



RECENZJA

osiągnięcia naukowego dr inż. Wojciecha Macyny nt.
**„Przechowywanie oraz manipulowanie danymi z uwzględnieniem
nowoczesnych rodzajów pamięci”**
*(na podstawie cyklu 12 publikacji powiązanych tematycznie
w związku z wystąpieniem o nadanie stopnia naukowego doktora
habilitowanego w dziedzinie nauk inżynieryjno-technicznych w dyscyplinie
„Informatyka techniczna i telekomunikacja”
oraz ocena jego pozostałych osiągnięć naukowo-badawczych,
dydaktycznych i organizacyjnych*

1. Wstęp

Niniejsza recenzja została przygotowana na podstawie Uchwały nr 302/14/RDND03/2024-2028 Rady Dyscypliny Naukowej Informatyka Techniczna i Telekomunikacja Politechniki Wrocławskiej z dnia 25 września 2025 roku w sprawie powołania komisji habilitacyjnej w postępowaniu o nadanie stopnia doktora habilitowanego, która to uchwała wyznaczyła mnie na recenzenta w w/w postępowaniu.

Dr inż. Wojciech Macyna w 1997 roku ukończył studia magisterskie z zakresu informatyki o specjalności systemy informacji naukowo-technicznej prowadzone na Wydziale Informatyki i Zarządzania Politechniki Wrocławskiej. Stopień doktora nauk technicznych w dyscyplinie informatyka uzyskał w 2004 roku na tymże Wydziale za rozprawę poświęconą weryfikacji dynamicznych więzów integralności specyfikowanych w logice temporalnej z operatorami metrycznymi. Od roku 2005 do chwili obecnej jego miejscem zatrudnienia jest Politechnika Wroclawska, gdzie dr inż. Macyna był najpierw asystentem (do 2009 roku), później adiunktem (do roku 2013), następnie st. wykładowcą (w latach 2013-2020), a później znów adiunktem.

WPLYNĘŁO
12-01-2026

RDND03/9/2026

2. Ocena osiągnięcia naukowego dr inż. Wojciecha Macyny

W skład monotematycznego cyklu publikacji dokumentującego osiągnięcie naukowe nt. „Przechowywanie oraz manipulowanie danymi z uwzględnieniem nowoczesnych rodzajów pamięci” wchodzi 12 publikacji w czasopiśmie i materiałach konferencyjnych o łącznej liczbie 117 stron. Zgodnie z autoreferatem Habilitanta, do artykułów tego cyklu zaliczono następujące prace (w przytoczonym spisie podano współczynnik wpływu **IF** czasopisma i liczbę cytowań w bazach Scopus i Google Scholar (**GS**), jak również punkty MNiSW przydzielone publikacji):

[A1] Jacek Cichon and Wojciech Macyna. "Approximate Counters for Flash Memory". In: *17th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications, RTCSA 2011, Toyama, Japan, pp.185-189*
ranking core: B, 70 pkt MNiSW, wkład: 50%, GS: 23, Scopus: 7

[A2] Maciej Pawlik and Wojciech Macyna. "Implementation of the Aggregated R-Tree over Flash Memory". In: *Database Systems for Advanced Applications - 17th International Conference, DASFAA 2012, Int. Workshops, Busan, South Korea. LNCS, Vol. 7240, 2012, pp.65-72*
ranking core: B, 70 pkt MNiSW, wkład: 50%, GS: 20, Scopus: 12

[A3] Wojciech Macyna and Krzysztof Majcher. "Cost-based storage of the R-tree aggregated values over flash memory". In: *2018 Int. Conference on Industrial Enterprise and System Engineering (ICoESE 2018), 2019, pp. 97-102*
20 pkt MNiSW, wkład: 80%, GS: 2, Scopus: 0

[A4] Macyna Wojciech and Kukowski Michal. "Flash-aware Clustered Index for Mobile Databases". In: *2018 Int. Conference on Industrial Enterprise and System Engineering (ICoESE 2018), 2019, pp. 25-30*
20 pkt MNiSW, wkład: 50%, GS: 0, Scopus: 0

[A5] Wojciech Macyna and Michal Kukowski. "Partially Indexing on Flash Memory". In: *Database and Expert Systems Applications - 30th Int. Conference, DEXA 2019, Linz, Austria. LNCS, vol. 11706, 2019, pp. 95-105.*
ranking core: B, 70 pkt MNiSW, wkład: 50%, GS: 1, Scopus: 1

[A6] Krzysztof Kwiatkowski and Wojciech Macyna. "CFTL - Flash Translation Layer for Column Oriented Databases". In: *Intelligent Information and Database Systems - 5th Asian Conference, ACIIDS 2013, Kuala Lumpur, Malaysia. LNCS, vol. 7802, 2013, pp. 146-155*
ranking core: B, 70 pkt MNiSW, wkład: 50%, GS: 1, Scopus: 0

[A7] Wojciech Macyna and Michal Kukowski. "Flash-Aware Storage of the Column Oriented Databases". In: *Fundam. Informaticae 173.1 (2020), pp. 47-72.*
IF: 1,333, 70 pkt MNiSW, wkład: 50%, GS: 2, Scopus: 0

[A8] Wojciech Macyna and Michal Kukowski. "Bulk Loading of the Secondary Index in LSM-Based Stores for Flash Memory". In: *New Trends in Database and Information Systems – ADBIS 2022 Short Papers, Doctoral Consortium and Workshops. Turin, 2022. Communications in Computer and Information Science, Springer, vol. 1652, 2022, pp.133-143*
ranking core: B, 70 pkt MNiSW, wkład: 50%, GS: 0, Scopus: 0

[A9] Wojciech Macyna, Michal Kukowski, and Michal Zwarzko. "Multi-core Adaptive Merging of the Secondary Index for LSM-Based Stores". In: *Database and Expert Systems Applications – 34th Int. Conference, DEXA 2023, Penang, Malaysia, 2023*. LNCS, vol. 14147, pp. 245–257. doi: 10.1007/978-
ranking core: B, 70 pkt MNiSW, wkład: 50%, GS: 2, Scopus: 1

