

Recenzja osiągnięcia naukowego i opinia w postępowaniu w sprawie nadania stopnia doktora habilitowanego nauk technicznych w dyscyplinie informatyka techniczna i telekomunikacja dr. inż. Jarosławowi Protasiewiczowi

1. Podstawowe dane o kandydacie

Dr inż. Jarosław Protasiewicz ukończył studia w r. 1998 na kierunku Automatyka i Robotyka na Wydziale Elektrycznym Politechniki Białostockiej. Stopień doktora nauk technicznych w zakresie informatyki uzyskał w Instytucie Badań Systemowych Polskiej Akademii Nauk w Warszawie w r. 2009. Od r. 2005 pracuje w Ośrodku Przetwarzania Informacji - Państwowym Instytucie Badawczym, początkowo jako programista, następnie Kierownik Laboratorium Inteligentnych Systemów Informatycznych. Od r. 2019 jest dyrektorem tego Instytutu.

2. Recenzja osiągnięcia naukowego

Jako osiągnięcie naukowe będące podstawą ubiegania się o nadanie stopnia doktora habilitowanego Habilitant, korzystając z art. 219 ust. 1 pkt. 2b ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2021 r. poz. 478 z późn. zm.), wskazał cykl prac naukowych, zatytułowany: „Systemy rekomendacji wiedzy z algorytmami inteligencji obliczeniowej we wspomaganiu doboru recenzentów i ekspertów oraz innowacyjności”. Na ten cykl składa się jedna monografia autorska oraz sześć publikacji. Wskazane w powyższym cyklu prace to:

1. Protasiewicz J., *Knowledge recommendation systems with machine intelligence algorithms - People and innovations*, Springer Nature, 2023 (ISBN: 978-3-031-32695-0)
2. Protasiewicz, J., Pedrycz, W., Kozłowski, M., Dadas, S., Stanisławek, T., Kopacz, A., & Gałęzewska, M. (2016). A recommender system of reviewers and experts in reviewing problems. *Knowledge-Based Systems*, 106, 164-178. [IF=8.8]
3. Mirończuk, M. M., & Protasiewicz, J. (2020). Recognizing innovative companies by using a diversified stacked generalization method for website classification. *Applied Intelligence*, 50(1), 42-60. [IF=5.3].
4. Protasiewicz, J., & Dadas, S. (2016, October). A hybrid knowledge-based framework for author name disambiguation. In 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 000594-000600).
5. Protasiewicz, J., Mirończuk, M., & Dadas, S. (2017). Categorization of multilingual scientific documents by a compound classification system. In *Artificial Intelligence and Soft Computing: 16th International Conference, ICAISC 2017, Zakopane, Poland, June 11-15, 2017*, Proceedings, Part II 16 (pp. 563-573). Springer International Publishing.
6. Protasiewicz, J. (2017, October). Inventorum: A platform for open innovation. In 2017 IEEE International Conference on Systems, Man and Cybernetics (SMC) (pp. 10-15).
7. Protasiewicz, J. (2017, October). Inventorum - A recommendation system connecting business and academia. In 2017 IEEE International Conference on Systems, Man and Cybernetics (SMC) (pp. 1920-1925).

Głównym celem naukowym zrealizowanym w ramach tego osiągnięcia było (i) opracowanie koncepcji dwóch systemów rekomendacji wiedzy: systemu rekomendacji recenzentów i ekspertów oraz systemu wspierającego innowacyjność; (ii) dobór i rozwój algorytmów niezbędnych do budowy tych systemów. Monografia [1], umieszczona przez Habilitanta na tej liście na pierwszej pozycji, prezentuje najpełniejszy opis obu systemów, a równocześnie przedstawia zbiorcze ujęcie rezultatów wcześniej opublikowanych w pozostałych pracach z tej listy. Dlatego w pierwszej kolejności omówię najważniejsze wyniki przedstawione w tych pracach.

W pracy [2] przedstawiono architekturę i szczegóły realizacji systemu rekomendacyjnego, wspomagającego wybór recenzentów i ekspertów zapraszanych do oceny projektów badawczych lub artykułów.

System składa się z trzech modułów: akwizycji danych, wyszukiwania informacji oraz rekomendacji. Moduł akwizycji danych pozyskuje dane strukturalne (np. publikacje kandydatów na recenzentów z ogólnodostępnych baz danych), dane niestructuralne (np. opisy osiągnięć poszczególnych osób) oraz informacje podane przez użytkowników. Moduł wyszukiwania informacji przekształca zgromadzone dane w czterech procesach: klasyfikacji, ujednoznacznienia autorów, ekstrakcji słów kluczowych i indeksowania pełnotekstowego. Końcowym efektem jest zbudowanie profili recenzentów, reprezentujących ich wiedzę.

Głównym modułem systemu jest moduł rekomendacji, do którego są kierowane dwa rodzaje zapytań o recenzentów. W zapytaniach formalnych słowa kluczowe opisują tematykę przedstawianej do recenzji publikacji (lub projektu), natomiast zapytanie tekstowe to pełny tekst (lub streszczenie) artykułu lub projektu do recenzji. Przedstawiono dwa podstawowe algorytmy rekomendacji.

