej celem jest ilościowa ocena wpływu błędów systemu rozpoznawania mowy na działanie modelu rozumienia języka naturalnego (NLU). Zaproponowana procedura obejmuje trzy etapy: wygenerowanie odwrotnej transkrypcji, pomiar zmian w reprezentacji znaczeniowej oraz analizę typów błędów ASR odpowiedzialnych za pogorszenie działania NLU. Habilitant wyróżnił różne kategorie zmian reprezentacji znaczeniowej, zaproponował odpowiadające im miary odporności i zbudował

model pozwalający wskazywać typy błędów rozpoznawania mowy o największym wpływie na działanie systemu.

W mojej ocenie najważniejszą cechą tej metody jest możliwość wykorzystania już istniejących zasobów treningowych NLU bez konieczności wykonywania dodatkowych nagrań. Podejście jest więc nie tylko poprawne metodologicznie, ale także atrakcyjne praktycznie. Habilitant był głównym autorem metody; zaproponował transkrypcję wsteczną jako mechanizm ewaluacji, zdefiniował kryteria odporności i miary, a także metodę identyfikacji błędów o największym wpływie na odporność modelu. Dodatkowym potwierdzeniem użyteczności rozwiązania jest włączenie jego implementacji do biblioteki *TextAttack*.

Praca FedCSIS 2023 stanowi naturalne rozwinięcie tego nurtu. Habilitant wraz ze współautorami przygotował otwarty wielojęzyczny zbiór danych i rozwiązania bazowe dla konkursu *Center for Artificial Intelligence Challenge on Conversational AI Correctness*. Jego osobisty wkład obejmował zdefiniowanie zadania, przygotowanie modeli NLU stanowiących rozwiązania odniesienia, opracowanie założeń konkursu i kryteriów ewaluacji. Tego rodzaju aktywność oceniam pozytywnie, gdyż pokazuje zdolność nie tylko do proponowania algorytmów, lecz także do projektowania procedur ewaluacyjnych użytecznych dla szerszej społeczności badawczej.

Ostatnia praca w cyklu, przyjęta na EACL 2026, dotyczy możliwości prowadzenia rozumowania na podstawie sygnału dźwiękowego przez multimodalne duże modele językowe. Autorzy przygotowali benchmark ART (*Audio Reasoning Tasks*) obejmujący dziewięć rodzajów zadań i 9000 przykładów, o łącznej długości ponad 30 godzin materiału dźwiękowego. Konstrukcja benchmarku celowo wyklucza zadania możliwe do rozwiązania przez prosty potok złożony z wyspecjalizowanego modułu i modelu tekstowego. Wyniki wskazują, że badane modele o otwartych wagach mają istotne trudności z zaproponowanymi zadaniami; w żadnym z nich dokładność nie przekroczyła 0.7, a w zadaniu *Sound Reasoning* żaden z modeli nie osiągnął 60% poprawnych odpowiedzi. Wkład Habilitanta obejmował sformułowanie problemu badawczego, zaplanowanie badań i wypracowanie założeń benchmarku.

2.3. Ocena spójności, oryginalności i znaczenia osiągnięcia

Przedstawiony cykl nie jest jednorodny technicznie: obejmuje analizę grafów, głębokie uczenie, kompresję modeli, analizę odporności systemów NLU oraz ewaluację multimodalnych modeli językowych. Ta różnorodność mogłaby być potencjalną słabością cyklu, jednak w dokumentacji Habilitant przekonująco pokazuje wspólną oś badawczą, którą jest automatyczna analiza konwersacji na różnych poziomach reprezentacji i w różnych modalnościach. Co ważne, kolejne prace nie są przypadkowym zbiorem publikacji, lecz odpowiadają na komplementarne pytania: kto uczestniczy w konwersacji, jak reprezentować i łączyć odniesienia do uczestników, jak analizować wypowiedź w systemie mowy, jak mierzyć odporność na błędy rozpoznawania oraz jak oceniać rozumowanie wymagające jednoczesnego użycia informacji tekstowej i dźwiękowej.

Za mocną stroną osiągnięcia uznaję również wyraźnie udokumentowany wkład osobisty Habilitanta. Dwie pierwsze prace są jednoautorskie, a w pracach wieloautorskich dokumentacja wskazuje konkretne elementy koncepcyjne i implementacyjne przypisane dr. Kubisowi. Szczególnie silny wkład widoczny jest w metodzie *back transcription* oraz w projektowaniu zadań i procedur ewaluacyjnych.

