



dr hab. inż. Mirosław Ochodek
Instytut Informatyki,
Wydział Informatyki i Telekomunikacji
Politechnika Poznańska,
tel. 61 665 2944
e-mail: miroslaw.ochodek@put.poznan.pl

Poznań, 17.10.2024

RECENZJA ROZPRAWY DOKTORSKIEJ DLA RADY DYSCYPLINY INFORMATYKA TECHNICZNA I TELEKOMUNIKACJA POLITECHNIKI WARSZAWSKIEJ

Tytuł rozprawy: **Cognitive Biases in Architectural Decision-Making: Impact and Debiasing Strategies**

Autorka rozprawy: **mgr inż. Klara Borowa**

Promotor rozprawy: **dr hab. inż. Andrzej Zalewski**

Dyscyplina: **Informatyka techniczna i telekomunikacja**

Cel i zakres tematyczny rozprawy

Celem badań przedstawionych w rozprawie doktorskiej było zbadanie wpływu błędów poznawczych na proces podejmowania decyzji architektonicznych w trakcie wytwarzania oprogramowania oraz opracowanie podejść/metod ograniczających taki wpływ (tzw. debiasing), co w konsekwencji może przekładać się na wzrost jakości proponowanych rozwiązań architektonicznych. Zdefiniowany cel badawczy łączy ze sobą aspekty istotne z perspektywy inżynierii oprogramowania z ugruntowanymi podstawami teoretycznymi z zakresu psychologii. Zakres rozprawy obejmuje logicznie powiązany ze sobą cykl badań dotyczący identyfikacji przesłanek jakimi kierują się praktycy podejmując decyzje architektoniczne oraz roli błędów poznawczych w tym procesie, poprzez badanie wpływu błędów poznawczych na powstawanie tzw. architektonicznego długu technologicznego, a kończąc na badaniu propozycji strategii/metod mających na celu ograniczenie udziału tychże błędów w podejmowaniu decyzji architektonicznych.

W rozprawie nie sformułowano tezy badawczej, określając w zastępstwie cztery główne pytania badawcze, które można sparafrazować w następujący sposób:

- RQ1: Jakie są główne uzasadnienia dla podejmowania decyzji architektonicznych przez praktyków?
- RQ2: W jaki sposób błędy poznawcze wpływają na podejmowanie decyzji architektonicznych?
- RQ3: Czy błędy poznawcze wpływające na podejmowanie decyzji architektonicznych powodują powstawanie architektonicznego długu technologicznego?
- RQ4: W jaki sposób można zredukować negatywny wpływ błędów poznawczych na podejmowanie decyzji architektonicznych?



Biorąc pod uwagę wcześniejsze badania dotyczące podejmowania i dokumentowania decyzji architektonicznych oraz badania dotyczące roli błędów poznawczych w innych obszarach inżynierii oprogramowania należy podkreślić, że zaproponowany cel badawczy jest interesujący i nowatorski, a wkład wynikający z przeprowadzonych badań empirycznych jest cenny dla rozwoju dyscypliny.

Zawartość i układ pracy

Recenzowana rozprawa doktorska napisana jest w języku angielskim, obejmuje 167 stron i podzielona jest na 10 rozdziałów oraz bibliografię. Pięć rozdziałów rozprawy zostało przygotowanych na podstawie wcześniej opublikowanych prac naukowych, których współautorką (a w przypadku czterech z nich także pierwszą autorką) jest doktorantka tj.:

- Borowa, K., Lewanczyk, R., Stpicyńska, K., Stradomski, P., & Zalewski, A. (2023). What rationales drive architectural decisions? An empirical inquiry. In *European Conference on Software Architecture* (pp. 303-318). Cham: Springer Nature Switzerland.
- Zalewski, A., Borowa, K., & Ratkowski, A. (2017). On cognitive biases in architecture decision making. In *Software Architecture: 11th European Conference, ECSA 2017, Canterbury, UK, September 11-15, 2017, Proceedings 11* (pp. 123-137). Springer International Publishing.
- Borowa, K., Zalewski, A., & Kijas, S. (2021). The influence of cognitive biases on architectural technical debt. In *2021 IEEE 18th International Conference on Software Architecture (ICSA)* (pp. 115-125). IEEE.
- Borowa, K., Dwornik, R., & Zalewski, A. (2021). Is knowledge the key? an experiment on debiasing architectural decision-making-a Pilot study. In *Product-Focused Software Process Improvement: 22nd International Conference, PROFES 2021, Turin, Italy, November 26, 2021, Proceedings 22* (pp. 207-214). Springer International Publishing.
- Borowa, K., Jarek, M., Mystkowska, G., Paszko, W., & Zalewski, A. (2022). Debiasing architectural decision-making: a workshop-based training approach. In *European Conference on Software Architecture* (pp. 159-166). Cham: Springer International Publishing.

Warto zaznaczyć, że pomimo oparcia istotnej części rozprawy o wcześniej opublikowane artykuły naukowe zachowuje ona jednorodny, spójny i kompletny charakter pozycji książkowej.

Pierwsze dwa rozdziały rozprawy stanowią wprowadzenie. Rozdział 1 zarysowuje cel i główne pytania badawcze rozprawy omawiając także jej strukturę, a rozdział 2 zawiera wprowadzenie teoretyczne przybliżając czytelnikowi zagadnienia związane z istnieniem i rolą błędów poznawczych (ze szczególnym uwzględnieniem badań prowadzonych w obszarze inżynierii oprogramowania) oraz typowymi strategiami redukcji ich wpływu. W kolejnych sześciu rozdziałach przedstawiono wyniki przeprowadzonych badań empirycznych układając je w logicznie powiązaną sekwencję.

W rozdziale 3 opisano badania, których celem było określenie jakimi przesłankami kierują się praktycy podejmując decyzje architektoniczne w projektach informatycznych. Zaprojektowano i przeprowadzono dwufazowe badanie empiryczne obejmujące badanie ankietowe przy udziale 63 uczestników oraz pogłębione wywiady z 13 uczestnikami wcześniejszego badania ankietowego. Przedstawione w rozdziale wyniki



badan wskazują, że praktycy najczęściej wybierają rozwiązania, które są im znane lub które wydają się być stosunkowo łatwe w implementacji. Wyniki badań wskazują również na istotne różnice w sposobie podejmowania decyzji w zależności od doświadczenia zawodowego decydenta.

W rozdziale 4 przedstawiono wyniki badań empirycznych, których celem było zidentyfikowanie najważniejszych błędów poznawczych obecnych w procesie podejmowania decyzji architektonicznych oraz rozpoznanie, jak błędy te oddziałują na różne aspekty podejmowanych decyzji. Badanie oparto o warsztat z udziałem 14 praktyków, podczas którego zidentyfikowano 12 kluczowych błędów poznawczych mających wpływ na podejmowanie decyzji architektonicznych. Wyniki badania wskazują, że błędy poznawcze są powszechne i mają istotny wpływ na decyzje architektoniczne, zakres rozważanych alternatyw oraz postrzeganie wymagań. Dodatkowo rozważano także możliwe podejścia do ograniczenia wpływu błędów poznawczych na proces podejmowania decyzji architektonicznych.

W rozdziale 5 przedstawiono wyniki badania dotyczącego wpływu błędów poznawczych na powstawanie architektonicznego długu technicznego (ATD). Bazując na wywiadach z 12 praktykami zaobserwowano widoczne powiązania pomiędzy wystąpieniem poszczególnych typów błędów poznawczych na etapie podejmowania decyzji architektonicznych a pojawieniem się długu technologicznego określonego rodzaju. Jako najbardziej istotne z tej perspektywy rodzaje błędów poznawczych uznano zakotwiczenie, nadmierny optymizm oraz efekt potwierdzenia. W ramach prowadzonych wywiadów zebrano także opinie uczestników dotyczące możliwych podejść do przeciwdziałania błędom poznawczym.