[A10] Wojciech Macyna and Carlos Ordonez. "A Bloom Filter Hierarchy for Non-key Search in Key-Value Stores". In: *2024 IEEE Int. Conference on Big Data (BigData)*. 2024, pp. 3712–3719
ranking core: B, 70 pkt MNiSW, wkład: 80%, GS: 0, Scopus: 0

[A11] Maciej Jurga and Wojciech Macyna. "Implementation of the Aggregated R-Tree for Phase Change Memory". In: *Database and Expert Systems Applications - 29th International Conference, DEXA 2018, Regensburg, Germany, 2018*. LNCS, vol. 11030, 2018, pp. 301–309
ranking core: B, 70 pkt MNiSW, wkład: 50%, GS: 2, Scopus: 2

[A12] Wojciech Macyna and Michal Kukowski. "Adaptive Merging on Phase Change Memory". In: *Fundam. Informaticae* 188.2 (2022), pp. 103–126
IF: 0,8, 70 pkt MNiSW, wkład: 50%, GS: 2, Scopus: 2

Wszystkie prace mają przydzielone punkty MNiSW. Dają one razem 740 punktów. Tylko dwa artykuły zostało opublikowanych w czasopismach z bazy JCR posiadających Impact Factor, z sumarycznym IF cyklu wynoszącym tylko 2,13. Obydwa artykuły opublikowano w czasopiśmie *Fundamenta Informaticae* należących do kwartyłu Q3. Najnowsza wartość wskaźnika Impact Factor dla tego czasopisma z czerwca 2025 r. to tylko 0.4, co świadczy mało pozytywnie o jego randze. Pozostałe dziesięć prac opublikowano w materiałach konferencyjnych, przy czym 8 z nich posiada kategorię B, co przekłada się na 70 punktów MNiSW za pracę.

Sumaryczna liczba cytowań prac cyklu według bazy Google Scholar wynosi 55, przy czym na dwie pierwsze prace (**A1** i **A2**), zresztą najwcześniejsze chronologicznie, przypada zdecydowana większość tych cytowań, bo aż 43. Pozostałe dziesięć późniejszych prac dorobiło się tylko 12 cytowań, co oznacza praktycznie jedno cytowanie przypadające w średnim na pracę. Spośród tych 12 cytowań, pięć cytowań to autocytowania. W rezultacie, dziesięć prac, od **A3** do **A12**, dorobiło się tylko 7 cytowań w pracach innych autorów, co zdecydowanie negatywnie świadczy o ich akceptacji i rozpoznawalności w międzynarodowej społeczności naukowej zajmującej się daną tematyką. Wniosek ten potwierdza również fakt, że trzy prace (**A4**, **A8**, **A10**) nie były w ogóle cytowane w literaturze, nawet przez samego habilitanta. Co symptomatyczne, również trzy inne prace cyklu nie były cytowane przez habilitanta. Oznacza to, że aż sześć prac cyklu, czyli połowa, nie było w ogóle cytowanych przez samego kandydata. Fakt ten świadczy negatywnie o spójności aktywności naukowej habilitanta. Z takim brakiem szacunku dla własnych prac spotkałem się po raz pierwszy w mojej długoletniej praktyce recenzowania wniosków habilitacyjnych.

Przechodząc do oceny merytorycznej powyższego cyklu publikacji, chciałbym stwierdzić, że generalnie pozytywnie oceniam aktualność i ważność tematyki badawczej podjętej w nim. Tematyka ta jest skoncentrowana na opracowaniu efektywnych algorytmów i technik przechowywanie oraz manipulowanie danymi z

uwzględnieniem nowoczesnych rodzajów pamięci. Właśnie ten aspekt technologiczny w postaci uwzględnienia nowych nośników pamięci uzasadnia potencjalnie innowacyjny charakter badań prowadzonych przez habilitanta przez okres około 15 lat. Wychodzą one naprzeciw wyzwaniom związanym z pojawieniem się takich rodzajów pamięci jak półprzewodnikowe pamięci *flash* oraz pamięci typu *phase change memory* (PCM). Oba te rodzaje reprezentują pamięć nieulotną (*non-volatile memory*), co jest cenną właściwością z punktu widzenia wielu aplikacji, przy czym pierwszy z nich charakteryzuje się dostępem blokowym, a drugi – adresacją bajtową. O ile pierwszy z tych nośników znalazł już szerokie zastosowanie przy budowie dysków SSD, stosowanych zarówno w powszechnie używanych systemach komputerowych typu laptopy czy desktopy, jak i platformach serwerowych, to drugi z nich znajduje się faktycznie na początku swojej drogi do szerokiego zastosowania w praktyce. Droga ta nie jest prosta, o czym świadczy chociażby fakt, iż produkcja układów 3D-Xpoint firmy Intel, w których po raz pierwszy technologia PCM została zastosowana komercyjnie, była zarzucona stosunkowo szybko pomimo wstępnych, gromkich deklaracji o ich ogromnym potencjale aplikacyjnym. Tym niemniej, nie odstraszyło to innych graczy od prac nad opracowaniem i wprowadzeniem na rynek układów pamięci opartych o technologię PCM. Wśród wyzwań związanych z wdrożeniem tych nowych rodzajów pamięci należy wymienić ograniczoną żywotność układów flash oraz znacznie dłuższy czas zapisu niż odczytu informacji, charakteryzujący oba rodzaje pamięci. Innym aktywnym obszarem badań naukowych w obszarze przechowywania i manipulowania danymi, rozpatrywanym przez dr. Macynę, jest rozwój algorytmów i technik opartych o wykorzystanie drzew LSM (*log-structured merge trees*) – popularnej struktury danych stosowaną w wielu systemach NoSQL, w tym MongoDB i Cassandra.

