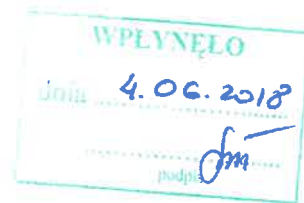


Wrocław, 31 maja 2018 r.

Dr hab. inż. Dariusz Król
Katedra Systemów Informatycznych
Wydział Informatyki i Zarządzania
Politechnika Wrocławska
E-mail: Dariusz.Krol@pwr.edu.pl



Recenzja rozprawy doktorskiej

mgra inż. Krzysztofa Kutta

pt. „Metody projektowania systemów z użyciem kolaboratywnej inżynierii wiedzy”

Niniejsza recenzja została przygotowana na prośbę Dziekana Wydziału Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej Akademii Górniczo-Hutniczej w Krakowie, dra hab. inż. Ryszarda Sroki, prof. n. AGH, wyrażoną w piśmie I.dz. WEAlIB-b/510-6.1/14 z dn. 26 marca 2018 r.

1. Charakterystyka rozprawy

Recenzowana rozprawa mgra inż. Krzysztofa Kutta pt. „Metody projektowania systemów z użyciem kolaboratywnej inżynierii wiedzy” obejmuje 161 stron tekstu i składa się ze streszczeń w języku angielskim i polskim, wstępu, siedmiu rozdziałów, zakończenia, czterech załączników oraz spisu literatury zawierającego 184 pozycje. Koncentruje się ona na problemach wspierania procesu kolaboratywnej inżynierii wiedzy dla trzech obszarów, tj. dla zarządzania jakością, zarządzania zmianami i zaangażowania użytkowników. Najbardziej istotnymi opracowanymi elementami są: 1) sposób podsumowania zmian w jedną metrykę WA 2) graf zmian, oraz 3) zbiór technik grywalizacji dostosowany do procesu. Pracę uzupełnia opis prototypowego zbioru narzędzi wspierającego proces kolaboratywnej inżynierii wiedzy w systemie wiki.

Przedstawione w pracy wyniki uzyskano w ramach projektu AGH finansowanego z dotacji ministerialnej dla młodych naukowców.

2. Problem badawczy i jego znaczenie

Głównym problemem podjętym w pracy jest wspieranie procesu inżynierii wiedzy. Od momentu powstania tego obszaru, który obecnie znajduje coraz szersze zastosowania, podobnie jak w przypadku inżynierii oprogramowania, pierwszoplanową rolę odgrywali inżynierowie. Jednakże dzisiaj rolę głównego uczestnika w projektowaniu systemów z użyciem kolaboratywnej inżynierii wiedzy praktycznie w pełni przejęli eksperci dziedzinowi. Zmiana ta pociągnęła za sobą nowe wyzwania, m.in. potrzebę dostosowania procesu do uczestników, którzy nie mają już dużej wiedzy technicznej, posiadają natomiast wiedzę i doświadczenie eksperckie, często interdyscyplinarne, łączące psychologię, kognitykę, informatykę czy matematykę. Problematyka ta uważana jest za jeden z ważniejszych obszarów i kierunków rozwojowych współczesnej informatyki. Te i jeszcze inne argumenty za tym, iż oceniana praca dotyczy ważnych problemów badawczych w informatyce i sztucznej inteligencji, znajdują się przede wszystkim w rozdziale 1.

Podstawowym zadaniem inżynierii wiedzy jest tworzenie systemów opartych na wiedzy tzw. baz wiedzy przy wykorzystaniu różnych modeli reprezentacji, metod pozyskiwania i przetwarzania wiedzy, dlatego też głównym celem badawczym niniejszej rozprawy było przygotowanie opisu iteracyjnego procesu kolaboratywnej inżynierii wiedzy w postaci środowiska dla tworzenia baz wiedzy. Opis ten został uzupełniony zbiorem metod i narzędzi wspierających tak sformułowany proces w celu stworzenia dobrej jakości bazy wiedzy w rozsądnym czasie i przy użyciu środków, które będą zadowalać jego uczestników. Zatem w dużym uproszczeniu można powiedzieć, że recenzowana rozprawa ma charakter przede wszystkim konstrukcyjny i stosowany. Problem, któremu poświęcona jest rozprawa, ma duże znaczenie praktyczne. Możliwość opracowania skutecznych metod projektowania systemów z użyciem kolaboratywnej inżynierii wiedzy może mieć bardzo duże znaczenie także w wielu innych działach informatyki, takich jak: bezpieczeństwo komputerowe, informatyka medyczna, inteligentna produkcja, inżynieria oprogramowania, symulacja komputerowa, sztuczna inteligencja i webmastering.