Kolejne trzy rozdziały - 6, 7 oraz 8 - skupiają się na rozwoju i badaniu podejść do przeciwdziałania występowania błędów poznawczych w trakcie podejmowania decyzji architektonicznych. W rozdziale 6 doktorantka przedstawiła wyniki pilotażowego badania mającego na celu ocenę skuteczności prostej interwencji „debiasingowej” polegającej na przybliżeniu decydentom w ramach prezentacji wiedzy na temat trzech wybranych typów błędów poznawczych oraz ich wpływu na podejmowanie decyzji architektonicznych. Badanie zostało przeprowadzone na dwóch grupach studentów, przy czym tylko jedna z nich wysłuchała prezentacji dotyczącej błędów poznawczych. Badanie obejmowało przeprowadzenie czterech spotkań projektowych w trybie zdalnym, podczas których uczestnicy dyskutowali bieżące decyzje projektowe, a ich wypowiedzi były nagrywane, transkrybowane i kodowane pod kątem występowania określonych błędów poznawczych. Wyniki badania wykazały, że prezentacja dotycząca błędów poznawczych nie przyniosła oczekiwanego efektu – grupa niepoddana interwencji używała więcej racjonalnych argumentów niż grupa, która wysłuchała prezentacji. Obie grupy wykazały dominujący wpływ błędu zakotwiczenia, a także efektu potwierdzenia oraz błędu nadmiernego optymizmu. Na podstawie wstępnych wyników uznano badane podejście do ograniczania wystąpienia błędów poznawczych jako niewystarczające sugerując konieczność zastosowania bardziej zaawansowanych technik debiasingowych.

W rozdziale 7 doktorantka przedstawiła badanie eksperymentalne mające na celu ocenę skuteczności zastosowania specjalizowanych warsztatów ograniczających wpływ błędów poznawczych na proces podejmowania decyzji architektonicznych. Badanie zostało przeprowadzone na grupie 44 studentów studiów magisterskich z podziałem na 12 zespołów. Każdy zespół opracował dwa projekty architektoniczne – jeden przed udziałem w warsztatach, a drugi po ich ukończeniu. Warsztaty obejmowały prezentację dotyczącą błędów



poznawczych oraz ćwiczenia mające na celu zastosowanie technik debiasingowych, takich jak anty-zakotwiczenie, przeciwdziałanie efektowi potwierdzenia oraz redukcji błędu polegającego na nadmiernym optymizmie. Analiza nagranych dyskusji wskazała na widoczną poprawę w sposobie argumentacji oraz jakości podejmowanych decyzji po warsztatach. Uczestnicy zaczęli częściej prezentować bardziej racjonalne argumenty i mniej poddawać się wpływowi błędów poznawczych. Zaobserwowano natomiast, że zmianie nie uległa sama liczba stwierdzeń obarczonych błędem poznawczym.

W rozdziale 8 przedstawiono wyniki badania mającego na celu sprawdzenie czy obserwacje poczynione w ramach eksperymentu z udziałem studentów można przenieść na grunt przemysłowej praktyki wytwarzania oprogramowania. Ponownie badanie zaprojektowano jako kontrolowany eksperyment podzielony na dwie fazy – badanie pilotażowe oraz badania właściwe. Łącznie w badaniu udział wzięło 18 praktyków pochodzących z trzech krajów. Dodatkowo na bazie wcześniejszych doświadczeń wprowadzono usprawnienia do zakresu warsztatu debiasingującego. Zaobserwowano, że po wzięciu udziału w warsztacie zmniejszeniu uległa liczba zaobserwowanych wystąpień błędów poznawczych. W przeciwieństwie do wcześniejszego badania nie zaobserwowano wzrostu liczby argumentów nie obciążonych błędami poznawczymi. Na podstawie zebranych doświadczeń opracowano także zbiór porad praktycznych związanych z prowadzeniem warsztatów debiasingowych w praktyce.