Warto jednocześnie odnotować trzy uwagi o charakterze dyskusyjnym. Po pierwsze, praca dotycząca kompresji modeli części mowy jest nieco słabiej związana z analizą konwersacji niż pozostałe elementy cyklu; jej związek z osiągnięciem wynika przede wszystkim z wykorzystania modeli językowych w urządzeniowych systemach mowy. Po drugie, część wyników została opublikowana w materiałach warsztatowych lub konferencyjnych o zróżnicowanym profilu, co utrudnia jednolitą ocenę ich oddziaływania. Z drugiej strony cykl zawiera także publikacje w uznanych międzynarodowych miejscach, w tym EMNLP, ICASSP i EACL, oraz artykuł czasopismowy w *Digital Scholarship in the Humanities*. Po trzecie, najnowsza praca dotycząca benchmarku ART jest bardzo świeża, dlatego nie sposób jeszcze ocenić jej długoterminowego wpływu na społeczność. Nie zmienia to jednak pozytywnej oceny oryginalności samego pomysłu i jakości przeprowadzonej ewaluacji.

Podsumowując, **oceniam główne osiągnięcie naukowe dr. Marka Kubisa jako znaczące i stanowiące istotny wkład w rozwój dyscypliny informatyka techniczna i telekomunikacja**, w szczególności w zakresie przetwarzania języka naturalnego, systemów konwersacyjnych oraz metod ewaluacji modeli językowych i multimodalnych.

3. Ocena istotnej aktywności naukowej dr. Marka Kubisa

Aktywność naukowa Habilitanta wykracza poza sześć publikacji przedstawionych jako główne osiągnięcie. W autoreferacie wykazuje on autorstwo lub współautorstwo 35 recenzowanych prac naukowych dotyczących przetwarzania języka naturalnego, z czego 28 ukazało się po obronie doktoratu. Habilitant zaznacza przy tym, że po pominięciu rozszerzonych wersji wcześniej opublikowanych prac liczby te wynoszą odpowiednio 28 i 21. Takie rozróżnienie świadczy o rzetelnym przedstawieniu własnego dorobku, natomiast dorobek Habilitanta pozostaje w dużej mierze niezauważony na arenie międzynarodowej, o czym może świadczyć relatywnie niska wartość współczynnika Hirsha (10 za Google Scholar z dnia sporządzania recenzji).

Istotnym elementem kariery naukowej dr. Kubisa jest łączenie badań akademickich z badaniami przemysłowymi. W latach 2013-2026 prowadził prace badawcze we współpracy z Samsung Electronics, dotyczące przede wszystkim systemów konwersacyjnych i rozpoznawania mowy. W 2017 roku odbył tygodniowy pobyt badawczy w siedzibie Samsung Electronics w Suwon w Korei. W latach 2020-2024 był promotorem pomocniczym w przewodzie doktorskim dr. Tomasza Ziętkiewicza prowadzonym w ramach programu doktoratów wdrożeniowych we współpracy UAM i Samsung Electronics Polska.

Słabym punktem wniosku jest, niestety, niska aktywność naukowa poza jednostką macierzystą, na więcej niż jednej uczelni lub instytucji naukowej, w szczególności zagranicznej. Co prawda, Habilitant podjął taki wysiłek, na co wskazuje współpraca z Al-Farabi Kazakh National University w Ałmatach, gdzie Habilitant pełni funkcję zagranicznego promotora naukowego w dwóch przewodach doktorskich. Natomiast brak jest dłuższych, prowadzących do wspólnych publikacji czy projektów pobytów czy staży naukowych za granicą. Mam nadzieję, że ta sytuacja ulegnie zmianie na przyszłych etapach kariery naukowej Habilitanta, do czego gorąco namawiam.

Dorobek pozacyklowy Habilitanta obejmuje m.in. tworzenie zasobów językowych dla języka polskiego, budowę i rozwój bazy PolNet, współtworzenie meta-korpusu literatury 19/20MetaPNC, zbioru POLygraph dotyczącego polskojęzycznych fake news oraz udział w pracach nad benchmarkiem międzyjęzykowego transferu wiedzy medycznej. Habilitant rozwijał również narzędzia do przeszukiwania baz leksykalnych, mierzenia podobieństwa semantycznego oraz rozpoznawania jednostek nazewniczych, a także prowadził prace nad normalizacją diachroniczną tekstu.

Na uwagę zasługuje aktywność projektowa. Habilitant uczestniczył w projektach dotyczących systemów konwersacyjnych, w tym POLINT-112-SMS i EUDAMU, brał udział w wyzwaniach badawczych dotyczących doboru odpowiedzi w dialogu, prowadził badania przemysłowe nad automatyczną korektą transkrypcji oraz wzbogacaniem korpusów głosowych z wykorzystaniem klonowania głosu. Wykazuje rolę wykonawcy w projektach FAKEBUSTER i DARIAH-PL oraz udział w projekcie aUPaEU dotyczącym partnerstw uniwersyteckich na rzecz akceleracji uniwersytetów europejskich.

Obecnie dr Kubis łączy pracę adiunkta w Zakładzie Sztucznej Inteligencji Wydziału Matematyki i Informatyki UAM z aktywnością w Centrum Sztucznej Inteligencji UAM, gdzie kieruje zespołem systemów konwersacyjnych. Jego zadania obejmują prowadzenie badań przemysłowych i wdrażanie wyników badań w praktyce gospodarczej. W mojej ocenie jest to działalność istotna zarówno naukowo, jak i z punktu widzenia transferu wiedzy.