Dla zapytań formalnych algorytm rekomendacji bazuje na podobieństwie, wyrażonym kosinusową miarą odległości, wektora słów kluczowych projektu i wektorów słów kluczowych recenzentów, z uwzględnieniem wag definiujących stopień przynależności tych słów kluczowych do poszczególnych recenzentów.

W przypadku zapytań tekstowych zastosowano wyszukiwarkę pełnotekstową, która wyznacza podobieństwo między tekstem zapytania a tekstami opisującymi profile recenzentów. Sposób wyznaczania współczynników rankingowych przy wykorzystaniu wybranej do badań wyszukiwarki Apache Lucene jest rozwinięciem kosinusowej miary podobieństwa.

Krytyczna ocena własności obu przedstawionych algorytmów dała podstawę do zastosowania w projektowanym systemie nowego, hybrydowego modelu rekomendacji, w którym połączono oba rankingi. Dobierając współczynniki wag użytkownik może preferować wyniki pierwszego lub drugiego algorytmu rekomendacji. Takie rozwiązanie umożliwi eksperymentalne strojenie systemu.

W pracy przedstawiono przykłady praktycznej weryfikacji przedstawionych algorytmów rekomendacji. Wykorzystano zestaw danych dotyczących projektów z konkursu FP7 Komisji Europejskiej. Omówiono też wdrożenie systemu rekomendowania recenzentów i ekspertów w Narodowym Centrum Badań i Rozwoju.

Oceniając znaczenie wyników tej pracy w pierwszej kolejności wymienić należy opracowanie koncepcji całego systemu rekomendacji recenzentów. Drugim znaczącym wynikiem badawczym był dobór algorytmów realizujących poszczególne zadania systemu. Wśród nich były autorskie rozwiązania: system klasyfikacji dokumentów wielojęzycznych, algorytm wydobywania słów kluczowych, a także cały algorytm rekomendacji. Trzecim wartym podkreślenia wynikiem badawczym było konfigurowanie i uczenie wszystkich zastosowanych algorytmów. Wykonanie tego etapu świadczy o bardzo dobrej znajomości metod uczenia maszynowego.

Praca [3] prezentuje system klasyfikacji oceniający, czy firma jest innowacyjna. Źródłem danych dla procesu oceny są strony internetowe firm. Podstawą klasyfikacji są trzy zestawy danych: podstawowy opis firmy, linki do firm współpracujących z firmą ocenianą zamieszczone na analizowanej stronie tej firmy oraz zestaw dokumentów wybranych spośród wszystkich dokumentów znajdujących się na stronie. Każdy zestaw danych jest oceniany przez oddzielny klasyfikator (naiwny klasyfikator Bayesa). We wcześniejszej pracy tych samych autorów końcowy wynik takiego zestawu klasyfikatorów był wyznaczany metodą ważonego głosowania. W ocenianej pracy autorzy przedstawili nową propozycję wywodzącą się z metod zespołowego uczenia - finalny wynik klasyfikacji jest wyznaczany przez meta-klasyfikator. W pracy rozpatrzono kilka metod uczenia tego meta-klasyfikatora: metody najbliższych sąsiadów, naiwnego modelu Bayesa, maszyny wektorów wspierających i drzew decyzyjnych.

Innym problemem badawczym rozwiązany w pracy był dobór liczby cech uwzględnianych przez klasyfikatory wejściowe, gdyż stwierdzono, że klasyfikatory dają najlepsze wyniki w różnych zakresach liczb cech. Do znalezienia optymalnej przestrzeni cech zastosowano algorytm genetyczny.

Do oryginalnych wyników tej pracy należy zaliczyć: (i) przedstawienie koncepcji oceny innowacyjności firm na podstawie ich stron internetowych, (ii) opracowanie metody uczenia meta-klasyfikatora, w tym przedstawienie formalnego opisu tego procesu.

W ocenie znaczenia wyników tej pracy na wyróżnienie zasługują też: (i) wybór dla potrzeb wykonywanych eksperymentów autorskiego zestawu testowego obejmującego ponad 2700 rzeczywistych witryn internetowych, dobranych na podstawie rankingów wiodących firm na rynku oraz (ii) wszechstronna analiza jakości wykonywanej klasyfikacji, prowadzona przy wykorzystaniu wielu miar i testów statystycznych.

W pracy [4] podjęto problem ujednoznaczniania nazwisk autorów, który pojawia się m.in. w przypadku występowania dwóch lub więcej autorów o tych samych nazwiskach i imionach. Zaproponowano dwupoziomowy proces ujednoznaczniania.

Pierwszy poziom tworzy algorytm regułowy, który iteracyjnie dopasowuje publikacje do istniejących profili badaczy. Heurystyczne reguły wykorzystują atrybuty, które można bezpośrednio wyodrębnić z publikacji, tj. imiona i nazwiska, współautorstwo, afiliacje, dziedziny nauki i słowa kluczowe.

