

Łódź, dnia 04.08.2023

dr hab. inż. Aneta Poniszewska-Marańda
Instytut Informatyki
Politechnika Łódzka

**RECENZJA ROZPRAWY DOKTORSKIEJ DLA RADY NAUKOWEJ DYSCYPLINY
„INFORMATYKA TECHNICZNA I TELEKOMUNIKACJA” POLITECHNIKI
WARSZAWSKIEJ**

Tytuł rozprawy: Wielowymiarowa eksploracja repozytoriów programowych w zakresie raportów zgłoszeń oraz ich obsługi

Autor rozprawy: mgr inż. Bartosz Wojciech Dobrzyński

Promotor rozprawy: prof. dr hab. inż. Janusz Sosnowski

Recenzja sporządzona na podstawie pisma Przewodniczącego Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja, dr hab. inż. Jarosława Arabasa, prof. uczelni, z dnia 12 czerwca 2023 r.

Ocena układu rozprawy doktorskiej, zawartość

Rozprawa zawiera sześć rozdziałów, bibliografię oraz załączniki. Posiada 141 stron, w tym 10 stron załączników. W rozprawie umieszczono 108 pozycji bibliograficznych, dobrze dobranych i aktualnych.

Rozdział pierwszy jest wprowadzeniem do pracy. Przedstawia motywację, jaka przyświecała Autorowi podczas realizacji pracy doktorskiej i pisania rozprawy oraz kontekst badań.

Praca dotyczy zagadnień związanych z procesem rozwoju i utrzymania oprogramowania, w aspekcie monitorowania tego procesu w obszarze prowadzenia projektów informatycznych. Monitorowanie tego procesu skupia się na wykorzystaniu współczesnych repozytoriów kodu oraz obsłudze zgłoszeń, celem zapewnienia wsparcia procesu wytwarzania oprogramowania. Kontekst badań jest osadzony we współczesnej literaturze, dotyczącej poruszanej tematyki. Autor w ogólny sposób odnosi się do wybranych pozycji literatury, jednak w wystarczający sposób jak na Wprowadzenie do pracy, prezentując ich główne założenia, zalety i wady.

Rozdział drugi prezentuje tezę pracy oraz sformułowanie problemu badawczego. Teza pracy została określona następująco:

„Analiza i ocena procesu wytwarzania oprogramowania wymaga opracowania reprezentatywnych modeli oraz metod eksploracji danych z repozytoriów zgłoszeń i repozytoriów kodu. Uwzględnienie różnych poziomów ekstrakcji i agregacji informacji oraz perspektyw obserwacji, poszerza zakres przedmiotowy monitorowania projektu i ułatwia identyfikację niedoskonałości”.

Z tezą powiązано zadania badawcze, dotyczące trzech zagadnień, które stanowią zakres niniejszej rozprawy.

Rozdział trzeci pracy prezentuje metodykę analizy repozytoriów kodu. Opis metodyki poprzedzony jest analizą stanu wiedzy w badanej tematyce repozytoriów kodu. Następnie

prezentowana jest charakterystyka repozytoriów programowych, systemów zgłoszeń ITS (Issue Tracking Systems), w szczególności z punktu widzenia zastosowania narzędzi JIRA. W kolejnej części rozdziału zaprezentowano wielowymiarową eksplorację cech repozytoriów, której dokonano według następujących kryteriów: ogólne charakterystyki, aktywności aktorów, profile komentarzy dodawanych do zgłoszeń oraz korelacja atrybutów zgłoszeń. Eksploracja ta dotyczy trzech obszarów: (1) ogólne raporty statystyczne wraz z czasami aktywności, (2) strukturalne i semantyczne analizy podstawowych cech raportowanych pól, (3) określenie różnic repozytoriów oraz ich niedoskonałości. Ogólne charakterystyki pól repozytorium przedstawiono za pomocą następujących cech: współczynnik wypełnienia poszczególnych pól, rozkład przyjmowanych wartości, najrzadziej i najczęściej występująca wartość w polu, liczba możliwych wartości, cechy czasu dla zgłoszeń. W analizie korelacji atrybutów zgłoszeń modele rozkładu cech repozytoriów ITS mogą być wyznaczone niezależnie lub zależnie od innych pól lub innych połączonych repozytoriów, np. zagregowane dane mogą być uzależnione od filtrów, takich jak typ zgłoszenia, priorytet, istotność, rola reportera.