Rozdziały 9 i 10 zamykają rozprawę podsumowując najważniejsze wyniki oraz ograniczenia badań w kontekście postawionych wcześniej centralnych pytań badawczych.

Ocena zastosowanego piśmiennictwa w ramach rozprawy doktorskiej

Dobór literatury w pracy naukowej jest kompleksowy (praca zawiera 182 odniesienia literaturowe) i obejmuje zarówno prace teoretyczne dotyczące błędów poznawczych, jak i prace badające ich wpływ na procesy inżynierii oprogramowania. Dodatkowo, rozprawa uwzględnia badania dotyczące długu technicznego oraz metod debiasingu, co pozwala na wszechstronne zrozumienie problematyki i identyfikację potencjalnych rozwiązań. Literatura została starannie wyselekcjonowana, aby wspierać badania empiryczne i dostarczać solidnych podstaw teoretycznych dla przeprowadzanych eksperymentów i analiz.

Bibliografia jest aktualna i obejmuje zarówno klasyczne, jak i nowsze źródła, co świadczy o solidnym przygotowaniu teoretycznym doktorantki. Autorka wykorzystuje odpowiednie źródła do wsparcia swoich argumentów, co zwiększa wiarygodność i głębokość analiz.

Ocena doboru i poprawności zastosowania metod badawczych

Badania zaprezentowane w rozprawie zostały oparte o empiryczne metody badawcze. Doktorantka zastosowała pełen wachlarz metod tego typu poczynając od badania ankietowego, poprzez wywiady i warsztaty a kończąc na kontrolowanych eksperymentach. Trzeba przyznać, że dobór metod i różnorodność podejść robi bardzo dobre wrażenie i jest godne uznania.

Jeśli chodzi o metody analizy danych. Doktorantka korzystała głównie z metod jakościowych opartych o generowanie hipotez i kodowanie otwartych tekstów (najczęściej pozyskanych w drodze transkrypcji zapisu



audio), aczkolwiek w przypadku badań opisanych w ostatnich rozdziałach rozprawy użyto także metody ilościowe wsparte wnioskowaniem statystycznym.

Uzyskane wyniki i ich znaczenie dla rozwoju dyscypliny naukowej oraz praktyki

W pracy przedstawiono szereg bardzo wartościowych wyników badawczych. Do najbardziej interesujących w moim odczuciu należy zbadanie wpływu błędów poznawczych na decyzje architektoniczne oraz ich obecność w ramach decyzji projektowych prowadzących do powstawania architektonicznego długu technologicznego. Pomimo tego, że badania te w znaczącej mierze mają charakter eksploracyjny, to ich wyniki już na tym etapie mają istotną wartość praktyczną. Jak pokazały badania związane z przeciwdziałaniem wpływu błędów poznawczych na podejmowanie decyzji architektonicznych już samo zwiększenie świadomości praktyków co do ich istnienia i roli w procesie decyzyjnym może nieść za sobą wymierne korzyści. Proponowane warsztatowe techniki debiasingu mogą zostać zastosowane w szkoleniu architektów oprogramowania a także studentów kierunków informatycznych dając nadzieje na polepszenie jakości sposobu argumentowania oraz dyskusowania na temat decyzji projektowych.

Uwagi krytyczne i pytania

W moim przekonaniu doktorantka wykazała się dużą dbałością o szczegóły w trakcie realizacji prowadzonych badań, niemniej jednak lektura rozprawy skłania mnie do zadania kilku pytań oraz sformułowania kilku ogólnych uwag co do jej treści.