Publikacje przedłożonego cyklu publikacji dotyczą więc trzech obszarów tematycznych a mianowicie:

1. **metod i technik przechowywanie danych w pamięci flash (prace A1-A7);**
2. **algorytmów i technik dostępu do danych z wykorzystaniem drzew LSM (prace A8-A10);**
3. **metody i techniki indeksowania danych w pamięci PCM (prace A11-A12).**

Obszar 1 (prace A1 – A7):

W obszarze tym stopniem akceptacji przez międzynarodową społeczność badawczą zdecydowanie wyróżniają się dwie najstarsze chronologicznie publikacje **A1**, **A2**. W pierwszej z nich, wychodząc naprzeciw problemom związanym z ograniczoną żywotnością pamięci flash, zaproponowano system liczników probabilistycznych do przechowywania stopnia zużycia bloków pamięci flash. Pozwoliło to znacznie zmniejszyć rozmiar licznika związanego z każdym blokiem pamięci flash oraz istotnie ograniczyć liczbę zmian w pamięci flash związaną z modyfikacją liczników. Co ważne, uzyskane wyniki zostały potwierdzone zarówno w sposób formalny, jak i eksperymentalny.

Natomiast w pracy **A2** skupiono się na opracowanie koncepcji przechowywania danych w pamięci flash z wykorzystaniem zagregowanych R-drzew. Ich zastosowanie wraz z uwzględnieniem charakterystyk pamięci flash umożliwiło

efektywną indeksację danych przechowywanych w pamięci flash. Opracowana koncepcja została rozwinięta w pracy **A3**, którą zaprezentowano na międzynarodowej konferencji w roku 2018, tj. aż sześć lat od momentu ukazania się pracy **A2**. W pracy **A3** zaproponowano metodę optymalizacji zagregowanych R-drzew pod względem wydajności dostępu do danych przechowywanych w pamięci flash. Wyniki uzyskane w pracy **A2** zostały również wykorzystane w artykule **A11**, o którym będzie mowa dalej. Należy tutaj zauważyć, że jest to wyjątkowy przypadek, gdy jakaś praca (mowa o **A2**) została zacytowana (i wykorzystana) w dwóch innych pracach kandydata. Pięć innych prac cyklu zostało przez niego zacytowanych tylko raz i aż sześć publikacji cyklu (tj. połowa) nie była w ogóle cytowana w innych pracach habilitanta.

Ten brak odwołania się do swoich poprzednich prac dotyczy w szczególności pracy **A4**, która w rzeczywistości kontynuuje tematykę indeksowania danych w pamięci flash. W pracy **A4**, którą zaprezentowano na tej samej konferencji co **A3**, zaproponowano nową metodę indeksowania zorientowaną na zastosowanie w mobilnych bazach danych. Pokazano, że metoda ta jest bardziej efektywna pod względem wymagań pamięciowych niż zastosowanie tradycyjnych B+-drzew. Nie wyjaśniono jednak, dlaczego nie wykorzystano opracowanej wcześniej koncepcji opartej na zagregowanych R-drzewach.

Praca **A4** nie była cytowana przez kogokolwiek, w tym przez samego habilitanta. Pod tym względem już korzystniej prezentują się publikacje **A5** i **A6**, które były choć raz cytowane przez samego kandydata przy braku cytowania przez innych autorów. W pierwszej z tych publikacji zaproponowano metodę indeksowania częściowego dla baz danych implementowanych w pamięci flash. Metoda ta pozwala zredukować liczbę operacji odczytu i kasowania danych, co prowadzi do redukcji czasu zapytań do bazy. W konkluzji pracy **A5**, opublikowanej w roku 2019, autorzy stwierdzają, że dane podejście oparte na indeksowaniu częściowym może otworzyć nowy kierunek badań. Tymczasem jednak nie znalazło ono odzewu w pracach innych autorów i zostało zacytowane tylko w innej pracy kandydata – w opublikowanym w roku 2022 artykule **A12** dotyczącym pamięci PCM, który został omówiony dalej. Natomiast w pracy **A6**, opublikowanej w roku 2013, zaproponowano sposób optymalizacji przechowywania informacji dla potrzeb kolumnowych baz danych wykorzystujących pamięci flash. Sposób ten oparto na wykorzystaniu warstwy pośredniej, zapewniającej zgodnie ze stwierdzeniem autorów efektywne zarządzanie pamięcią flash. I znów niestety mamy do czynienia z niespójnością w autoreferacie, gdzie kandydat stwierdza, że publikacja **A6** stanowi początek badań nad przechowywaniem danych kolumnowych w pamięci flash, przy czym drugą ich część stanowi opublikowana w roku 2020 praca **A7**, w której niestety sami autorzy w ogóle nie odwołują się do publikacji **A6**, jakby o niej zapomnieli. W pracy **A7**, ostatniej z publikacji wchodzących w skład pierwszego obszaru tematycznego cyklu, zaproponowano strukturę zwaną CF-drzewem (Column-Flash tree), na której oparto przechowywanie kolumnowych baz danych. Dzięki temu sposób przechowywania może być dostosowany do zmiennych wzorców zapytań oraz charakterystyk dysków flash.