Rozdział czwarty pracy przedstawia analizę tekstową raportów zgłoszeń opartą na wielowymiarowej eksploracji tekstów dostępnych w repozytoriach zgłoszeń, m.in. w tzw. polach opisowych, takich jak tytuł, opis, komentarze zgłoszenia. Podobnie, jak w rozdziale trzecim, analiza tekstowa poprzedzona jest analizą stanu wiedzy w badanej tematyce, dotyczącej eksploracji tekstów w systemach śledzenia zgłoszeń (ITS) oraz systemach kontroli wersji (SVC). Następnie dokonano analizy słownikowej opisów błędów zgłoszeń w repozytoriach. Autor przeanalizował dostępne teksty i doszedł do wniosku, że słowa występujące w zgłoszeniach są zbiorem słów języka naturalnego (należące do tezaursu) oraz słów niebędących częścią tezaursu, które mogą reprezentować między innymi żargon IT/projektowy, nazwy funkcji aplikacji/systemu, odniesienia do kodu aplikacji/systemu, referencje/linki do załączników, referencje/linki do zewnętrznych zasobów, fragmenty kodu, logi aplikacji. Ponadto, tekst może zawierać błędy w pisowni, np. tzw. literówki lub zapożyczenia z języków obcych, co jest bardzo częste w projektach IT. Autor wyróżnił cztery słowniki: słownik słów należących do tezaursu (NLW), słownik słów projektowych (FW), słownik słów niesklasyfikowanych (NCW), słownik słów zgodnych z notacją *CamelCase* (CCW). Słownik NLW w większości przypadków zawiera elementy występujące w słowniku tezaursu. Autor opracował zbiór wyrażeń regularnych identyfikujących badane teksty. Wyrażenia te zostały użyte w implementacji algorytmu przetwarzania tekstu podlegającego metodom klasyfikacji, który został zaprezentowany w kolejnym podrozdziale. Następnie przedstawiono dwa algorytmy w postaci pseudokodu, obrazujące funkcję generowania słowników NLW, FW, CCW, NCW oraz funkcję pomocniczą, nazwaną „is_valid_tag”. W kolejnej części rozdziału zaprezentowano autorski algorytm klasyfikacji zadań dodawanych do repozytorium zgłoszeń. Jest on wykonywany w dwóch fazach: (1) wstępne przetwarzanie tekstu, (2) klasyfikacja. Wstępne przetwarzanie tekstu obejmuje następujące kroki: (1) pobranie zgłoszeń z repozytorium poprzez API JIRA, (2) stworzenie zbioru oryginalnych encji tekstów oznaczonych identyfikatorem zgłoszenia lub komentarza, (3) redukcja tekstu poprzez usunięcie słów przystankowych oraz cyfr niepołączonych z innymi znakami, (4) transformacja oryginalnych encji tekstów z wykorzystaniem wyrażeń regularnych oraz wyznaczenie opracowanych cech tekstowych. Te kroki algorytmu zaprezentowano w postaci pseudokodu. Zaproponowany algorytm klasyfikacji tekstów oparty jest na wyborze najlepszego klasyfikatora, który na wejściu przyjmuje listę algorytmów do walidacji, zbiór trenujący i listę cech, które mają zostać zweryfikowane. Wynikiem działania algorytmu jest klasyfikator (algorytm i zestaw cech) z najlepszym wynikiem precyzji oraz macierzą z wynikami wszystkich przeprowadzonych testów.

W kolejnej sekcji zaprezentowano wyniki eksperymentów przeprowadzone dla dwóch zadań klasyfikacji: typu zgłoszenia i kategorii komentarza oraz ich analizę. Ostatni podrozdział rozdziału czwartego zawiera dyskusję zagadnień poruszonych w tym rozdziale, w tym między innymi zagadnień klasyfikacji tekstów zgłoszeń w repozytoriach oraz przeprowadzonych eksperymentów dla zaproponowanych algorytmów.