Drugi poziom ujednoznaczniania stanowi algorytm grupowania hierarchicznego, który jest zmodyfikowaną wersją hierarchicznego grupowania aglomeracyjnego (*Hierarchical Agglomerative Clustering*). Zaproponowano dwa warianty oceny podobieństwa tworzonych klastrów: podobieństwa określonego według reguł heurystycznych, przy wykorzystaniu m.in. wag wymienionych poprzednio atrybutów publikacji oraz podobieństwa ocenianego przez klasyfikatory trenowane na danych uzyskanych z pierwszego poziomu ujednoznaczniania.

Przeprowadzone badania pokazały, że algorytm pierwszego poziomu, czyli oparty na regułach pozwala dopasować publikacje jedynie do już istniejących profili autorów. Natomiast algorytm grupowania jest w stanie nie tylko rozpoznać znanego autora, ale także stworzyć nowy profil.

Opracowany algorytm został wszechstronnie zbadany eksperymentalnie na zbiorze kilkuset autorów i tysięcy publikacji; w badaniach sprawdzono kilka klasyfikatorów i kryteriów stopu algorytmu.

Podsumowując, znaczącym wynikiem artykułu jest algorytm dwupoziomowego ujednoznaczniania nazwisk autorów. Do ważnych elementów wniesionych przez autorów do konstrukcji tego algorytmu należą m.in. wybór reguł opartych na wiedzy, dobór klasyfikatorów, opracowanie grupowania hierarchicznego z adaptacyjnymi kryteriami zatrzymania algorytmu.

W pracy [5] autorzy podjęli problem uwzględnienia wielojęzyczności publikacji w procesie ich klasyfikacji. Przedstawiono propozycję autorskiego systemu klasyfikacji dokumentów (publikacji naukowych), których pewne fragmenty (tytuł, streszczenie, słowa kluczowe) są zapisane w różnych językach. Publikacje powinny być klasyfikowane według dziedzin nauki, których dotyczy ich treść.

W pracy zaproponowano trzystopniowy proces klasyfikacji. W pierwszym etapie klasyfikowany dokument jest dzielony na części zapisane w różnych językach i są tworzone wektorowe reprezentacje tych części. W drugim etapie oddzielne wielojęzyczne klasyfikatory klasyfikują poszczególne części przypisując im prawdopodobieństwa przynależności do wszystkich możliwych kategorii klasyfikacji. W ostatnim kroku wielojęzyczny moduł decyzyjny klasyfikuje cały dokument na podstawie wyników klasyfikacji poszczególnych jego części.

Realizacja tej koncepcji wymagała m.in. wyboru klasyfikatorów części dokumentów oraz wyboru modułu decyzyjnego, który wypracowuje końcową klasyfikację. Wybór klasyfikatorów obejmował porównanie wielomianowego naiwnego klasyfikatora Bayesa z klasyfikatorem w postaci sieci LSTM (*Long Short-Term Memory*), natomiast jako moduł decyzyjny wybrano zespół klasyfikatorów regresji logistycznej.

W przeprowadzonych badaniach eksperymentalnych wykazano przewagę zaproponowanej klasyfikacji trójstopniowej nad klasyczną klasyfikacją całych dokumentów.

Głównym osiągnięciem artykułu jest przedstawienie koncepcji złożonego, autorskiego systemu klasyfikacji wielojęzycznych publikacji naukowych i dobór algorytmów realizujących poszczególne etapy klasyfikacji. Bardzo ważnym krokiem w realizacji całego systemu było zaplanowanie i wykonanie eksperymentów pozwalających wybrać najlepszy, spośród wielu badanych, wariant złożonego systemu.

Artykuł konferencyjny [6] przedstawia ogólnie system informacyjny Inventorum, którego celem jest zbieranie i udostępnianie informacji dotyczących innowacji, firm zainteresowanych działalnością innowacyjną, projektów, ekspertów, a także konferencji. Dla użytkowników systemu tworzone są opisujące ich profile, które są następnie podstawą rekomendacji kierowanych do tych użytkowników ofert i informacji.

Praca nie przedstawia wyników badań naukowych, jest natomiast raportem prezentującym cele, funkcje i ogólną architekturę systemu, a także wykorzystane do jego budowy technologie programowe.

Z tego powodu dyskusyjne jest włączenie pracy do listy prac wchodzących w skład osiągnięcia naukowego. Ewentualnie należy ją rozpatrywać łącznie z pracą [7].

Praca [7] przedstawia drugi główny składnik osiągnięcia naukowego Habilitanta, a mianowicie projekt systemu informacyjnego dla wspierania innowacyjności, wymiany informacji i kooperacji pomiędzy instytucjami nauki i biznesu. W projektowanym systemie wyodrębniono trzy grupy potencjalnych użytkowników - są nimi naukowcy, jednostki naukowe oraz przedsiębiorstwa. Ponadto są gromadzone dane o innowacjach, projektach i konferencjach. Przedstawiono procesy zbierania danych z otwartych baz danych, z zasobów Internetu, w szczególności stron internetowych przedsiębiorstw, a także zbierania danych z udziałem samych użytkowników. Na podstawie danych określających dorobek naukowców, specyfikę działalności jednostek naukowych oraz zakresy działalności przedsiębiorstw są tworzone profile tych grup użytkowników.