- 1) Reprodukowalne badania naukowe (ang. Reproducible Research). Istotną cechą prawidłowo wykonanych badań empirycznych jest ich przejrzystość (transparentność) oraz reprodukowalność. W związku z tym intensywnie promowaną praktyką badawczą w obszarze empirycznej inżynierii oprogramowania jest uzupełnianie raportów z prowadzonych badań o tzw. pakiety reprodukcyjne, pozwalające zweryfikować narracje autorów prac naukowych korzystając bezpośrednio z materiałów źródłowych. W przypadku badań opisanych w ostatnich rozdziałach rozprawy opublikowano tego typu pakiety, co należy docenić. Niestety podobne pakiety nie są dostępne dla wcześniejszych badań. W moim odczuciu uzupełnienie tych materiałów byłoby niezwykle wartościowe.
- 2) W ramach badania ankietowego opisanego w rozdziale 3 dotyczącego przesłanek do podejmowania decyzji architektonicznych pozyskano szeroki zakres danych demograficznych o uczestnikach (m.in. wiek, płeć, wielkość firmy oraz branża, w której funkcjonuje), natomiast w ramach przedstawionych analiz skupiono się wyłącznie na powiązaniu doświadczenia decydentów (mierzonego w latach pracy) z rodzajem przesłanek prowadzących do podejmowania przez nich decyzji architektonicznych. Dlaczego doktorantka wybrała tylko ten aspekt jako potencjalnie powiązany z wyborem przesłanek?
- 3) W tym samym badaniu ankietowym pytano o najczęściej występujące przesłanki do podejmowania decyzji architektonicznych nie pytając o same decyzje, co w moim odczuciu może rodzić pewne istotne implikacje w kontekście badań częstotliwościowych. Np. prowadzi do niejawnego założenia, że wszyscy uczestnicy podejmowali zbliżone rodzaje decyzji projektowych. Odwołując się do definicji decyzji architektonicznej zaproponowanej przez Jana Boscha należy zauważyć, że nie każda decyzja projektowa jest decyzją



architektoniczną. Według Boscha decyzje architektoniczne to takie, które mają istotny wpływ na strukturę komponentów i konektorów czy też niosą ze sobą reguły lub ograniczenia dla tego w jaki sposób rozwijany jest system w ramach danej architektury.¹ Zatem rodzi się pytanie - czy i na jakiej podstawie możemy założyć, że uczestnicy badania podejmowali tego typu decyzje i były to decyzje o zbliżonej randze – np. czy możemy stwierdzić, że programista posiadający nikłe doświadczenie zawodowe podejmował faktycznie decyzje, które mają znamiona decyzji architektonicznych. Jakie znaczenie może mieć ten problem dla poprawności konkluzji z tego badania?

Podobnie ocena częstotliwości wystąpień poszczególnych przesłanek do podejmowania decyzji architektonicznych wydaje się bazować na niejawnym założeniu o jednorodności kontekstów projektów, w których brali udział uczestnicy badania. Czy możemy założyć, że takie same przesłanki będą miały miejsce przy podejmowaniu decyzji architektonicznych dla prostej aplikacji internetowej jak i dla złożonego systemu bankowego? Czy nie znając kontekstu podejmowania decyzji możemy w bezpieczny sposób wysnuć wnioski na temat częstotliwości/popularności przesłanek do podejmowania decyzji architektonicznych w ramach tego badania?

4) W badaniu dotyczącym występowania błędów poznawczych w procesie decyzyjnym (rozdział 4 rozprawy) uczestnicy zapoznawali się z listą 105 błędów poznawczych. Biorąc pod uwagę ich znaczną liczbę rodzi się pytanie, czy kolejność prezentowania błędów na liście mogła przełożyć się na wyniki badania częstotliwości wystąpienia tych błędów poznawczych? Czy badano ten aspekt lub przeciwdziałano jego wpływowi np. podając uczestnikom losowo posortowane listy?