Rozdział piąty pracy prezentuje wielowymiarową analizę procesu obsługi zgłoszeń w repozytoriach programowych. Analiza taka może być przeprowadzona na dwóch poziomach: (1) ogólnym, tzw. gruboziarnistym i (2) szczegółowym, tzw. drobnoziarnistym. W związku z tym w pierwszej części rozdziału zaprezentowano gruboziarnistą oceną efektywności procesu obsługi zgłoszeń według dwóch metryk: (1) czasu obsługi zgłoszeń, poprzez wyznaczone profile czasowe obsługi zgłoszeń oraz (2) zakres nierozwiązanych zgłoszeń, poprzez liczby nierozwiązanych zgłoszeń. Trzecia metryka to rozkład zmian kodu względem zgłoszeń. Ponadto, Autor zaprezentował pojęcie długu błędów i algorytm detekcji długu błędów wraz ze wzrostami w różnych okresach czasu. Algorytm ten bazuje na czterech danych wejściowych: data początku weryfikowanego okresu, data końca weryfikowanego okresu, interwał czasowy (miesiąc bądź sprint) i próg, powyżej którego liczba nierozwiązanych błędów będzie zliczana. Dodatkowo, opisano sposób minimalizacji ryzyka związanego z długiem błędów.

W drugiej części rozdziału zaprezentowano autorski model obsługi zgłoszeń w repozytoriach programowych. Jest to grafowy model, określony jako IHG, *Issue Handling Graph*. Model został zdefiniowany, opisany i zaprezentowany w formie graficznej oraz poparty przykładami na podstawie projektów informatycznych, które są rozpatrywane przez całą rozprawę, a które stanowią punkt wyjścia do badań i analiz rozwiązań zaproponowanych przez Autora. W sekcji 5.2.1 przedstawiono szczegóły modelu IHG w formie opisu przepływu zgłoszeń w modelu. W sekcji 5.2.2. przedstawiono algorytmy analizy grafu IHG, wśród których wyróżniono trzy ich grupy: algorytmy przetwarzania danych, algorytmy agregacji danych, algorytmy pomocnicze. W ramach algorytmów przetwarzania danych zaprezentowano: algorytm pobierania danych, dotyczących poszczególnych zgłoszeń; algorytm budowania ścieżek z listy zgłoszeń JIRA. W ramach algorytmów agregacji danych zaprezentowano: algorytm budowy grafu IHG, który agreguje dane ścieżek wygenerowanych przez poprzedni algorytm. W ramach algorytmów pomocniczych zaprezentowano: algorytm wizualizacji grafu IHG, wygenerowanego przez algorytm budowy grafu, na który składa się algorytm filtracji wierzchołków, algorytm filtracji krawędzi, algorytm generowania pliku stanów, algorytm generowania pliku ze statystykami unikalnych ścieżek grafu, algorytm wyszukiwania i wyznaczanie anomalii (pętli) ścieżek.

Zaprezentowane algorytmy wspierają generowanie szeregu charakterystyk ścieżek oraz stanów grafu, które mogą być grupowane w statystyki ścieżek (czasowe, strukturalne, ilościowe, komentarzy zgłoszeń przechodzących przez ścieżkę) i statystyki stanów (czasowe, ilościowe, strukturalne). Opracowane algorytmy mogą służyć do generowania grafu pełnego lub z ograniczeniami, np. dla określonego typu zgłoszenia, stanu, grupy użytkowników tworzących zgłoszenia lub czasu, w którym zostały one dodane do repozytorium.