Analiza profili użytkowników jest następnie podstawą rekomendacji informacji przedstawianych użytkownikom. Zadaniem algorytmu rekomendacji jest dopasowanie możliwości jednostek naukowych i ekspertów do potrzeb przedsiębiorstw. Drugą funkcją algorytmu jest rekomendowanie użytkownikom opracowanych już innowacji, projektów oraz programów konferencji, które mogą być przedmiotem ich zainteresowania.

Do realizacji takiego dopasowania autor zaproponował algorytm filtrowania opartego na treści (*content based filtering*), z użyciem kosinusowej miary podobieństwa dopasowywanych obiektów. Miara ta wyznaczana jest na wektorach cech tworzonych dla profili porównywanych użytkowników systemu. W przypadku obiektów charakteryzowanych zestawem dokumentów tekstowych zastosowano wyszukiwarki pełnotekstowe, które indeksują dokumenty według bardziej złożonych algorytmów rankingowych, ale również przy wykorzystaniu kosinusowej miary podobieństwa.

Przedstawiona koncepcja została zaimplementowana w praktycznie zrealizowanym i dostępnym poprzez Internet systemie o nazwie Inventorum.

Oryginalnym osiągnięciem badań przedstawionych artykule jest opracowanie koncepcji unikalnego systemu dla wspierania innowacyjności. Całość tego osiągnięcia uzupełnia dobór algorytmów realizujących funkcje systemu, w szczególności algorytmu rekomendacji, którego rolą jest inicjowanie współpracy przedsiębiorstw i naukowców.

Monografia [1] jest podsumowaniem wyników badań przedstawionych w zbiorze artykułów tworzących osiągnięcie naukowe Habilitanta. Jej treść skupia się na opisie koncepcji, a następnie realizacji systemu rekomendacji recenzentów i ekspertów oraz systemu wspierającego innowacyjność, obejmuje także analizę algorytmów wybranych do stworzenia tych systemów.

Monografia liczy siedem rozdziałów. Rozdział pierwszy jest rozdziałem wprowadzającym. Drugi rozdział poświęcono przeglądowi literatury dotyczącej tematyki prowadzonych badań.

W trzecim rozdziale monografii przedstawiono koncepcję systemu rekomendacji recenzentów i ekspertów. Opis systemu jest nieco szerszym rozwinięciem treści przedstawionych w artykule [2] z listy prac stanowiących osiągnięcie naukowe Habilitanta. Szerzej omówiono rolę takiego systemu oraz eksperymenty wykonane na stworzonym systemie. Ocena wyników badawczych, które przyniosła realizacja tego systemu została już omówiona w ramach recenzji pracy [2].

Czwarty rozdział monografii przedstawia proces tworzenia systemu informacyjnego dla wspierania innowacyjności Inventorum, który został już omówiony w recenzji pracy [7]. Treść tego rozdziału ujmuje przy tym w nieco szerszym zakresie opis funkcjonalności systemu. We wdrożonym do użytku systemie udostępniono mianowicie trzy formy rekomendacji: „szybkie” rekomendacje po rozpoczęciu pracy w systemie, przedstawiające użytkownikowi najnowsze i spersonalizowane informacje rekomendowane przez system; pełne rekomendacje – generowane przez przedstawiony w pracy [7] algorytm rekomendacji oraz możliwość samodzielnego wyszukiwania informacji w systemie przez użytkownika przy pomocy udostępnionej wyszukiwarki pełnotekstowej.

Znaczenie wyników badawczych uzyskanych w trakcie realizacji systemu Inventorum przedstawiono już w recenzji pracy [7].

Rozdział piąty monografii zawiera przegląd wybranych algorytmów pozyskiwania danych i wydobywania z nich informacji, zastosowanych przez autora w systemach omówionych w dwóch poprzednich rozdziałach. Algorytmy te były przedmiotem wcześniejszych publikacji autora, ale autor przedstawił w monografii ich poszerzone ujęcie.

Pierwsze dwa przedstawiane algorytmy realizują akwizycję i przetwarzanie danych. Algorytm ekstrakcji zastosowano do pobierania danych z baz danych publikacji, a następnie transformacji danych, z użyciem pośredniej reprezentacji w XML, do postaci przyjętej w lokalnej bazie danych. Algorytm wyszukiwania pełnotekstowego jest podstawą działania robota indeksującego (*crawler tematyczny*), który zbiera dodatkowe informacje o publikacjach ze stron internetowych naukowców. Analiza tekstu stron i wydobywanie informacji o publikacjach jest realizowane z wykorzystaniem modelu CRF (*Conditional Random Fields*).

Kolejno przedstawione zostały algorytmy przekształcania danych w informacje wykorzystywane do budowy profili kandydatów na recenzentów. W pierwszej kolejności rozważono problem klasyfikacji publikacji według hierarchicznej struktury dziedzin i dyscyplin naukowych. Do badań wybrano następujące algorytmy: wielomianowy naiwny klasyfikator bayesowski, maszyna wektorów wspierających, perceptron wielowarstwowy. Do kategoryzacji publikacji została wykorzystana ontologia OSJ (*Ontology of Scientific Journals*), tworząca trójpoziomowe drzewo dziedzin, poddziedzin i dyscyplin nauki. Do przypisania publikacji do właściwych kategorii zastosowano zespół wymienionych wyżej klasyfikatorów, modelujących kolejne poziomy wspomnianego drzewa. Wyróżniono płaską i hierarchiczną organizację takich klasyfikatorów.