Obszar 3 (prace A11 – A12):

Drugim nowoczesnym rodzajem pamięci rozpatrywanym w pracach kandydata jest pamięć PCM. Metody indeksowania danych przechowywanych w pamięci PCM stanowią temat badań, których wyniki przedstawiono w pracach **A11** oraz **A12**, z których pierwsza nawiązuje do publikacji **A2**, zaś druga – do publikacji **A5**, obu należących do pierwszego obszaru tematycznego. Podobnie jak **A2**, praca **A11**, opublikowana w 2018 r., również wykorzystuje zagregowane R-drzewa do optymalizacji przechowywania i manipulacji danymi. Zaproponowana metoda indeksowania wykorzystuje pamięć podręczną PCM do przechowywania zagregowanych wartości, pozwalając na redukcję kosztownych operacji zapisu i dzięki temu przewyższa inne metody, w tym bazujące na kolejce FIFO. Należy jednak zauważyć, że uzyskane w **A11** wyniki, pomimo tego, że upłynęło już około 7 lat od momentu opublikowania, znalazły odzew tylko w opublikowanym w roku 2023 przeglądzie prac dotyczącym indeksowania danych dla nowoczesnych rodzajów pamięci (wśród 76 innych prac). Natomiast praca **A12**, opublikowana w roku 2022, rozwija koncepcję indeksowania częściowego, którą kandydat zaproponował w publikacji **A5** z roku 2019. Zaproponowana metoda indeksowania częściowego umożliwia indeksowanie tylko takich danych, które są wynikiem działania zapytań o zakres danych (tzw. *range queries*), co ogranicza zakres jej zastosowania. Jej efektywność, w tym możliwość adaptacji do charakterystyk obciążenia, została zbadana i potwierdzona eksperymentalnie z wykorzystaniem autorskiego symulatora pamięci PCM. Niestety, w pracy brak informacji o dostępności kodu symulatora, na przykład, w ramach repozytorium github czy ewentualnie w innej formie, co w chwili obecnej staje się normą dla tego typu oprogramowania, pozwalając na odtwarzalność wyników (*reproducibility*). Być może jest to jedna z przyczyn, że praca **A12** została zacytowana tylko w jednej pracy autorów z Chin, datowanej rokiem 2024, w której wspomniano o metodzie zaproponowanej przy udziale habilitanta, lecz pierwszeństwo oddano wykorzystaniu B+-drzew.

Obszar 2 (prace A8 – A10):

W przypadku danego obszaru tematycznego rzuca się przede wszystkim w oczy słabe powiązanie z pracami w pozostałych dwóch obszarach przedłożonego cyklu, gdyż trzy prace wchodzące w skład tego obszaru nie są cytowane przez żadną z pozostałych publikacji cyklu, przy czym tylko jedna praca z pozostałych dwóch obszarów cyklu (**A12**) jest cytowana przez pracę wchodzącą w skład danego obszaru (publikację **A9**). Jest to o tyle dziwne, że już pierwsza pozycja, czyli **A8**, dotyczy pamięci flash rozpatrywanych w pierwszym obszarze tematycznym. Zaproponowano w niej metodę ułatwiającą ładowanie wielu elementów indeksu wtórnego na raz do drzewa LSM, zapewniającego indeksowanie danych znajdujących się w pamięci flash. Drzewa takie są popularną strukturą danych, stosowaną w wielu systemach NoSQL takich jak MongoDB czy Cassandra, które są z kolei wykorzystywane przez wiele popularnych serwisów (np. Facebook, LinkedIn, Instagram czy Twitter). Dane przechowywane na drzewie LSM są z reguły reprezentowane jako elementy typu klucz-wartość. Ponieważ często okazuje się jednak, że indeksowanie po kluczu nie jest wystarczające w tego typu bazach danych, konieczne jest stosowanie indeksu wtórnego dla danych niekluczowych. Metoda zaproponowana w pracy **A8**, opublikowanej w roku 2022, jest zoptymalizowana pod kątem pamięci flash. Efektywność działania metody

potwierdzono w sposób eksperymentalny dla różnego zestawu danych oraz różnych zakresów zapytań. Tym niemniej, dana praca nie znalazła odzewu w pracach innych badaczy i nie jest w ogóle cytowana w literaturze.

Brak cytowani charakteryzuje również pracę **A10**, chociaż jest on tutaj łatwiej wytłumaczalny, gdyż opublikowano ją stosunkowo niedawno – w roku 2024 jako jedyną wspólną publikację z badaczem spoza macierzystej uczelni. W publikacji tej zaproponowano wykorzystanie hierarchicznych filtrów Blooma do wyszukiwania elementów niekluczowych w bazach danych opartych o drzewo LSM. Dzięki zintegrowaniu tych filtrów z drzewem LSM liniowy czas przeszukiwania wartości niekluczowych na drzewie LSM skraca się do logarytmicznego bez potrzeby zastosowania osobnej struktury dla indeksu wtórnego. Przeprowadzone eksperymenty pokazały, że czas przeszukiwania jest nawet o 80% krótszy od czasu przeszukiwania bez użycia proponowanej metody. Co istotne, autorzy stwierdzili, że metoda ta może zostać użyta wszędzie tam, gdzie konieczne jest przeszukiwanie w nieuporządkowanym zbiorze danych przechowywanym w wielu plikach.