W kolejnych sekcjach opisano charakterystyki profili stanów oraz ścieżek obsługi zgłoszeń opracowanych na podstawie modelu IHG – charakterystyki czasowe stanów grafu IHG w sekcji 5.2.3 oraz profile ścieżek obsługi zgłoszeń w sekcji 5.2.4, gdzie wyróżniono dwie klasy profili ścieżek: wydajnościowo-czasowe i strukturalne. Opisywane zagadnienia poparto przykładami konkretnych projektów programistycznych. Podczas generowania profili ścieżek zgłoszeń zaobserwowano pojawienie się anomalii w procesie obsługi zgłoszeń. W związku z tym zaproponowano autorską metodę wykrywania tych anomalii, opartą na użyciu specjalnie zdefiniowanych wyrażeń regularnych. Kolejne sekcje zawierają analizę czasową obsługi zgłoszeń w ścieżkach oraz dyskusję zaproponowanych rozwiązań i otrzymanych wyników eksperymentów przeprowadzonych dla tych rozwiązań.

Szósty rozdział rozprawy zawiera podsumowanie pracy, opis osiągnięć pracy oraz perspektywy rozwoju tematu pracy w niedalekiej przyszłości.

Siódmy rozdział pracy prezentuje zastosowaną bibliografię, która zawiera 108 pozycji bibliograficznych.

Ostatnia część rozprawy zawiera załączniki do głównej części pracy, które prezentują odpowiednio: profile ścieżek obsługi zgłoszeń dla projektu Groovy, grafy modelu IHG dla projektu MongoDB, zestaw atrybutów ścieżek obsługi zgłoszeń oraz szczegółowe charakterystyki ścieżek obsługi zgłoszeń.

Ocena zastosowanego piśmiennictwa

Rozprawa zawiera 108 pozycji bibliograficznych, w tym 11 pozycji internetowych, które z reguły stanowią dokumentację techniczną wybranych narzędzi i technologii. Zdecydowana większość pozycji to artykuły naukowe, opublikowane w czasopismach naukowych lub w materiałach konferencyjnych konferencji naukowych w ostatnich kilku-kilkunastu latach. Wśród pozycji są również prace opublikowane przez Autora – w sumie 4 prace, których jest współautorem.

Autor przedstawił analizę źródeł literaturowych częściowo w rozdziale 2 – na temat repozytoriów programowych oraz częściowo w rozdziale 3 rozprawy – na temat eksploracji tekstów w systemach śledzenia zgłoszeń (ITS) i w systemach kontroli wersji (SVC).

Przeprowadzona analiza obecnego stanu wiedzy w poruszanej tematyce oraz istniejących rozwiązań wskazuje, że Autor poprawnie rozumie problemy, związane z tematem pracy. Posiada wiedzę na temat rozwiązań krajowych i światowych w poruszonym zakresie. Co prawda część istotnych zagadnień oraz problemów została jedynie zasygnalizowana, jednakże taka forma pozwala mieć nadzieję, że Autor posiada również wiedzę szczegółową na ich temat.

Wskazanie i ocena celu pracy

Autor sformułował tezę pracy i problem badawczy w rozdziale 2. rozprawy. W rozdziale tym zarysowano problemy, związane z procesem wytwarzania oprogramowania w aspekcie zgłoszeń błędów i nowych funkcji do realizowanych projektów.

Z tezą pracy powiązано zadania badawcze, dotyczące trzech obszarów: (1) eksploracji zawartości informacyjnej repozytoriów zgłoszeń (poprzez opracowanie charakterystyk strukturalnych, czasowych, zdefiniowanie profili różnych aspektów informacyjnych oraz opracowanie algorytmów analiz statystycznych i tekstowych), (2) opracowania rozszerzonego modelu grafowej obsługi zgłoszeń (IHG, *Issue Handling Graph*) oraz metod analizy zorientowanej na badanie profili obsługi różnych typów zgłoszeń i detekcji anomalii w repozytoriach, (3) weryfikacji opracowanej metodologii na danych z repozytoriów projektów typu open source oraz projektów komercyjnych.

Niemniej jednak cel pracy nie został jasno i precyzyjnie zidentyfikowany i zaprezentowany. Jedynie można domniemywać, że celem pracy jest zrealizowanie zdefiniowanych zadań badawczych.