5) Badając związek błędów poznawczych z występowaniem architektonicznego długu technologicznego jako najbardziej istotne błędy poznawcze uznano te o największej częstotliwości występowania. Czy spojrzenie na problem istotności wyłącznie z perspektywy prawdopodobieństwa (częstotliwości) jest wystarczające? Czy nie powinno zostać ono wzbogacone o aspekty związane z oceną konsekwencji, podobnie jak ma to miejsce w przypadku analizy ryzyka, gdzie ocenami istotność ryzyka w oparciu o wiele czynników np. prawdopodobieństwo i wpływ jego wystąpienia?

6) Kontrolowany eksperyment z udziałem studentów opisany w rozdziale 7 zakładał sekwencję wykonania dwóch zadań rozdzielonych warsztatem debiasingowym. Czy i w jaki sposób kontrolowano wpływ czynnika jakim jest naturalne uczenie się/zdobywanie praktycznego doświadczenia w trakcie eksperymentu (niezależnie od warsztatu)? Problem ten nie jest zarysowany w podrozdziale na temat zagrożeń do poprawności. Dodatkowo wyniki eksperymentu z udziałem praktyków sugerują, że osoby doświadczone mają mniej problemów z formułowaniem argumentów, które nie są obciążone błędami poznawczymi co można by uznać za potencjalny argument świadczący o istotności zdobywania doświadczenia w tym kontekście.

7) W ramach eksperymentu z udziałem praktyków prezentowanym w rozdziale 8 rozprawy przeprowadzono 15 testów statystycznych dla poszczególnych miar (Tabela 8.3) każdorazowo zakładając stały poziom

¹ Bosch, J. (2004). Software architecture: The next step. In *Software Architecture: First European Workshop, EWSA 2004, St Andrews, UK, May 21-22, 2004. Proceedings 1* (pp. 194-199). Springer Berlin Heidelberg.



istotności. Czy poczynione obserwacje nie podlegają zagrożeniu do poprawności wynikającego z wielokrotnego testowania (ang. multiple testing)?

8) W pracy można znaleźć także drobne defekty techniczne, które nie mają jednak istotnego wpływu na jakość merytoryczną samej rozprawy, np. problemy w numeracji rozdziałów czy tabel oraz pewne niespójności np. w rozdziale 10 podsumowującym wyniki jest mowa o 11 typach błędów poznawczych, natomiast w rozdziale 4 w Tabeli 4.2 prezentowanych jest 12 typów.

Podsumowanie

Recenzowana rozprawa stanowi istotny i oryginalny wkład w rozwój dziedziny badań w zakresie inżynierii oprogramowania, a w szczególności zagadnień związanych z architekturą oprogramowania. Autorka rozprawy w przekonujący sposób wykazała zdolność do prowadzenia badań naukowych, od etapu planowania i projektowania eksperymentów, poprzez ich realizację, analizę danych, aż po interpretację wyników i formułowanie praktycznych rekomendacji. Warto podkreślić, że prowadzenie tego typu badań jest zadaniem niezwykle trudnym, niewybaczającym błędów i wymagającym doskonałej organizacji i systematyczności, zwłaszcza biorąc pod uwagę, że w badaniach przedstawionych w rozprawie bardzo często uczestniczyły osoby zawodowo zajmujące się wytwarzaniem oprogramowania. Warto podkreślić także, że wyniki badań przedstawionych w rozprawie zyskały uznanie komitetów programowych sztanदारowych konferencji międzynarodowych w obszarze badań nad architekturą oprogramowania takich jak ICSA i ECSA.

Biorąc pod uwagę powyższe, stwierdzam, iż przedstawiona do recenzji rozprawa doktorska autorstwa mgr inż. Klary Borowej pt. „Cognitive Biases in Architectural Decision-Making: Impact and Debiasing Strategies” spełnia wymogi formalne stawiane rozprawom doktorskimi i wnoszę o jej dopuszczenie do publicznej obrony jednocześnie rekomendując ją do wyróżnienia.

dr hab. inż. Mirosław Ochodek

Politechnika Poznańska