Opracowanie koncepcji profili, wybór algorytmów do ich budowy i konfiguracja tych algorytmów na podstawie eksperymentalnych badań na rzeczywistych danych stanowią istotny element dorobku badawczego Habilitanta.

Wśród cech danych, które są podstawą klasyfikacji wyróżniono wielojęzyczność publikacji. Prezentowany autorski system klasyfikacji dokumentów wielojęzycznych został już przedstawiony w zrecenzowanej wcześniej pracy [5] i tam wskazano oryginalne elementy, które wnoszą ten system klasyfikacji do dorobku Habilitanta.

Następnym prezentowanym algorytmem był algorytm ujednoznaczniania autorów publikacji. Algorytm ten został już omówiony w ramach recenzji pracy [4], gdzie wskazano też znaczące wyniki badawcze wnoszone do dorobku Habilitanta przez ten algorytm.

Kolejny przedstawiony algorytm jest autorskim algorytmem wydobywania słów kluczowych PKE (*Polish Keyword Extractor*). Opracowany algorytm korzysta z lematyzatora języka polskiego, filtru części mowy, dwu metod wyboru fraz kandydujących na słowa kluczowe i dwu metod oceny wybranych kandydatów. Własności algorytmu zostały ocenione eksperymentalnie na abstraktach publikacji. Algorytm PKE porównano z algorytmem KEA (*Keyphrases Extraction Algorithm*) wydobywania słów kluczowych z tekstów napisanych w języku angielskim. Opracowanie algorytmu jest oryginalnym elementem badawczego dorobku Habilitanta.

Ostatni punkt piątego rozdziału monografii prezentuje model oceny innowacyjności przedsiębiorstwa. Model ten został wcześniej przedstawiony w pracy [3] i w recenzji tej pracy omówiono oryginalne wyniki, które ten model wnosi do dorobku Habilitanta.

Szósty, ostatni rozdział monografii dotyczy implementacji programowej systemu rekomendacji recenzentów i ekspertów oraz systemu wspomagającego innowacyjność. Przedstawiono architekturę systemów od strony technicznej, zastosowane do ich budowy technologie i narzędzia programowe, a także funkcje udostępniane przez interfejsy obu systemów. Podano też informacje o wdrożeniach tych systemów, zgromadzonych danych oraz statystykach korzystania z systemów.

Oceniając oryginalność monografii należy podkreślić, że stanowi ona zbiorcze ujęcie wyników wcześniej opublikowanych w oddzielnych pracach i ocenionych w recenzjach tych prac. Jej opracowanie wnosi przy tym nowe wartości w dorobku Habilitanta, co uzasadniają następujące argumenty:

- Monografia udostępnia całościowy opis dwóch systemów rekomendacji wiedzy: systemu rekomendacji recenzentów oraz systemu wspomagania innowacyjności – wspólnym elementem obu systemów jest budowa profili użytkowników systemów.
- Monografia przedstawia szersze opracowanie stanu wiedzy w podejmowanej dziedzinie – obejmuje bowiem analizę literatury w znacznie dłuższym zakresie czasu niż wcześniejsze publikacje autora.
- W monografii przedstawiono pełne informacje o doświadczeniach wynikających z wdrożenia i dotychczasowego użytkowania wymienionych systemów – monografia może

stanowić kompendium nie tylko teoretycznej, ale również praktycznej wiedzy dla osób podejmujących badania w rozważanej dziedzinie.

Podsumowanie punktu 2 recenzji - dorobku stanowiącego osiągnięcie naukowe

Podsumowując ocenę cyklu prac wskazanych jako osiągnięcie naukowe Habilitanta, do najważniejszych wyników przedstawionych w tych publikacjach zaliczam:

- 1) Opracowanie koncepcji dwóch systemów informatycznych realizujących rekomendację wiedzy: systemu rekomendacji recenzentów i ekspertów oraz systemu informacyjnego umożliwiającego wspieranie innowacyjności.
- 2) Wybór i rozwój algorytmów realizujących funkcje wymienionych systemów, w szczególności następujących algorytmów:
 - a) algorytmy akwizycji danych,
 - b) algorytmy przekształcania danych w informacje wykorzystywane do budowy profili, w tym algorytm klasyfikacji publikacji,
 - c) autorski system klasyfikacji dokumentów wielojęzycznych,
 - d) algorytm ujednoznacznianie autorów publikacji,
 - e) autorski algorytm wydobywania słów kluczowych z tekstów,
 - f) model rekomendacji recenzentów,
 - g) model oceny innowacyjności przedsiębiorstwa.

Należy podkreślić, że podstawą wyboru zastosowanych algorytmów i modeli były badania oceniające przydatność wielu metod w projektowanych systemach. Rozwój wybranych algorytmów, niezbędny do ich dostosowania do realizacji oczekiwanych zadań, wymagał szerokiego warsztatu naukowego Habilitanta, w szczególności bardzo dobrej znajomości metod uczenia maszynowego.