Na stronach 18-19, w rozdziale 3, Autor definiuje cele badawcze w odniesieniu do repozytoriów projektowych oraz ich analizy. Cele te częściowo pokrywają się zadaniami badawczymi określonymi w rozdziale 2.

Wskazanie i ocena zastosowanych metod badawczych

Autor w swojej pracy stosuje następujące metody badawcze: analiza dostępnej literatury w temacie rozprawy, analiza przedmiotu badań, czyli repozytoriów programowych, obserwacja prowadzonych projektów informatycznych, metody statystyczne i ilościowe w analizach danych uzyskanych z repozytoriów programowych, prace eksperymentalne.

Uważam, że wszystkie zastosowane metody badawcze są poprawnie wybrane i zastosowane. Stanowią kompletny warsztat badawczy, niezbędny dla przeprowadzenia badań w ramach tematu rozprawy.

Autor opiera się również na własnym praktycznym doświadczeniu, wyniesionym z pracy nad projektami komercyjnymi. Takie doświadczenie stanowi niewątpliwie wartość dodaną w pracy nad poruszonym tematem oraz nad przygotowywaniem rozprawy, w szczególności w obszarze inżynierii oprogramowania.

Ocena części rozprawy dotyczącej omówienia wyników badań

Tematyka badań Autora obejmuje wielowymiarową analizą i eksplorację procesu obsługi zgłoszeń w repozytoriach programowych. Badania przeprowadzone w ramach realizacji pracy obejmują dwa powiązane ze sobą obszary: (1) eksploracja zawartości informacyjnej programowych repozytoriów zgłoszeń i (2) wielowymiarowa analiza obsługi zgłoszeń.

Wyniki badań zostały zaprezentowane przez Autora częściowo w rozdziale 3 i 4 oraz w rozdziale 5 rozprawy, odpowiednio na temat analizy repozytoriów programowych, analizy tekstowej raportów zgłoszeń oraz analizy obsługi zgłoszeń. Ponadto, dyskusja na temat uzyskanych wyników badań została zaprezentowana pod koniec rozdziału 4 i 5 oraz w podsumowaniu samej rozprawy.

Sposób prezentacji podejmowanych w pracy zagadnień oraz uzyskanych wyników badań jest na ogół jasny i zrozumiały, chociaż Autor stosuje dużo różnorodnych symboli na oznaczenie poszczególnych projektów oraz ich elementów. Wprowadzone symbole są ogólne, czasami nieintuicyjne, np. nazwy projektów używanych w badaniach. Ponadto, część zagadnień, a także wyników pracy Autora poruszona jest w rozprawie bardzo pobieżnie.

Niemniej jednak tekst rozprawy został przygotowany starannie. Strona redakcyjna nie budzi większych zastrzeżeń – moje uwagi zostały przedstawione w kolejnym punkcie recenzji, dotyczącym nieprawidłowości w rozprawie.

Praktyczne zastosowanie uzyskanych wyników badań

Autor w swoich badaniach łączył aspekty teoretyczne i praktyczne, Ponadto, jak sam wielokrotnie stwierdza w rozprawie, w trakcie prac opiera się na swoim doświadczeniu praktycznym, wyniesionym z wieloletniej pracy w tzw. przemyśle, przy realizacji komercyjnych projektów informatycznych.

W ramach realizacji pracy doktorskiej Autor przeanalizował istotną liczbę repozytoriów projektów, zarówno open source, jak i komercyjnych, celem zbadania zagadnień związanych z eksploracją powiązanych ze sobą procesów wytwarzania i utrzymania oprogramowania oraz

analizą istniejących w nich danych, W związku z tym opracował modele danych i schematy eksploracji, które zostały wsparte odpowiednimi metrykami, mającymi na celu wspieranie: (1) ekstrakcji cech charakterystycznych (składniowych, semantycznych, czasowych i statystycznych) repozytoriów śledzenia problemów, odniesionych do różnych perspektyw obserwacji, (2) śledzenie efektywności obsługi zgłoszonych problemów, (3) wykrywanie niedoskonałości repozytoriów programowych oraz procesu obsługi zgłoszeń.

