

## **Recenzja rozprawy mgr inż. Macieja Niemira pt. „*Doskonalenie jakości elektronicznego katalogu produktów z wykorzystaniem sztucznej inteligencji*”**

### **1 Podstawa prawna i ocena spełnienia wymagań formalnych**

Bezpośrednią podstawą prawną do przygotowania recenzji stanowi pismo Dziekana Wydziału Inżynierii Zarządzania Politechniki Poznańskiej, dr hab. inż. Marcina Butlewskiego, prof. uczelni. Pismo to informuje o powołaniu mnie, na podstawie Uchwały Rady Dyscypliny Nauk o Zarządzaniu i Jakości Politechniki Poznańskiej, na recenzenta rozprawy Pana mgra inż. Macieja Niemira pt. „*Doskonalenie jakości elektronicznego katalogu produktów z wykorzystaniem sztucznej inteligencji*”, przygotowanej w trybie doktoratu wdrożeniowego, pod opieką promotora dr hab. inż. Beaty Mrugalskiej, prof. uczelni. Opiekunem ze strony instytucji wdrażającej był dr inż. Krzysztof Muszyński. Postępowanie będzie przeprowadzone w dziedzinie nauki społecznej, dyscyplinie nauki o zarządzaniu i jakości.

Wymagania dotyczące rozpraw doktorskich, które stanowią podstawę niniejszej recenzji, zostały określone w ustawie z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce (Dz. U. 2018 poz. 1668 z późniejszymi zmianami, dalej: Ustawa). Zgodnie z art. 187 Ustawy, rozprawa doktorska musi spełniać określone kryteria formalne. W szczególności ust. 3 wskazuje, że rozprawa doktorska może przyjąć formę pracy pisemnej. Wraz z powiadomieniem o powołaniu mnie na recenzenta, otrzymałem wydrukowany maszynopis, co potwierdza spełnienie tego wymogu. Ponadto, ust. 4 art. 187 wymaga, aby praca doktorska zawierała streszczenie w języku angielskim. Autor zamieścił takie streszczenie na stronie 8 swojej dysertacji, spełniając również ten wymóg formalny. W związku z tym mogę stwierdzić, że Pan mgr inż. Maciej Niemir przygotował rozprawę zgodnie z wymogami formalnymi określonymi w Ustawie, a dostarczona 28 października 2024 roku dokumentacja jest wystarczająca do sporządzenia recenzji.

Ustępy 1 i 2 artykułu 187 Ustawy wskazują również czym powinna charakteryzować się pod względem merytorycznym przedłożona praca pisemna, żeby można było uznać ją za rozprawę doktorską. Stosownie do tych zapisów, zasadniczym celem dalszych rozważań będzie odpowiedź na następujące pytania:

- Czy przedłożony maszynopis prezentuje ogólną wiedzę teoretyczną kandydata w dyscyplinie nauki o zarządzaniu i jakości oraz umiejętność samodzielnego prowadzenia pracy naukowej?
- Czy przedmiotem rozprawy jest oryginalne rozwiązanie problemu naukowego lub oryginalne rozwiązanie w zakresie zastosowania wyników własnych badań naukowych w sferze gospodarczej lub społecznej.

Szczegółowa ocena przedłożonej rozprawy Pana mgra inż. Macieja Niemira znajduje się w kolejnych częściach tego opracowania.

## 2 Ocena trafności wyboru problemu badawczego, jego oryginalności i znaczenia

Doktorant trafnie określił lukę i zdefiniował odpowiadający jej praktyczny problem badawczy jako doskonalenie jakości elektronicznych katalogów produktów wykorzystujące technologie sztucznej inteligencji (SI). Tematyka ta jest dobrze dostosowana do obecnych trendów w handlu elektronicznym, który coraz bardziej polega na dokładnych danych o produktach, wpływających na wydajność procesów biznesowych, zadowolenie klienta i konkurencyjność. Problem dotyczy wyraźnej potrzeby branży, a biorąc pod uwagę rozpowszechnienie internetowych rynków i katalogów cyfrowych, poprawa jakości danych ma fundamentalne znaczenie dla sukcesu handlu elektronicznego.

Tradycyjne metody i techniki walidacyjne są dobrze ugruntowane, jednak wykorzystanie SI w praktyce zarządzania danymi i zapewniania jakości jest coraz częstsze. Kompleksowa integracja SI z procesami walidacji tego typu informacji jest więc aktualna i istotna. Połączenie przez Doktoranta klasycznych metod walidacji opartych na regułach z technikami opartymi na technologii SI wraz z opracowaniem wielopoziomowego modelu walidacji danych produktowych stanowi oryginalne, innowacyjne i nowatorskie podejście.