Reasumując, **aktywność naukową Habilitanta poza głównym osiągnięciem oceniam pozytywnie**, choć z uwzględnieniem wskazanych wyżej mankamentów. Ma ona charakter ciągły, interdyscyplinarny i międzynarodowy, a jej ważną cechą jest łączenie badań podstawowych, budowy zasobów i narzędzi oraz badań przemysłowych.

4. Ocena pozostałej aktywności akademickiej dr. Marka Kubisa

Działalność dydaktyczna Habilitanta jest dobrze rozwinięta. Prowadzi on zajęcia z przedmiotów *sztuczna inteligencja*, *analiza i projektowanie obiektowe* oraz *systemy dialogowe*, a także seminarium magisterskie poświęcone systemom konwersacyjnym. Jest promotorem dziesięciu prac magisterskich dotyczących m.in. analizy i ujednoznaczniania wypowiedzi użytkownika, sterowania inicjatywą systemu konwersacyjnego, symulacji zachowania użytkownika, rozpoznawania emocji w mowie, modelowania satysfakcji użytkownika i wzbogacania danych uczących. Zakres tych tematów jest spójny z profilem badawczym Habilitanta i potwierdza zdolność do przekazywania studentom aktualnej wiedzy z obszaru NLP i systemów dialogowych.

Pozytywnie oceniam działalność organizacyjną. W latach 2009-2019 dr Kubis był członkiem komitetu organizacyjnego konferencji *Language & Technology Conference (LTC)*, od 2011 roku pełniąc funkcję sekretarza komitetu. Łącznie uczestniczył w organizacji sześciu edycji tej konferencji. Od 2023 roku współorganizuje coroczne *Symposium on Challenges for Natural Language Processing (CNLPS)* w ramach FedCSIS.

Habilitant wielokrotnie pełnił funkcje recenzenta oraz przewodniczącego obszaru tematycznego (*Area Chair*) w procedurze ACL Rolling Review, a także recenzował prace dla konferencji i czasopism z zakresu lingwistyki komputerowej. Szczególnie rola Area Chair w rozpoznawalnej na arenie światowej organizacji, jaką jest ACL, zasługuje na uznanie. Był członkiem komitetów programowych ścieżek przemysłowych i sesji demonstracyjnych czołowych konferencji z obszaru NLP, w tym ACL, EACL, EMNLP, NAACL i COLING. Uczestniczył również w komitetach doradczych i działaniach mentoringowych skierowanych do doktorantów, postdoków i początkujących badaczy, m.in. w ramach warsztatów YRRSDS i Student Research Workshops.

Na odnotowanie zasługuje także aktywność redakcyjna. Dr Kubis jest współredaktorem czterech tomów serii *Lecture Notes in Computer Science* poświęconych technologiom języka naturalnego i materiałom konferencji LTC. Jest to działalność pracochłonna, wymagająca doświadczenia naukowego i organizacyjnego, a jednocześnie korzystna dla rozwoju środowiska badawczego.

Podsumowując, **pozostałą aktywność akademicką Habilitanta oceniam bardzo dobrze**. Działalność dydaktyczna, promotorska, recenzencka, konferencyjna i redakcyjna jest systematyczna i wyraźnie wykracza poza minimalne oczekiwania stawiane kandydatom do stopnia doktora habilitowanego.

5. Konkluzja

Biorąc pod uwagę przedstawiony cykl sześciu powiązanych tematycznie publikacji, dorobek naukowy niewchodzący w skład tego cyklu, działalność projektową i wdrożeniową, współpracę międzynarodową, a także osiągnięcia dydaktyczne i organizacyjne, stwierdzam, że dr Marek Kubis posiada dojrzały i spójny profil naukowy w obszarze przetwarzania języka naturalnego i systemów konwersacyjnych.

Główne osiągnięcie naukowe obejmuje oryginalne metody analizy sieci konwersacji, grafowe metody ujednolicania odniesień, metody oceny odporności NLU na błędy rozpoznawania mowy oraz nowe procedury benchmarkowania systemów konwersacyjnych i multimodalnych modeli językowych. Osiągnięcia te są uzupełnione aktywnością naukową realizowaną przede wszystkim w UAM, we współpracy z przemysłem (Samsung Electronics).

W konsekwencji **stwierdzam, że osiągnięcia naukowe dr. Marka Kubisa spełniają wymagania określone w art. 219 ust. 1 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce**. W szczególności przedstawione osiągnięcie stanowi znaczący wkład w rozwój dyscypliny informatyka techniczna i telekomunikacja, a całokształt działalności Habilitanta potwierdza istotną aktywność naukową i akademicką.

Wnioskuje zatem o pozytywne zakończenie postępowania i popieram nadanie dr. Markowi Kubisowi stopnia doktora habilitowanego w dziedzinie nauk inżynieryjno-technicznych, w dyscyplinie informatyka techniczna i telekomunikacja.



Tomasz Trzcinski