3. Wkład autora

W rozdziale 1.3 wymieniono siedem oryginalnych rozwiązań, które opracowano, zweryfikowano i porównano z metodami znanymi z literatury. Są to: 1) struktura inżynierii wiedzy, 2) analiza wymagań stawianych przed systemami w procesie projektowania ich z użyciem kolaboratywnej inżynierii wiedzy, 3) opis metodyki zwinnej kolaboratywnego tworzenia baz wiedzy, 4) ontologia zmian wprowadzanych do bazy wiedzy, 5) grafowy, semantyczny dziennik zmian, 6) metryki oparte na zbiorze technik grywalizacji i 7) prototypowy zbiór narzędzi wspierający proces kolaboratywnej inżynierii wiedzy w systemie wiki.

Wykazano, że proponowane metody lepiej spełniają wymagania stawiane w procesie kolaboratywnej inżynierii wiedzy niż znane, dotychczas stosowane rozwiązania. Uważam, że szczególnie wartościowy jest opracowany i pozytywnie zweryfikowany eksperymentalnie prototypowy zbiór narzędzi wspierający proces kolaboratywnej inżynierii wiedzy oparty na mechanizmie wiki.

Zgodnie z moją wiedzą w literaturze przedmiotu nie raportowano wcześniej rozwiązań tak uniwersalnych jak propozycje przedstawione w niniejszej rozprawie. Także przyszłe prace badawcze, np. nad nowymi metodami wsparcia procesu projektowania, mogą zostać z łatwością umieszczone w kontekście już opracowanych metod.

Jestem przekonany, że wybrane i opisane w rozprawie metody są bardzo starannie przemyślane, a zaprezentowane w dalszej części recenzji uwagi krytyczne mają charakter dyskusji z Autorem niż rzeczywistego wskazywania poważnych błędów. Nie mam też wątpliwości, że przedstawione w rozprawie pomysły mają dużą szansę wybrzmieć szerokim echem w środowisku naukowym.

Obszerna bibliografia zawiera w większości prace aktualne, opublikowane w przeciągu ostatnich dwóch dekad. Dzięki tak starannemu i bogatemu przeglądowi tych prac, głównie w rozdziałach 2-3, czytelnik może łatwo ocenić wkład Autora do omawianej dyscypliny.

Autor w rozprawie powołuje się zaledwie na jedną własną, współautorską publikację [102]. Pomimo opublikowania 6 innych prac bezpośrednio dotyczących tematu rozprawy, nie zostały one ujęte w spisie cytowanej literatury. Jednakże, dopiero w takim całościowym ujęciu dorobek publikacyjny Autora należy uznać za istotny i dobrze powiązany z przedmiotem rozprawy.

4. Poprawność

Badania zostały poprawnie zaplanowane i przeprowadzone. Bardzo dobrze przedstawiono w pracy: problem naukowy (rozdział 1), istniejący stan wiedzy (rozdziały 2-3), zaproponowane metody (rozdziały 4-7), ich walidację i porównanie z rozwiązaniami konkurencyjnymi (rozdział 8) oraz wnioski, wynikające z przeprowadzonych prac (rozdział 9). W mojej opinii Autor dotrzymał staranności w opracowaniu metod, dzięki czemu wyniki pięciu eksperymentów są wiarygodne i interesujące. Ponadto, Autor wykazał się dużą umiejętnością przekazywania uzyskanych wyników w sposób jasny i czytelny. Duża liczba starannie wykonanych tabel, rysunków, listingów i wykresów ułatwia śledzenie otrzymanych rezultatów.

Poniżej zamieszczam listę zauważonych przeze mnie punktów niejasnych lub dyskusyjnych, ale nie wpływają one w żadnym stopniu na moją bardzo dobrą ocenę tekstu pracy.

Uwagi o charakterze ogólnym.

- Zasadniczo, praca posiada tylko jedną poważną wadę, którą stanowi brak próby przedstawienia analitycznego rozwiązania problemu. Rozprawa, jak już pisałem, ma charakter konstrukcyjny i składa się z kilku powiązanych sekwencyjnie modułów, ale w pracy brak jest sformułowania modelu matematycznego, który, dla danego zbioru parametrów, przewidywałby zachowanie się metod na poszczególnych etapach. Autor nie podjął próby przeprowadzenia rozważań teoretycznych. Wydaje się, że istnieją pewne formalizmy, np. dla ontologii zmian, które można by spróbować dostosować do specyfiki opisywanych metod wspierających proces. Zdaję sobie jednak doskonale sprawę z tego, że sformułowanie takiego modelu jest zadaniem niezwykle trudnym ze względu na złożony i bardzo zróżnicowany charakter wybranych obszarów wsparcia.
- Drugą, ale mniej istotną wadą jest brak jasno sprecyzowanej tezy. Może być ona pewnym założeniem, twierdzeniem, pytaniem lub konkluzją odnośnie przedmiotu badań, którą należy w rozprawie udowodnić na podstawie przeprowadzonego studium danego zagadnienia. Jest ona formułowana na podstawie znajomości, zawartej w literaturze przedmiotu, ogólnej wiedzy na dany temat i z reguły dotyczy pewnego fragmentu naszej niewiedzy o tym zagadnieniu. Jest to coś, co wydaje się nam, że wiemy na ten temat, ale