Zaproponowany model IHG wspierany jest przez opracowane autorskie algorytmy analizy powiązane ze zdefiniowanymi profilami strukturalnymi i statystycznymi i ich miarami, obejmującymi różne aspekty procesów obsługi zgłoszeń. Ponadto, w trakcie analiz ścieżek obsługi zgłoszeń wielu projektów Autor zaobserwował anomalie, dla których opracował metody i schematy analiz.

Stworzona metoda została wykorzystana przez Autora w praktyce do zbadania i oceny procesu obsługi zgłoszeń w projekcie komercyjnym, a następnie od przeprowadzania w nim usprawnień. Według zamieszczonych przez Autora informacji, usprawnienia te dotyczyły optymalizacji złożoności zgłoszeń i optymalizacji procesu obsługi zgłoszeń. Ułatwiło to i przyspieszyło realizację procesu obsługi zgłoszeń w repozytorium projektu.

Wyniki rozprawy mają zatem znaczenie dla dalszego rozwoju nauki w dziedzinie inżynierii oprogramowania oraz posiadają wartość aplikacyjną i praktyczną dla rzeczywiście realizowanych projektów informatycznych.

Ewentualne nieprawidłowości, które pojawiły się w rozprawie

Analiza rozprawy nasunęła mi kilka uwag krytycznych:

- Wykresy na rysunkach 1-4 i 10 nie posiadają oznaczenia nazw osi – powoduje to, że zaprezentowane wykresy są niezrozumiałe.
- Grafy przedstawione na rysunkach 5 i 11-14 są niestety bardzo mało czytelne.
- Niewłaściwe użycie słowa „funkcjonalność” zamiast słowa „funkcja” – na wielu stronach, w całej pracy.
- Niewłaściwe użycie słowa „sekcja” w stosunku do podrozdziału – na wielu stronach, w całej pracy, np. „sekcja 5.1.” na str. 108.
- Zbyt ogólne i przez to niezrozumiałe nazwy niektórych rozdziałów i podrozdziałów oraz sekcji pracy.
- Forma niektórych zdań nasuwa wątpliwości stylistyczne.
- Standardowo tzw. „podpisy” dla tabel umieszczane są nad tabelami, natomiast w pracy znajdują się pod tabelami – na wielu stronach, w całej pracy.
- Co oznacza skrót „Alg”? oczywiście można się łatwo domyślić, ale następujące przykładowe sformułowania nie są eleganckie: „dlatego opracowałem Alg. 1 wspierający ten proces”, „Alg. 2, zaimplementowany został z wykorzystaniem biblioteki”, „Jako parametr wejściowy Alg. 1 przyjmuje”, „przedstawiony w pseudokodzie Alg. 6”, „(wynik działania Alg. 11)”.
- Podobnie, co oznacza skrót „Tab”? – na przykład sformułowania „Statystyka zaprezentowana w Tab. 12 daje istotny wgląd”, „budowanego za pomocą wyrażeń regularnych Tab. 11”.
- Brak spisu rysunków i spisu tabel na końcu pracy. Brak indeksu stosowanych w pracy symboli.
- Numeracja rysunków, tabel, algorytmów jest niezgodna z ogólnie przyjętymi zasadami dla tzw. długich tekstów, jakim jest rozprawa doktorska.

- Praca powinna być napisana w formie bezosobowej, chociaż w przeważającej części.
- Częste stosowanie sformułowania „oparte o” – poprawnie powinno się stosować sformułowanie „oparte na”.
- Błędy interpunkcyjne.
- Błędy językowe, np. „Model IHG umożliwia wyprowadzenie innych zagregowanych profili np. warianty ścieżek np. ze wspólnym stanem początkowym, wspólnym stanem początkowym i pośrednim, wspólnym stanem początkowym i końcowym itp.”.

Przedstawione uwagi nie obniżają jednakże mojej pozytywnej oceny pracy.