Ostatnia z analizowanych prac, artykuł **A9** pochodzący z roku 2023 i również dotyczący wykorzystania drzew LSM do przechowywania danych w bazach NoSQL, jest bliski tematycznie uprawianej przeze mnie problematyce obliczeń równoległych. Artykuł dotyczy zastosowania architektur wielordzeniowych do implementacji indeksowania częściowego dla baz danych opartych na drzewach LSM. Indeksowanie takie, zdaniem autorów użyte po raz pierwszy dla generowania indeksu wtórnego dla drzew LSM, okazuje się stosunkowo dobrze dostosowane do użycia architektury wielordzeniowej wykorzystującej nawet kilkadziesiąt wątków. Z mojej perspektywy, szereg aspektów pracy budzi jednak wątpliwości. W szczególności, brak w niej jakichkolwiek szczegółów dotyczących implementacji przetwarzania/programowania równoległego. Wykresy pokazujące uzyskaną skalowalność czasu obliczeń są mało czytelne, a w pracy brak jest jakichkolwiek danych liczbowych na temat uzyskanych przyspieszeń. Nie podano typu zastosowanego procesora (tylko nazwę serwera), co utrudnia identyfikację wykorzystywanego CPU. Na podstawie danych dostępnych w sieci można dokonać takiej identyfikacji, z której wynika, że procesory w tym serwerze mogą mieć co najwyżej 10 rdzeni, a nie 20, jak twierdzą autorzy, którzy najprawdopodobniej mieli na myśli liczę wątków dostępnych dzięki trybowi hyperthreading. W oczy rzuciło mi się również niezgodne z rzeczywistością bardzo duże wartości objętości pamięci podręcznych L1 i L2 jako odpowiednio 640 KB oraz 2560 KB, co szczególnie razi w przypadku pamięci L1, której pojemność kształtują się w granicach od kilku do 128 KB na rdzeń. Wyjaśnienie okazało się dostatecznie proste – w obu przypadkach dopisane zostało zbędne zero do podanych wartości. Już mniejszą wagę ma nazwanie podanych w pracy wartości jako odnoszących się do linii pamięci cache, a nie całych pamięci podręcznych. Pojemności tych linii są jeszcze mniejsze – rzędu kilkudziesięciu bajtów dla linii pamięci L1. Wskazane błędy wydawać się mogą błaha, świadczą jednak niezbyt pozytywnie o jakości recenzowania prac przyjmowanych na konferencję DEXA 2023 (Penang, Malezja), w której materiałach praca **A9** została opublikowana.

Podsumowanie oceny jednotematycznego cyklu publikacji

Przedłożony cykl 12 publikacji tworzy całość dotyczącą ważnej i aktualnej problematyki badawczej ukierunkowanej na zwiększenie efektywności procesów przechowywania i przetwarzania danych z uwzględnieniem wyłaniających się technologii pamięci. W publikacjach tych habilitant przedstawił metody i technik przechowywania danych w pamięci flash (prace **A1-A7**) oraz metody i techniki indeksowania danych w pamięci PCM (prace **A11-A12**), a także algorytmów i technik dostępu do danych z wykorzystaniem drzew LSM (prace **A8-A10**). Zakres uzyskanych wyników jest dostatecznie szeroki, przy czym zostały one otrzymane w ciągu dostatecznie długiego okresu czasu obejmującego około 15 lat. Niełatwym zadaniem jest w tych okolicznościach obiektywna ocena znaczenia tych wyników dla dyscypliny informatyka techniczna i telekomunikacja.

W takiej sytuacji na pomoc przychodzi analiza, z jednej strony, rangi i jakości wydawnictw, w których ukazały się pozycje cyklu. Odpowiednio wysoka ranga wydawnictw wraz z towarzyszącym procesem recenzowania niejako gwarantuje adekwatny poziom wydawanych prac. W przypadku przedłożonego cyklu żadna z prac nie ukazała się w czasopiśmie o dostatecznie wysokim wskaźniku Impact Factor (w przypadku prac **A7** i **A12**) czy też w materiałach konferencji kategorii **A** lub **A+**. W rezultacie żadnej z prac nie osiągnęła progu 100 punktów w ranking MNIŚW, nie mówiąc już o wyższej punktacji 140 lub 200 punktów. Niezbyt pozytywną ocenę materiałów konferencyjnych, w których opublikowane zostało 10 prac cyklu, potwierdza omawiany wyżej przypadek pracy **A9**, której tekst został zaakceptowana na konferencję w postaci obarczonej jawnymi brakami i błędami.

Zdając sobie sprawę, że powyższy ranking nie stanowi absolutnej wyroczni, odwołamy się do drugiego sposobu oceny publikacji w postaci ich cytowalności jako wyznacznika poziomu naukowego oraz oddziaływaniu na badania naukowe i tym samym wkładu do aktualnego stanu wiedzy w dyscyplinie. W tym przypadku jedynie dwie pierwsze prace cyklu, opublikowane w latach 2011-2012 uzyskały relatywnie wysoką cytowalność – odpowiednio 23 i 20 cytowań, wszystkie w pracach innych autorów. Co warto zauważyć, pierwsza z tych prac praktycznie nie znalazła odzewu w pracach związanych z tematyką nowych nośników pamięci (z wyjątkiem jednej pracy chińskich autorów) i tylko druga pozycja była cytowana przez badaczy zajmujących się tematyką związaną z pamięcią flash. **Pozostałe 10 prac cyklu, które ukazały się przestrzeni lat 2013-2024, charakteryzuje się bardzo niską cytowalnością – zostały one wspomniane tylko siedem razy w sześciu pracach innych autorów, co zdecydowanie negatywnie świadczy o stopniu ich akceptacji i rozpoznawalności w międzynarodowej społeczności naukowej zajmującej się daną tematyką badań.** Chciałbym również zwrócić uwagę na fakt, że połowa prac przedłożonego cyklu (aż sześć) nie było w ogóle cytowanych przez samego kandydata. Fakt ten świadczy negatywnie o spójności aktywności naukowej habilitanta, który mało aktywnie starał się znajdować wspólne wątki w swoich badaniach i wykorzystywał synergię już otrzymanych wyników do uzyskania nowych rezultatów.

Konkluzja oceny jednotematycznego cyklu publikacji

Podsumowując ocenę cyklu publikacji pod wspólnym tytułem „*Przechowywanie oraz manipulowanie danymi z uwzględnieniem nowoczesnych rodzajów pamięci*” należy stwierdzić, że składające się na ten cykl prace dotyczą ważnego i

aktualnego problemu badawczego. Jednakże na podstawie analizy poziomu naukowego tych prac i stopnia ich oddziaływania na badania naukowe i wkładu do stanu wiedzy w rozpatrywanym obszarze stwierdzam, że opiniowany cykl publikacji nie stanowi istotnego wkładu w rozwój dyscypliny naukowej „**Informatyka techniczna i telekomunikacja**”, a tym samym nie spełnia wymagań stawianych ustawowo osiągnięciom w przewodzie habilitacyjnym.