Bardzo ważnym atutem badań prowadzonych przez Habilitanta była eksperymentalna, praktyczna weryfikacja opracowanych rozwiązań, bazująca na specjalnie przygotowanych dużych zbiorach rzeczywistych danych.

Na koniec należy podkreślić, że zaprojektowane systemy zostały wykonane i są od szeregu lat użytkowane.

Jeśli chodzi o publikacyjną rangę artykułów przedstawionych jako osiągnięcie naukowe, to dwa spośród nich zostały wydane w czasopiśmie naukowych (*Knowledge-Based Systems* (IF=8.8), *Applied Intelligence*, (IF=5.3)). Trzy artykuły zostały wydane w materiałach międzynarodowej konferencji *IEEE International Conference on Systems, Man, and Cybernetics*, indeksowanych w bazie WoS; jeden artykuł w materiałach międzynarodowej konferencji *Artificial Intelligence and Soft Computing*. Monografia, stanowiąca całościowe ujęcie uzyskanych wyników, została wydana w wydawnictwie Springer w ramach serii „*Studies in Computational Intelligence*”.

Habilitant jest samodzielnym autorem trzech spośród siedmiu publikacji przedstawionych jako osiągnięcie naukowe. W czterech pozostałych publikacjach Habilitant wymienia w opisie swojego wkładu w powstanie tych prac: „współtworzenie metodologii badań” lub „opracowaniu koncepcji i zaprojektowanie systemu”, a także „napisanie publikacji”. Deklarowany udział w opracowaniu tych publikacji wynosi się od 40 do 80%.

Ranga części wydawnictw, w których ukazały się publikacje przedstawione jako osiągnięcie naukowe nie jest wprawdzie zbyt wysoka, ale prace te przedstawiają zestaw opracowanych przez autora, spójnych tematycznie metod wykorzystanych do stworzenia dwóch unikalnych systemów rekomendacji wiedzy.

Podsumowując tę część recenzji stwierdzam, że przedstawiony cykl publikacji stanowi oryginalny wkład naukowy w dyscyplinie informatyka techniczna i telekomunikacja, w obszarze rekomendacji wiedzy, klasyfikacji i innych metod uczenia maszynowego. W mojej ocenie uzyskane wyniki spełniają wymagania osiągnięcia naukowego w rozumieniu art. 219 ust. 1 pkt. 2 ustawy Prawo o szkolnictwie wyższym i nauce z dnia 20 lipca 2018 r. z późniejszymi zmianami.

3. Ocena pozostałego dorobku naukowego

Pozostały dorobek publikacyjny po uzyskaniu stopnia doktora

Większa część badań prowadzonych przez Habilitanta po doktoracie dotyczyła tematyki przedstawionej w monografii i sześciu artykułach wskazanych jako osiągnięcie naukowe. Tematyka ta została więc scharakteryzowana w poprzednim punkcie niniejszej recenzji. Dodatkowe autorskie i współautorskie publikacje Habilitanta z tego obszaru obejmowały: 3 artykuły w czasopismach (*IEEE Access*, *Expert Systems with Applications*), 2 monografie (wydane przez *National Information Processing Institute*) oraz 9 artykułów konferencyjnych.

Oddzielny obszar badań prowadzonych przez Habilitanta po doktoracie był związany z Jego działalnością zawodową i dotyczył projektowania i wdrażania systemów informatycznych umożliwiających gromadzenie, przetwarzanie i udostępnianie danych dotyczących nauki i szkolnictwa wyższego w Polsce. Tematyki tej dotyczyło 9 współautorskich publikacji Habilitanta, z czego 8 wydano w materiałach konferencyjnych.

Ponadto Habilitant wydał po doktoracie 3 publikacje bliższe tematycznie pracy doktorskiej, a dotyczące wykorzystania sieci neuronowych w algorytmach prognozowania zużycia energii elektrycznej. Jedną z tych prac to publikacja autorska, a dwie współautorskie.

W sumie po doktoracie Habilitant wydał 33 publikacje, w tym 7 prac w czasopismach, 3 monografie (w tym 1 autorska) oraz 23 publikacji konferencyjnych (w tym 4 autorskie). Siedem spośród tych publikacji tworzy osiągnięcie naukowe.

Podsumowanie całego dorobku publikacyjnego

Autorstwo lub współautorstwo publikacji naukowych

Po doktoracie Habilitant wydał 33 publikacje, w tym 7 przedstawionych jako osiągnięcie naukowe. Publikacje te obejmują:

- 7 prac w czasopismach,
- 3 monografie naukowe, w tym 1 autorska,
- 23 publikacje konferencyjne, w tym 4 autorskie.

Pozostałe liczbowe wskaźniki dorobku naukowego

- sumaryczny *Impact Factor* publikacji stanowiących osiągnięcie naukowe – 14.1
- łączny *Impact Factor* wszystkich publikacji, z uwzględnieniem roku publikacji – 22.726
- liczba cytowań według WoS – 248, Scopus – 317, Google Scholar – 583
- indeks Hirscha według WoS – 5, Scopus – 5, Google Scholar – 9

W mojej ocenie liczba publikacji i wskaźniki bibliometryczne Habilitanta są na dobrym poziomie.

