

Białystok, 22 stycznia 2024 roku

Prof. dr hab. Jarosław Stepaniuk

Katedra Systemów Informacyjnych i Sieci Komputerowych

Wydział Informatyki, Politechnika Białostocka

**Recenzja osiągnięcia naukowego**  
**Pana dr Andrzeja Janusza pod tytułem**  
**„Wybrane metody reprezentacji obiektów i pojęć w danych dla**  
**algorytmów uczenia maszynowego”**

Pan dr Andrzej Janusz jest aktualnie zatrudniony na stanowisku adiunkta na Wydziale Matematyki, Informatyki i Mechaniki Uniwersytetu Warszawskiego. W czerwcu 2023 roku wystąpił z wnioskiem o przeprowadzenie postępowania w sprawie nadania stopnia doktora habilitowanego w dziedzinie nauk inżynieryjno-technicznych w dyscyplinie informatyka techniczna i telekomunikacja.

Habilitant przedstawił, do oceny, cykl powiązanych tematycznie 11 publikacji wieloautorskich w języku angielskim opublikowanych w latach 2014 – 2023. W tym artykuły w takich czasopismach jak Information Sciences, Engineering Applications of Artificial Intelligence czy też International Journal of Approximate Reasoning oraz publikacje w materiałach konferencji międzynarodowych.

Szczegółowa lista powiązanych tematycznie publikacji Habilitanta (wraz z informacją o Jego wkładzie w przypadku wszystkich jedenastu prac współautorskich) została przedstawiona w autoreferacie. Poniżej przedstawię skróconą listę:

- (O1) Andrzej Janusz, Sebastian Stawicki, Hung Son Nguyen: Adaptive Learning for Improving Semantic Tagging of Scientific Articles.
- (O2) Andrzej Janusz, Dominik Ślęzak: Rough Set Methods for Attribute Clustering and Selection.
- (O3) Andrzej Janusz, Dominik Ślęzak: Computation of Approximate Reducts with Dynamically Adjusted Approximation Threshold.

- (O4) Andrzej Janusz, Marek Grzegorowski, Marcin Michalak, Łukasz Wróbel, Marek Sikora, Dominik Ślęzak: Predicting seismic events in coal mines based on underground sensor measurements.
- (O5) Sebastian Stawicki, Dominik Ślęzak, Andrzej Janusz, Sebastian Widz: Decision bireducts and decision reducts - a comparison.
- (O6) Andrzej Janusz, Dominik Ślęzak: Investigating Similarity between Hearthstone Cards: Text Embeddings and Interchangeability Approaches.
- (O7) Mateusz Przyborowski, Tomasz Tajmajer, Łukasz Grad, Andrzej Janusz, Piotr Biczuk, Dominik Ślęzak: Toward Machine Learning on Granulated Data - a Case of Compact Autoencoder-based Representations of Satellite Images.
- (O8) Dominik Ślęzak, Marek Grzegorowski, Andrzej Janusz, Michał Kozielski, Sinh Hoa Nguyen, Marek Sikora, Sebastian Stawicki, Łukasz Wróbel: A framework for learning and embedding multi-sensor forecasting models into a decision support system: A case study of methane concentration in coal mines.
- (O9) Andrzej Janusz, Dominik Ślęzak: Analytics over Multi-sensor Time Series Data - A Case-Study on Prediction of Mining Hazards.
- (O10) Andrzej Janusz, Daniel Kałuża, Maciej Matraszek, Łukasz Grad, Maciej Świechowski, Dominik Ślęzak: Learning Multimodal Entity Representations and Their Ensembles with Applications in a Data-Driven Advisory Framework for Video Game Players.
- (O11) Andrzej Janusz, Dominik Ślęzak, Sebastian Stawicki, Krzysztof Stencel: A Practical Study of Methods for Deriving Insightful Attribute Importance Rankings using Decision Bireducts.