3. Ocena istotnej aktywności naukowej w tym realizowanej w więcej niż jednej uczelni lub instytucji naukowej, w szczególności zagranicznej

Dorobek naukowo-badawczy

Oprócz oceny samego cyklu publikacji, warto odnotować następujące statystyki dokumentujące osiągnięcia naukowo-badawcze Habilitanta po uzyskaniu stopnia doktora w postaci dorobku publikacyjnego:

- Liczba punktów wg listy MNiSW uzyskanych z cyklu publikacji zgłoszonego jako osiągnięcie naukowe Habilitanta: 740
- Sumaryczny Impact Factor dla osiągnięcia naukowego: 2,13
- Sumaryczna liczba cytowań dla osiągnięcia naukowego: 52 oraz 25 według Google Scholar oraz Scopus, odpowiednio
- Sumaryczny Impact Factor publikacji: 2,13
- Liczba publikacji po doktoracie: 24, w tym:
 - monografie: 0
 - artykuły w czasopiśmie znajdujących się w bazie JCR: 2
 - artykuły w materiałach konferencji międzynarodowych: 17
 - rozdziały w monografiach: 5
- Indeks Hirscha (Scopus): 3
- Indeks Hirscha (Google Scholar): 4
- Liczba cytowań (Scopus): 35, w tym 29 bez autocytowań
- Liczba cytowań (Google Scholar): 64

Po doktoracie Habilitant opublikował tylko 24 prac, co biorąc pod uwagę okres, który upłynął od doktoratu (21 lat), daje niewiele więcej niż jedną (dokładnie 1,14) pracę na rok, co świadczy o relatywnie niskiej efektywności publikacyjnej kandydata, przy czym w bazie Google Scholar brak prac kandydata w latach 2005-2010, a także 2013-2014, 2017 i 2021, tj. w ciągu 10 spośród rozpatrywanych 21 lat. Istotną słabością publikacji habilitanta jest fakt, że tylko 2 prace zostały opublikowane w czasopiśmie z Impact Factor, przy czym w czasopiśmie o niskiej wartości współczynnika IF, które lokuje się w trzecim kwartylu.

Wśród tych opublikowanych prac Habilitant wydzielił sześć publikacji **B1-B6** spoza przedłożonego cyklu, które można potraktować jako tzw. drugie osiągnięcie naukowe. Są to następujące prace:

[B1] L. Bellatreche, F. Djellali, W. Macyna, and C. Ordonez. “Energy-Aware Query Processing: A Case Study on Join Reordering”. In: IEEE Int. Conference on Big Data - BigData 2023, Sorrento, Italy. IEEE, 2023, pp. 3743–3752

ranking core: B, GS: 4

[B2] L. Krzywiecki, K. Majcher, and W. Macyna. “Efficient Probabilistic Methods for Proof of Possession in Clouds”. In: Data Mining and Big Data, First Int. Conference, DMBD 2016, Bali, Indonesia. LNCS, Springer, 2016, pp. 364–372.

GS: 2

[B3] C. Ordonez and W. Macyna. “Optimizing Energy Consumed by Analytics in the Cloud”. In: 2024 IEEE Int. Conference on Big Data (BigData). IEEE, 2024, pp. 5201–5210.

ranking core: B, GS: 1

[B4] C. Ordonez, W. Macyna, and L. Bellatreche. “Data engineering and modeling for artificial intelligence”. In: Data Knowl. Eng. 153 (2024), 2 strony

GS: 1

[B5] C. Ordonez, W. Macyna, and L. Bellatreche. “Energy-Aware Analytics in the Cloud”. In: Proc. Int. Workshop on Big Data in Emergent Distributed Environments, BiDEDE 2024, Santiago, Chile, June 9-15, 2024. ACM, 2024, 6 stron.

ranking core: A+, GS:4

[B6] C. Ordonez, R. Varghese, N. Phan, and W. Macyna. “Growing a FLOWER: Building a Diagram Unifying Flow and ER Notation for Data Science”. In: Proc. 2024 Workshop on Human-In-the-Loop Data Analytics, HILDA 24, Santiago, Chile, 14 June 2024. ACM, 2024, 6 stron

ranking core: A+, GS: 1

Przy rozpatrzeniu tych prac rzuca się w oczy, że króciutka pozycja **B4** to tzw. editorial czyli po prostu wstęp do zbioru artykułów innych autorów i nie może być traktowana jako pełnowartościowa publikacja naukowa. Tematyka pozostałych pięciu prac jest niejednorodna i dotyczy aż trzech różnych zagadnień. Pierwsze z nich to weryfikacja poprawności przechowywania danych w systemach chmurowych z wykorzystaniem funkcji haszujących. Tematykę tę reprezentuje tylko praca **B2**, o której innowacyjności w momencie opublikowania trudno sądzić, gdyż nie zawiera ona żadnych porównań pod względem efektywności zaproponowanej metody z innymi rozwiązaniami, w tym szczególnie opartymi na metodach kryptograficznych czy kodach nadmiarowych. Dodajmy, że bibliografia w tej pracy zwiera tylko trzy pozycje, do tego dostatecznie stare już w momencie publikacji pozycji **B2**.

Drugie z rozpatrywanych zagadnień dotyczy zarządzania zbiorami danych w projektach naukowych w obszarze Data Science i reprezentowane jest tylko przez jedną pracę **B6**. Kandydat zgłosił ją jako opublikowaną w materiałach konferencji najwyższej kategorii A+, co jednak w mojej opinii nie jest w pełni uzasadnione. Kategoria A+ faktycznie dotyczy tylko głównej konferencji ACM SIGMOD/PODS, podczas gdy publikacja **B6** została opublikowana w ramach tzw. *companion*, czyli towarzyszącego zasadniczej konferencji workshopu HILDA'24 (2024 Workshop on Human-In-the-Loop Data Analytics), co dużo bardziej odpowiada jakości tej pracy. Zaprezentowana w niej koncepcja zunifikowanych diagramów została bowiem

przedstawiona na tyle ogólnie i mało konkretnie, że trudno ocenić jej efektywność i zakres zastosowania w praktyce. Dotyczy to w szczególności możliwości jej wykorzystania w zagadnieniach uczenia maszynowego, co deklarują autorzy w konkluzji pracy.