Inne formy działalności naukowo-badawczej Habilitanta

Habilitant kierował czterema projektami obejmującymi realizację systemów informatycznych dotyczących nauki, szkolnictwa wyższego i innowacyjnej gospodarki:

- 1) Kierownik projektu RAD-on „Zintegrowany system usług dla nauki – etap II” Program Operacyjny Polska Cyfrowa w ramach Poddziałania 2.3.1 „Cyfrowe udostępnienie informacji sektora publicznego ze źródeł administracyjnych i zasobów nauki” (typ I projektu: Cyfrowe udostępnienie ISP ze źródeł administracyjnych. Projekt realizowany w latach 2017-2021.
- 2) Kierownik projektu „Ogólnopolska platforma informacyjna o potencjale innowacyjnym Inventorum” – projekt realizowany w latach 2014 – 2015 w ramach Programu Innowacyjna Gospodarka na lata 2007–2013: priorytet 2 „Infrastruktura sfery B+R”, działanie 2.3 „Inwestycje związane z rozwojem infrastruktury informatycznej nauki”, poddziałanie 2.3.2 „Projekty w zakresie rozwoju zasobów informacyjnych nauki w postaci cyfrowej”.
- 3) Kierownik projektu „System wspomagania wyboru recenzentów” – projekt realizowany w latach 2011-2012 w ramach projektu systemowego „Wsparcie systemu zarządzania badaniami naukowymi oraz ich wynikami”, w ramach Programu Operacyjnego Innowacyjna Gospodarka 2007–2013 (priorytet I, działanie 1.1, poddziałanie 1.1.3).
- 4) Kierownik projektu „System informacji o szkolnictwie wyższym “POL-on” – projekt realizowany w latach 2010-2014 Program Operacyjny Kapitał Ludzki: priorytet IV. Nauka i Szkolnictwo Wyższe , Działanie 4.1. „Wzmocnienie i rozwój potencjału dydaktycznego uczelni oraz zwiększenie liczby absolwentów kierunków o kluczowym znaczeniu dla gospodarki opartej na wiedzy”, poddziałanie 4.1.3 „Wzmocnienie systemowych narzędzi zarządzania szkolnictwem wyższym”.

Prowadzenie tak znaczących zadań jest świadectwem wysokich kompetencji i kwalifikacji Habilitanta w zakresie realizacji złożonych projektów informatycznych. Należy również podkreślić, że dwa z tych projektów dotyczą realizacji systemów, których koncepcja i zastosowane algorytmy są tematem publikacji wskazanych jako osiągnięcie naukowe Habilitanta.

Spośród jeszcze innych form działalności naukowej należy zauważyć, że Habilitant odbył wizytę studyjną i wygłosił wykład w *University of Alberta* w Kanadzie (r. 2016).

Podsumowanie pozostałego dorobku naukowego

Oceniam dorobek naukowy Habilitanta mieszczący się poza ocenionym już osiągnięciem naukowym jako dość dobry. Liczba publikacji i inne wskaźniki bibliometryczne są na dobrym poziomie, ale całość obniża nieco brak samodzielnych publikacji w czasopismach wyróżnionych wskaźnikiem IF (*Impact Factor*).

4. Ocena dorobku dydaktycznego i pozostałych osiągnięć

Habilitant prowadzi zajęcia dydaktyczne od 2002 roku na Wydziale Informatyki w Wyższej Szkole Informatyki Stosowanej i Zarządzania w Warszawie. Początkowo prowadził zajęcia laboratoryjne z przedmiotu Sieci neuronowe i zbiory rozmyte, a od r. 2008 prowadzi również wykłady z tego przedmiotu. W latach 1999 – 2001 prowadził zajęcia laboratoryjne z przedmiotów Programowanie w języku Pascal oraz Programowanie w języku C na Wydziale Technik Informatycznych w Politechnice Warszawskiej.

Habilitant był promotorem ponad 25 prac dyplomowych z zakresu sztucznej inteligencji, uczenia maszynowego i inżynierii oprogramowania w Wyższej Szkole Informatyki Stosowanej i Zarządzania w Warszawie, na Wydziale Informatyki i Grafiki. Obecnie prowadzi sześć prac dyplomowych.

Habilitant był promotorem pomocniczym rozprawy doktorskiej „Iteracyjne metody wykrywania hierarchicznych struktur jednostek nazewniczych”, obronionej w Instytucie Badań Systemowych Polskiej Akademii Nauk (r. 2022). Aktualnie Habilitant jest promotorem pomocniczym dwóch doktoratów wdrożeniowych realizowanych w ramach Szkoły Doktorskiej Technologii Informatycznych i Biomedycznych Instytutów PAN oraz doktoratu wdrożeniowego realizowanego w ramach Szkoły Doktorskiej Akademii Leona Koźmińskiego.