musimy to udowodnić, lub coś, czego nie wiemy, a chcemy wiedzieć. Sformułowanie tezy badawczej i jej rozwinięcie w postaci hipotez badawczych należy do trudniejszych etapów pracy nad rozprawą doktorską. Autor bardzo szczegółowo przedstawił cel i oryginalne osiągnięcia pracy oraz przegląd literatury, i te elementy mogą być pewną wskazówką do sformułowania tezy pracy.

Mam także kilka uwag szczegółowych.

- Tabele 3.1 oraz 9.1 mimo tego, iż częściowo pokrywają się niektórymi kolumnami (SMW, KnowWE, OntoWiki) i wierszami (R1-R23), zawierają błędne sumaryczne liczby spełnionych wymagań. Wystarczy porównać dwie pierwsze kolumny (SMW i KnowWE). Z obu tabel odczytamy 2 x 17 spełnionych wymagań. Jednakże druga tabela została rozszerzona o wiersze (R24-R27), w których w pierwszej kolumnie występują dwa spełnione wymagania, a w drugiej jedno spełnione wymaganie, więc ich sumaryczne liczby powinny się zmienić. Przy tej okazji, warto zauważyć, iż w drugiej tabeli warto było pozostawić kolumnę Loki oraz dodać w ostatniej kolumnie odsyłacze do odpowiednich stron w pracy. Mimo tych niedokładności można odczytać, że proponowana platforma BiFröST uzyskała spełnienie dodatkowych wymagań w stosunku do Loki dla R6, R13, R20, R24, R25, R26, R27 i to należy uznać za duży sukces.
- Rozdział 3.3.4 charakteryzujący środowisko Knowledge Wiki nie przystaje do pozostałych opisów, np. OntoWiki. Praktycznie taki opis powinien być na tyle szczegółowy, by dać podstawy do uzupełnienia tabel 3.1 i 9.1 podsumowujących spełnienie wymagań.
- Ograniczenie z wstępnie określonych w rozdziale 4.2 sześciu do czterech obszarów wspierania procesu inżynierii wiedzy nie było bardzo przekonujące. W szczególności metody wspierania obszaru reprezentacji wiedzy są bardzo interesujące na tle najnowszych osiągnięć w sztucznej inteligencji.
- W rozdziale 6.2 Autor powiązał TOP 5 użytkowników o niskiej jakości wprowadzanych zmian z największą liczbą usuwanych elementów. To powiązanie jest dyskusyjne, podanie nieokreślonego warunku „za często” za wystarczający, nie jest uzasadnione.

- Wprowadzony w rozdziale 5.3 instrument opiniowania i dyskusji jest interesujący, jednakże bazuje na umiejętności współpracy i komunikacji, a w przypadku jej braku lub wysokich kompetencji homogenicznej grupy użytkowników, może prowadzić do niekończącej wymiany opinii. Czy należałoby wprowadzić (automatyczny) mechanizm blokujący taki proceder? Wydaje się, że moderator może nie być w stanie wystarczająco szybko reagować na takie sytuacje bez wsparcia systemowego.
- Problem lepszej bazy wiedzy jest rozpatrywany jako nowy stan bazy z niepogarszającą się liczbą testów jednostkowych oraz wskaźnikiem WA o wartości min. 3. Czy rzeczywiście wprowadzenie niespójności w bazie wiedzy zostanie odzwierciedlone tymi dwoma wskaźnikami? Przy braku analizy formalnej wskaźników i ich zależności nie można tego potwierdzić. Ponadto, testy jednostkowe sprawdzają jedynie proste własności. Czy w tej sytuacji nie powinno rozważyć się użycia testów funkcjonalnych? Z pewnością można natomiast jakość bazy wiedzy zweryfikować innymi miarami, np. podobieństwem semantycznym lub entropią, oraz sprawdzonymi narzędziami, np., DQF, MQM, KBQ, itp.
- Dlaczego wskaźnik użyteczności (credibility) jest wyliczany co 10 minut, a nie przy każdej zmianie, tak jak inne wskaźniki?
- Dla metryki WA wartość 3 odgrywa zasadniczą rolę. Dlaczego? Niestety w pracy nie znalazłem odpowiedzi na to pytanie.
- W rozdziałach 6-7 niektóre listingi nie zostały oznaczone podpisem, a nie ma ku temu żadnych przeciwwskazań. Można też było dodać do pracy nie tylko spisy rysunków i tabel, ale także spis listingów.
- Natknąłem się na nieuporządkowane listy cytowań, np. [150, 62], [154, 10], [77, 20, 141], błędy językowe, np. „hange ontologies”, „alongside with”, „has been present”, „develiped”, „HypperTwitter”, „possibility to also use”, „In particular.” oraz błąd odsyłacza hipertekstowego, np. <http://www.planetmedia.es/proyecto-irene/>, który pokazuje błąd 404.