Trzecie z zagadnień obejmuje coraz bardziej istotną problematykę zużycia energii w nowoczesnych systemach informatycznych i przedstawione jest przez trzy prace – **B1**, **B3** oraz **B5**. W przypadku tej ostatniej pracy mamy do czynienia z podobną sytuacją jak w przypadku pozycji **B6**, gdyż **B5** zaprezentowano nie na zasadniczej konferencji ACM SIGMOD/PODS rangi Core A+, lecz w ramach towarzyszącego jej workshopu BiDEDE 2024 (Workshop on Big Data in Emergent Distributed Environments), co również dużo bardziej odpowiada jakości pracy **B5**. Zawiera ona przegląd wyzwań i podejść powszechnie znanych w szerokim środowisku badaczy, którzy od ponad 20 lat intensywnie zajmują się problematyką efektywności energetycznej systemów informatycznych. Podobnie przeglądowy charakter ma praca **B3**. Jedynie w publikacji **B1** podjęto próbę uwzględnienia specyfiki relacyjnych baz danych wykorzystywanych np. w hurtowniach danych, gdzie często występuje potrzeba łączenia wielu tabel bazodanowych w jednym zapytaniu i zweryfikowano różne techniki łączenia tabel w relacyjnych bazach danych pod kątem efektywności zapytań oraz zużycia energii. Chociaż wytyczony w ten sposób kierunek badań może stanowić podstawę do opracowania bardziej kompleksowego podejścia, w tym mniej lub bardziej sformalizowanych metod optymalizacji zużycia energii w tego typu aplikacjach, możliwość ta nie znalazła do tej pory odzwierciedlenia w badaniach prowadzonych przez habilitanta.

Podsumowując więc moją analizę sześciu prac **B1-B6** dr Macyny, które nie weszły w skład przedłożonego cyklu dwunastu prac **A1-A12**, muszę niestety stwierdzić, że nie zasłużyły one na moją pozytywną ocenę zarówno pod względem oryginalnych osiągnięć naukowych, jak też ich znaczenia dla praktyki wykorzystania systemów informatycznych. Należy przy tym przyznać, że prace te (oraz wcześniej oceniana publikacja **A10**) pozwalają spełnić wymóg formalny, aby istotna działalność naukowa kandydata była realizowana w więcej niż jednej uczelni lub instytucji naukowej. Prace **B1**, **B3**, **B5** i **B6** oraz publikacja **A10** powstały bowiem w wyniku współpracy międzynarodowej z Houston University of Technology i ISAE ENSMA w Poitiers.

W recenzowanych przez mnie dotychczas wnioskach habilitacyjnych istotny element stanowiła aktywność badawcza i rozwojowa kandydata w projektach, z reguły powiązana z zasadniczą działalnością naukową habilitanta. W początkach swojej kariery naukowej po uzyskaniu stopnia doktora (lata 2005-2012) dr Macyna uczestniczył jako wykonawca w trzech projektach, Były to:

1. „Informatyczny system zarządzania dla małych i średnich przedsiębiorstw zintegrowany z modułem wspomaganie decyzji” w ramach Sektorowego Programu Operacyjnego Wzrost Konkurencyjności Przedsiębiorstw, beneficjenci: Comarch i Akademia Ekonomiczna w Poznaniu, okres realizacji: IX 2005 – VI 2007. Rola kandydata: zaprojektowanie architektury nowego systemu ERP o nazwie Altum dla Comarch.
2. „Efektywny system wspomaganie decyzji oparty o controllingową hurtownię danych” w ramach Sektorowego Programu Operacyjnego Wzrost

Konkurencyjności Przedsiębiorstw, beneficjenci: Comarch, Politechnika Łódzka, okres realizacji IV.2007 – VIII.2008. Rola habilitanta: główny projektant hurtowni danych dla systemu Comarch ERP Altum.

3. projekt POIG 01.03.01-02-002/08 „Czujniki i sensory do pomiarów czynników stanowiących zagrożenia w środowisku - modelowanie i monitorowanie zagrożeń” w ramach Programu Operacyjnego Innowacyjna Gospodarka, lata 2008–2012. Rola: wykonawca zadania poświęconego analizie on-line.

Niestety, projekty te nie miały charakteru badawczego i co najwyżej zawierały elementy prac rozwojowych. Co istotne, tematyka tych projektów wydaje się być niezwiązana z problematyką badań naukowych prowadzonych przez habilitanta, co silnie ogranicza ich znaczenie dla pozytywnej oceny istotnej działalności naukowej realizowanej przez dr. inż. Wojciecha Macynę. Ponadto, przez ostatnie 13 lat nie wykazano żadnej aktywności projektowej. Należy ubolewać, że kandydat nie wykorzystał możliwości, jakie dla działalności sensu stricto badawczej w rozpatrywanym obszarze, otwiera aktywność projektowo-technologiczna realizowana we współpracy z innymi instytucjami i szczególnie firmami. W szczególności, możliwości takie otwierał fakt, iż ulokowany w Gdańsku oddział firmy Intel zajmował się opracowaniem oprogramowania dla pierwszej komercyjnie dostępnej na rynku pamięci PCM w postaci 3D-Xpoint.

Ocena aktywności Habilitanta w rozpatrywanym obszarze byłaby niepełna bez uwzględnienia wygłoszenia w okresie po doktoracie 11 referatów na specjalistycznych międzynarodowych konferencjach naukowych, współprzewodniczenia komitetowi programowemu workshopu DEMAI 2024 oraz zredagowania wydania specjalnego czasopisma „Data & Knowledge Engineering” na podstawie prac zaakceptowanych na workshop DEMAI 2023, a także wykonania 2 recenzji prac dla zagranicznych czasopism naukowych (Cluster Computing i IEEE Access) oraz około 30 recenzji dla międzynarodowych konferencji naukowych. Biorąc pod uwagę czas, który upłynął od doktoratu, dorobek ten należy uznać za mocno umiarkowany.