Ocena, czy rozprawa stanowi oryginalne rozwiązanie problemu naukowego

Uważam, że teza rozprawy oraz obszar naukowy rozprawy zostały określone jasno i precyzyjnie. Praca ma charakter teoretyczno-implementacyjny. Część teoretyczna obejmuje:

- przeprowadzenie analizy dostępnych źródeł literaturowych na temat poruszanych w rozprawie zagadnień, dotyczących procesu eksploracji repozytoriów programowych w projektach informatycznych, w szczególności problemu analizy raportów zgłoszeń i ich obsługi,
- autorskie opracowanie profili statystycznych zawartości repozytoriów, profili aktywności aktorów (uczestników) projektu oraz wielowymiarowa eksploracja opisów i komentarzy zgłoszeń,
- zdefiniowanie ogólnych i szczegółowych charakterystyk procesu obsługi zgłoszeń,
- opracowanie autorskiej metody obsługi zgłoszeń, opartej na grafowym modelu IHG,
- opracowanie algorytmów analizy schematów obsługi zgłoszeń w kontekście zaproponowanej ich taksonomii.

Część implementacyjna obejmuje stworzenie narzędzia o nazwie *IssueAnalyzerTool*, które zostało użyte do przeprowadzenia analiz zaproponowanych rozwiązań teoretycznych na wybranym reprezentatywnym zbiorze repozytoriów programowych.

Tematyka rozprawy sytuuje ją w obszarze badawczym, związanym z poszukiwaniem efektywnych metod, dotyczących budowy systemów informatycznych, w szczególności eksploracji repozytoriów programowych celem szybkiego i efektywnego wyszukiwania i obsługi zgłoszeń na temat błędów i nowych funkcji w realizowanych projektach.

Uważam, że podjęcie tematu pracy doktorskiej w rozpatrywanym obszarze i zakresie jest celowe i w pełni uzasadnione potrzebą zapewniania i zwiększania jakości współcześnie tworzonego oprogramowania, a w szczególności dużych, złożonych oraz dynamicznych w swoim działaniu systemów informatycznych. Jest to w pełni uzasadnione ze względów teoretycznych, poznawczych i praktycznych na tle obecnego stanu wiedzy w rozpatrywanym obszarze.

Ocena, czy rozprawa prezentuje ogólną wiedzę teoretyczną kandydata w dyscyplinie oraz umiejętność samodzielnego prowadzenia pracy naukowej

W moim przekonaniu Autor w wystarczającym stopniu przeanalizował aktualny stan wiedzy w rozpatrywanym temacie, jasno sformułował problemy, a następnie w zadawalającym stopniu rozwiązał je w swojej pracy.

Oryginalnym rezultatem rozprawy, a tym samym samodzielnym i oryginalnym dorobkiem Autora według mojej oceny jest opracowanie metody wspomagającej budowę systemów informatycznych, która umożliwia wielowymiarową eksploatację repozytoriów programowych, w szczególności w zakresie raportów zgłoszeń i ich obsługi.

Dorobek Autora powiększa implementacja zaproponowanych rozwiązań poprzez stworzenie prototypu narzędzia o nazwie *IssueAnalyzerTool* do analizy zgłoszeń w repozytoriach w sposób zautomatyzowany. Zaproponowana metoda została ponadto zweryfikowana poprzez zastosowanie jej w rzeczywistości realizowanych projektach informatycznych.

Uważam, że przedstawiona mi do recenzji rozprawa prezentuje ogólną wiedzę teoretyczną kandydata w dyscyplinie informatyka techniczna i telekomunikacja oraz umiejętność samodzielnego prowadzenia pracy naukowej.

Wniosek końcowy

W mojej ocenie, Pan mgr Bartosz Dobrzyński w swojej rozprawie trafnie zdefiniował problem badawczy, a następnie go rozwiązał w odpowiednim zakresie. Pozytywnie oceniam wszystkie wymienione powyżej elementy pracy.

Uważam, że rozprawa spełnia wymagania stawiane rozprawom doktorskim przez obowiązujące przepisy o stopniach i tytułach naukowych. W związku z tym wnoszę o dopuszczenie jej do publicznej obrony.