Pomimo ww. uwag krytycznych ogólny odbiór tekstu jest bardzo dobry. Rozprawa napisana została w języku angielskim w sposób wyjątkowo staranny. Poziom edycyjny tekstu jest wysoki.

W zasadzie, poza kilkoma podanymi wcześniej błędnymi zapisami, w tekście nie występują potknięcia językowe czy stylistyczne. Biorąc pod uwagę objętość pracy, liczba pomyłek jest mała i w żadnym przypadku nie umniejsza wartości pracy, którą oceniam wysoko.

5. Wiedza kandydata

Praca w ogólności odzwierciedla duży poziom wiedzy i doświadczenia praktycznego Autora w przedmiotowej tematyce badań. Bogata i umiejętnie cytowana literatura robi bardzo dobre wrażenie. W szczególności rozdziały 2 i 3 prezentują aktualny stan wiedzy. W rozdziale 2 Autor obszernie opisuje inżynierię wiedzy, poczynsz od rozproszonej, przez kooperatywną, kolaboratywną i kolektywną inżynierię wiedzy. Szczególną uwagę przywiązuje do zaprezentowania aktualnych problemów i wyzwań występujących w tym obszarze badań. Z kolei zagadnieniom semantycznej bazy wiedzy, a w szczególności opartych na mechanizmie wiki, Autor poświęca rozdział 3. Pozostałe rozdziały 4-9 są oryginalnym wkładem Autora w opisywanej tematyce.

O wiedzy i umiejętnościach kandydata świadczą nie tylko uzyskane wyniki w rozprawie, ale także jego publikacje i cytowania. Mgr inż. Krzysztof Kutt jest współautorem 14 referatów konferencyjnych, 1 posteru, 1 rozdziału w książce oraz 1 artykułu (w druku) w czasopiśmie z listy JCR (IF 2016=3,997). Według Google Scholar publikacje te były cytowane 19 razy (stan na dzień 30-05-2018 r.).

6. Inne uwagi

Zalety i mocne strony rozprawy:

- opracowano skuteczne metody i narzędzia wspierania procesu inżynierii wiedzy dla trzech obszarów wsparcia, tj. dla zarządzania jakością, zarządzania zmianami i zaangażowania użytkowników, w tym prototypowy zbiór narzędzi wspierający proces kolaboratywnej inżynierii wiedzy oparty na mechanizmie wiki,
- przedstawiono bardzo dobry opis stanu wiedzy, poparty bogatą literaturą,
- wykonano staranne porównanie ze znanymi metodami alternatywnymi: SMW, KnowWE, OntoWiki i Loki,
- przeprowadzone badania eksperymentalne są rozległe i skrupulatne,

- od strony edycyjnej praca jest praktycznie bezbłędna.

Wady i słabe strony rozprawy to dwa błędy metodologiczne:

- brak próby przedstawienia analitycznego rozwiązania problemu,
- brak sformułowanej tezy pracy.

7. Podsumowanie

Mgr inż. Krzysztof Kutt wykazał się dobrą znajomością literatury przedmiotu rozprawy, umiejętnością projektowania systemów z użyciem kolaboratywnej inżynierii wiedzy, konstruowania metod i narzędzi dla zarządzania jakością bazy wiedzy, zarządzania zmianami i zaangażowania użytkowników, a także zaplanowania i przeprowadzenia ewaluacji eksperymentalnej uzyskanych wyników.

Zdecydowanie uważam, że:

- praca zawiera oryginalne rozwiązanie problemu naukowego,
- kandydat posiada ogólną wiedzę teoretyczną w dyscyplinie informatyka,
- kandydat posiada umiejętność samodzielnego prowadzenia pracy naukowej.

Ważąc zalety i uwagi krytyczne, moim zdaniem, przedstawiona rozprawa w pełni spełnia wymogi Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym (Dz.U. Nr 65 poz. 595 z późniejszymi zmianami) i wnoszę o jej przyjęcie oraz dopuszczenie jej Autora do publicznej obrony.

Danisz Paul