Ocena dorobku dydaktycznego i organizacyjnego

Zgodnie z obowiązującą ustawą, zarówno działalność dydaktyczna, jak i organizacyjna, nie są obowiązkowymi elementami oceny wniosku habilitacyjnego. Tym niemniej zwyczajowo stanowią one fragment całości oceny dorobku kandydata. O ile dorobek organizacyjny dr. Macyny nie jest zbyt bogaty i ogranicza się do opieki nad praktykami studenckimi na kierunku „Informatyka Algorytmiczna” na Wydziale Informatyki i Telekomunikacji Politechniki Wrocławskiej, to dorobek dydaktyczny habilitanta zasługuje na zdecydowanie pozytywną ocenę. Dr Macyna był promotorem pomocniczym pracy doktorskiej mgr inż. Michała Kukowskiego: „Indeksowanie baz danych na nowoczesnych typach pamięci” obronionej w 2024 roku na Wydziale Informatyki i Telekomunikacji Politechniki Wrocławskiej. Był także promotorem około 30 prac magisterskich, przy czym wyniki trzech prac zostały opublikowane na konferencjach międzynarodowych, a kilka prac zostało wdrożonych w firmie Comarch. Zasadnicza działalność dydaktyczna kandydata była jednak związana z prowadzeniem zajęć z przedmiotów informatycznych w Politechnice

Wrocławskiej: na kierunku informatyka (I i II stopnia) wydziału Podstawowych Problemów Techniki, na kierunku Matematyka Stosowana oraz Informatyka Algorytmiczna na Wydziale Informatyki i Telekomunikacji. Były to zarówno wykłady z takich przedmiotów, jak bazy danych i systemy informacyjne, metody wytwarzania oprogramowania, technologie programowania, stream programming (w j. angielskim), aplikacje bazodanowe i zarządzanie projektami informatycznymi, jak również różnorakie ćwiczenia i laboratoria. Przygotował również sylabusy do następujących zajęć: kurs programowania, metody wytwarzania oprogramowania oraz aplikacje bazodanowe.

Konkluzja oceny istotnej aktywności naukowej

Po doktoracie habilitant opublikował tylko 24 prac, co biorąc pod uwagę okres, który upłynął od doktoratu (21 lat), daje niewiele więcej niż jedną (dokładnie 1,14) pracę na rok, co świadczy o relatywnie niskiej efektywności publikacyjnej kandydata, przy czym w bazie Google Scholar brak prac kandydata w ciągu dziesięciu spośród rozpatrywanych 21 lat. Istotną słabością publikacji habilitanta jest ich niska cytowalność ze współczynnikiem Hirscha równym 3 lub 4 w zależności od bazy, a także fakt, że tylko dwie prace zostały opublikowane w czasopiśmie z Impact Factor, przy czym w czasopiśmie o niskiej wartości współczynnika IF. Oprócz 12 prac wchodzących w skład jednotematycznego cyklu publikacji, dr Macyna przedłożył do oceny sześć innych prac, które nie zasłużyły jednak na moją pozytywną ocenę zarówno pod względem oryginalnych osiągnięć naukowych, jak też ich znaczenia dla praktyki wykorzystania systemów informatycznych. Działalność badawcza i publikacyjna dr Macyny nie znalazła wsparcia w realizowanych przez niego nielicznych projektach, które nie miały charakteru badawczego i co najwyżej zawierały elementy prac rozwojowych, a ich tematyka była niezwiązana z problematyką badań naukowych prowadzonych przez habilitanta. Również rozpoznawalność kandydata przez międzynarodową społeczność naukową, wyrażającą się m.in. w udziale w Komitecie Programowym jednej konferencji naukowej i redagowaniu numeru specjalnego jednego czasopisma, jest niewysoka. Na moją jednoznacznie pozytywną ocenę zasługuje tylko działalność dydaktyczna habilitanta. Jednocześnie stwierdzam, iż działalność naukowa dr inż. Wojciecha Macyny spełnia wymóg formalny, aby była realizowana w więcej niż jednej uczelni lub instytucji naukowej.

Podsumowując w tym kontekście ocenę aktywności naukowej dr. inż. Wojciecha Macyny po doktoracie stwierdzam zatem, że moja ocena całokształtu jego dorobku naukowego, w tym publikacyjnego i projektowo-technologicznego, jest negatywna i dorobek ten nie spełnia wymagań stawianych w postępowaniu o nadanie stopnia doktora habilitowanego w dziedzinie nauk technicznych.

4. Podsumowanie recenzji

Jednotematyczny cykl publikacji dr. inż. Wojciecha Macyny pod wspólnym tytułem „Przechowywanie oraz manipulowanie danymi z uwzględnieniem nowoczesnych rodzajów pamięci”, stanowiący podstawę do ubiegania się o uzyskanie stopnia naukowego doktora habilitowanego, tworzy całość dotyczącą ważnej i aktualnej

problematyki badawczej ukierunkowanej na zwiększenie efektywności procesów przechowywania i przetwarzania danych z uwzględnieniem wyłaniających się technologii pamięci.

Cykl ten nie stanowi jednak znacznego wkładu habilitanta w rozwój dyscypliny naukowej informatyka. Również moja ocena całokształtu jego dorobku naukowego, w tym publikacyjnego i projektowo-technologicznego, jest negatywna. Tym samym nie są spełnione wymagania określone przez aktualnie obowiązującą Ustawę o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki.

A handwritten signature in black ink, consisting of several fluid, overlapping loops and a long horizontal stroke extending to the right.

Wyrazykowski R.