Należy podkreślić, że są to w wielu przypadkach publikacje w renomowanych czasopismach. Informacja o punktach przypisanych do prac z wykazu osiągnięć naukowych Habilitanta (na podstawie listy załączonej do komunikatu Ministra Edukacji i Nauki) jest następująca:

- (O1) 70 punktów – konferencja Conference on Computer Science and Intelligent Systems
- (O2) 40 punktów – czasopismo Applied Artificial Intelligence
- (O3) 20 punktów – konferencja International Symposium on Foundations of Intelligent Systems
- (O4) 140 punktów – czasopismo Engineering Applications of Artificial Intelligence
- (O5) 140 punktów – czasopismo International Journal of Approximate Reasoning
- (O6) 70 punktów – konferencja IEEE Conference on Systems, Man and Cybernetics

- (O7) 70 punktów – konferencja IEEE International Conference on Big Data  
(O8) 200 punktów – czasopismo Information Sciences  
(O9) 20 punktów – jako rozdział w monografii  
(O10) 200 punktów – czasopismo Information Sciences  
(O11) 200 punktów – czasopismo Information Sciences

Za najważniejsze wyniki przedstawione w opublikowanych pracach uważam:

- Opracowanie nowych algorytmów redukcji zbioru atrybutów (redukty aproksymacyjne, mRMR), wykorzystujące jako kryterium stopu statystyczny test permutacyjny.
- Opracowanie algorytmu generowania reduktów decyzyjnych przy wykorzystaniu metod grupowania atrybutów w tablicach decyzyjnych.
- Przedstawienie metody mierzenia stopnia wymienialności atrybutów w kontekście zadanego problemu predykcyjnego.
- Nowe, efektywne obliczeniowo algorytmy konstruowania bireduktów decyzyjnych.
- Metody skalowalnej ekstrakcji cech szeregów czasowych zaimplementowane w modelu obliczeń MapReduce.

Patrząc na całkowity dorobek publikacyjny, Habilitant jest autorem lub współautorem ponad 75 artykułów naukowych opublikowanych w czasopismach oraz w materiałach pokonferencyjnych międzynarodowych konferencji naukowych. Wskaźniki naukometryczne zostały przedstawione poniżej na podstawie trzech baz danych: Scopus, Web of Science oraz Google Scholar.

Baza danych	Liczba publikacji	Cytowania łącznie	Cytowania bez autocytowań	Indeks Hirscha
Scopus	68	891	673	17
Web of Science	50	530	436	12
Google Scholar	72	1197	-	18

Uważam, że indeks Hirscha jest wysoki jak na kandydata do stopnia doktora habilitowanego w dyscyplinie informatyka techniczna i telekomunikacja.

Habilitant pełnił, w latach 2015 – 2017, rolę współprzewodniczącego komitetu programowego International Symposium Advances in Artificial Intelligence and Applications (AAIA, część konferencji FedCSIS). Habilitant pełnił również rolę współprzewodniczącego komitetu programowego International Joint Conference on Rough Sets (IJCRS) w roku 2016.

Natomiast w latach 2017 – 2019 pełnił rolę członka komitetu programowego tej cyklicznej konferencji. Był przewodniczącym komitetów organizacyjnych sesji specjalnych dotyczących organizowanych przez niego konkursów eksploracji danych przy konferencjach FedCSIS (2014 – 2022), IEEE BigData (2019 – 2022), ISMIS (2017), IJCRS (2015) oraz PAKDD (2015).

Habilitant po otrzymaniu stopnia naukowego doktora brał udział w szeregu projektów badawczo-rozwojowych finansowanych przez Narodowe Centrum Badań i Rozwoju (NCBR) oraz Centrum Rozwoju Małych i Średnich Przedsiębiorstw (PARP). Najważniejsze, moim zdaniem, projekty zostały wymienione na poniższej liście, wraz z informacją o roli pełnionej przez Habilitanta:

- ICRA: Nowoczesne narzędzia inżynierskie do wspomaganie decyzji przeznaczone dla dowódców podczas działań ratowniczo-gaśniczych PSP w obiektach budowlanych. Projekt finansowany przez NCBR i realizowany w latach 2013 – 2015 przez konsorcjum w składzie: Szkoła Główna Służby Pożarniczej w Warszawie, Wydział MIMUW, Centrum Naukowo-Badawcze Ochrony Przeciwpożarowej - Państwowy Instytut Badawczy oraz firma Dituel. Habilitant pełnił rolę wykonawcy.
- DISESOR: Zintegrowany, szkieletowy system wspomaganie decyzji dla systemów monitorowania procesów, urządzeń i zagrożeń. Projekt finansowany przez NCBR i realizowany w latach 2015 – 2016 przez konsorcjum w składzie: Instytut Technik Innowacyjnych EMAG, Wydział Mechaniczny Technologiczny Politechniki Śląskiej, Wydział MIMUW oraz firma Sevitel Sp. z o.o. Habilitant pełnił rolę wykonawcy.
- Toolbox4HR: Opracowanie innowacyjnego systemu automatyzującego wczesne procesy rekrutacji opartego o samouczące się algorytmy i Big Data. Projekt finansowany przez PARP i realizowany w latach 2016 – 2017 przez konsorcjum w składzie: firma Toolbox for HR Sp. z o.o. oraz MIMUW. Habilitant pełnił rolę wykonawcy.
- Grail: Opracowanie modułowego narzędzia do implementacji i optymalizacji zaawansowanej sztucznej inteligencji szerokiego zastosowania w grach wideo. Projekt finansowany przez NCBR i realizowany w latach 2017 – 2018 przez firmę Silver Bullet Labs Sp. z o.o. Habilitant pełnił rolę wykonawcy i był członkiem kadry kluczowej projektu.
- SENSEI: System wspomagający rozwój umiejętności graczy. Projekt finansowany przez NCBR i realizowany w latach 2018 – 2020 przez firmę Esensei Sp. z o.o.

Habilitant pełnił rolę kierownika naukowego części badawczej i był członkiem kadry kluczowej projektu.

- LITL: Labelling in the Loop. Opracowanie opartego na uczeniu maszynowym systemu automatycznego i półautomatycznego oznaczania danych w dużych zbiorach danych. Projekt finansowany przez NCBR i realizowany w latach 2020 – 2022 przez firmę QED Software. Habilitant pełnił rolę wykonawcy badań i był członkiem kadry kluczowej projektu.
- BrightBox: Narzędzia klasy explainable AI służącego do poprawy interpretowalności i przewidywalności działania metod uczących oraz diagnostyki poprawności działania wyuczonych modeli AI/ML. Projekt finansowany przez NCBR i realizowany w latach 2020 – 2022 przez firmę QED Software. Habilitant pełni rolę kierownika badań i był członkiem kadry kluczowej projektu.

Habilitant od 2012 roku jest członkiem International Rough Set Society. Od 2014 roku jest członkiem Polskiego Towarzystwa Informatycznego. Jest także członkiem europejskiej organizacji COST -- European Cooperation in Science and Technology. W ramach współpracy z tą organizacją prowadził zajęcia podczas letniej szkoły Advanced Big Data Training School for Life Sciences w Barcelonie. Habilitant brał udział w stażu naukowym w Kanadzie na University of Halifax. W okresie lipiec – sierpień 2017 roku, pracował w grupie badawczej Prof. Stana Matwina nad metodami uczenia się reprezentacji obiektów w zbiorach danych.

### **Konkluzja**

Podsumowując, uważam, że zgłoszony przez Habilitanta powiązany tematycznie cykl jedenastu publikacji pt. „Wybrane metody reprezentacji obiektów i pojęć w danych dla algorytmów uczenia maszynowego” wykazuje znaczny wkład Habilitanta w rozwój dyscypliny naukowej informatyka techniczna i telekomunikacja w zakresie metod sztucznej inteligencji. Jednocześnie stwierdzam, że Habilitant spełnił wymagania obowiązującej Ustawy niezbędne do uzyskania stopnia naukowego doktora habilitowanego w dziedzinie nauk inżyniersko – technicznych w dyscyplinie informatyka techniczna i telekomunikacja.



Jarosław Stepaniuk