Recenzowanie publikacji w czasopismach międzynarodowych

Habilitant wykonał kilkadziesiąt recenzji artykułów dla następujących czasopism naukowych: wydawnictwo Elsevier: *Applied Soft Computing* (2 recenzje), *Expert Systems With Applications* (47), *Knowledge-Based Systems* (11), *Information Sciences* (1); wydawnictwo ACM: *Transactions on Information Systems* (1), *Transactions on Knowledge Discovery from Data* (1); wydawnictwo IEEE: *IEEE Transactions on Cybernetics* (3); wydawnictwo Springer: *Granular Computing* (2), *Progress in Artificial Intelligence* (1), *Scientometrics* (3).

Recenzowanie wniosków o projekty

Habilitant uczestniczył kilkakrotnie w zespołach oceniających wnioski projektowe w programach europejskich. W latach 2020, 2022, 2023 recenzował projekty w programach Horizon Europe - CNECT / DG for Communications Networks, Content and Technology, a w r. 2023 dodatkowo recenzował projekty w programie Digital Europe - CNECT / DG for Communications Networks, Content and Technology.

Członkostwo w międzynarodowych organizacjach i towarzystwach naukowych

- Członek (*senior member*) Institute of Electrical and Electronics Engineers (IEEE) od 2011 r.

Udział w pracach komisji i zespołów związanych z nauką

- Sekretarz Komisji Wyborczej do Rady Doskonałości Naukowej w r. 2019 oraz 2023.
- Kapituła Rankingu Szkół Wyższych od roku 2020, prowadzonego Fundacją Edukacyjną „Perspektywy” – udział w opracowaniu rankingu szkół wyższych od roku 2020.
- Członek Zespołu Doradczego do opracowania innowacyjnych kierunków rozwoju usług metrologicznych Głównego Urzędu Miar – podjęta tematyka: aspekty sztucznej inteligencji w metrologii.
- Członek zespołu do oceny wykorzystania środków finansowych w ramach przedsięwzięcia „Polski MOOC” (2018-2020).
- Członek zespołu do spraw przygotowania Programu INFRASTARt (2020).

Współpraca z otoczeniem społecznym i gospodarczym

- 1) Poza kierowaniem, opracowaniem założeń funkcjonalnych i współudziałem w opracowaniu rozwiązań algorytmicznych dla czterech wymienionych już systemów informatycznych Habilitant powiększył swój dorobek technologiczny w następujących działaniach:
 - współudział w opracowaniu architektury systemu informatycznego „Obsługa Strumieni Finansowania” w latach 2005 – 2011 jako programista i projektant;
 - wdrożenie metodyki Agile – Scrum w Laboratorium Inteligentnych Systemów Informatycznych OPI PIB.

- 2) Współpraca z otoczeniem gospodarczym w ramach pełnionej funkcji dyrektora OPI PIB – współudział w opracowanie strategii OPI PIB na lata 2021 – 2023 i dalej, zakładającej poszerzenie działalności gospodarczej instytutu.

Osiągnięcia organizacyjne

Z pracą w OPI PIB wiąże się istotne osiągnięcia organizacyjne Habilitanta, którym było utworzenie i kierowanie (w latach 2011 - 2019) Laboratorium Inteligentnych Systemów Informatycznych. 80-osobowy zespół laboratorium zajmował się wytwarzaniem oprogramowania oraz realizacją badań w zakresie uczenia maszynowego i inżynierii oprogramowania.

Podsumowanie dorobku dydaktycznego i pozostałych osiągnięć

Habilitant ma solidne doświadczenie dydaktyczne – od 25 lat prowadzi zajęcia, najpierw 2 lata w Politechnice Warszawskiej, a następnie do chwili obecnej w Wyższej Szkole Informatyki Stosowanej i Zarządzania. Przygotował tam i prowadził wykłady i zajęcia laboratoryjne.

Jest bardzo aktywny we współpracy z doktorantami. Był promotorem pomocniczym w jednym, ukończonym już przewodzie doktorskim, a aktualnie jest promotorem pomocniczym w trzech otwartych przewodach.

Jest bardzo aktywnym recenzentem ważnych międzynarodowych czasopism naukowych, wykonał dla nich ponad 60 recenzji. Wniósł duży wkład w pracach zespołów recenzujących projekty badawczo-rozwojowych Komisji Europejskiej.

Jest aktywnym członkiem szeregu komisji i zespołów związanych z organizacją nauki. Legitymuje się szeregiem osiągnięć we współpracy z otoczeniem społecznym i gospodarczym.

Podsumowując, oceniam pozytywnie dorobek dydaktyczny Habilitanta i jego aktywność w zakresie organizacji nauki.

5. Wniosek końcowy

Biorąc pod uwagę pozytywną ocenę osiągnięcia naukowego pt.: „Systemy rekomendacji wiedzy z algorytmami inteligencji obliczeniowej we wspomaganii doboru recenzentów i ekspertów oraz innowacyjności” oraz pozytywną ocenę dorobku naukowego, dydaktycznego i organizacyjnego Habilitanta, stwierdzam że dr inż. Jarosław Protasiewicz spełnia wymagania stawiane przez ustawę kandydatom do stopnia naukowego doktora habilitowanego.

W związku z tym popieram wniosek o nadanie dr. inż. Jarosławowi Protasiewiczowi stopnia naukowego doktora habilitowanego w dyscyplinie informatyka techniczna i telekomunikacja.