Badania Pana mgr inż. Macieja Niemira są dobrze dopasowane do nowoczesnych procesów biznesowych, które kładą nacisk na transformację cyfrową i ciągłą poprawę jakości. Opracowane i przedstawione w dysertacji wyniki badań i ich praktyczne zastosowanie mają niewątpliwie duże znaczenie dla zainteresowanych podmiotów gospodarczych. Dokładne dane o produktach mają kluczowe znaczenie dla sukcesu handlu elektronicznego, wpływając na takie obszary, jak satysfakcja klienta, funkcjonalność wyszukiwania i zarządzanie zapasami. Poprawa jakości danych bezpośrednio wpływa na wyniki biznesowe. W tym kontekście przedstawiony problem badawczy mieści się niewątpliwie w centrum zainteresowania dyscypliny nauki o zarządzaniu i jakości.

Ogólnie rzecz biorąc, problem badawczy jest dobrze dobrany, aktualny i ma znaczny potencjał innowacji i zastosowań w świecie rzeczywistym.

## 3 Ocena określenia celu, zadań i pytań badawczych

Autor dysertacji przed ustaleniem celu rozprawy doktorskiej dokonał całościowego i systematycznego przeglądu literatury związanego z wybraną tematyką badań. Przegląd ten został profesjonalnie przygotowany z wykorzystaniem powszechnie stosowanej metodyki PRISMA. Na uwagę i wysoką ocenę zasługuje nie tylko przeprowadzenie analizy zgodnie z najlepszymi międzynarodowymi standardami, dokładna dokumentacja przeprowadzonych działań, ale także ciekawa wizualizacja i wartościowa analiza uzyskanych rezultatów.

Cel pracy został zwięźle określony i obejmował poprawę jakości danych o produktach poprzez opracowanie i zastosowanie w praktyce modelu walidacji, który obejmuje rozwiązania SI. Sformułowanie celu podkreśla wdrożeniowy charakter rozprawy i jasno przedstawia kierunek badań i rozważań zawartych w dysertacji. Choć w dalszej części rozprawy podano więcej szczegółów, wydaje się, że wskazanie we wprowadzeniu, jakie konkretnie techniki SI będą wykorzystane byłoby wygodne dla czytelnika.

Również dość ogólne wyrażenie „poprawa jakości danych o produktach”, mogłoby być bardziej doprecyzowane już w pierwszym rozdziale pracy.

Zadania badawcze są odpowiednie i logicznie uporządkowane sekwencyjnie. Obejmują cały proces badawczy od przeglądu literatury po opracowanie modelu i ocenę algorytmów walidacyjnych. Każde zadanie jest zwięźle określone i wyraźnie wspiera ogólny cel. Podobnie jak w przypadku celu pracy, również tutaj wskazanie konkretnych metod lub narzędzi, które zostały użyte (np. rodzaje technik SI, typy analiz statystycznych) ułatwiły by lekturę dysertacji. Dodanie większej liczby szczegółów metodologicznych wzmocniłoby ich przejrzystość i kompleksowość.

Pytania badawcze są dobrze dostosowane do celu i problemu badawczego, obejmując kluczowe obszary, takie jak: normy branżowe (Pytanie 1), problemy z jakością danych o produktach (Pytanie 2), struktura modelu walidacji (Pytanie 3) i wpływ algorytmów walidacji na jakość danych (Pytanie 4). Pytania są trafne, jasne i ukierunkowane na aspekty praktyczne przeprowadzenia badań i wdrożenia ich wyników. Także tutaj nieznaczne zwiększenie szczegółowości przyczyniłoby się do lepszego zrozumienia zawartości umieszczonej w dalszych rozdziałach dysertacji.

#### 4 Ocena metodyki

Autor rozprawy zastosowaną w pracy metodykę badań przedstawił w podrozdziale 1.7, uzupełniając klarowny i zwięzły opis o schemat realizacji pracy doktorskiej (Rysunek 3). Doktorant zdecydował się na zastosowanie metodyki jakościowo-ilościowej zgodnie z modelem CRISP-DM (*Cross Industry Standard Process for Data Mining*) wykorzystywanego do wydobywania wiedzy z danych (Chapman, 2000). Podejście mieszane, od analizy jakościowej do ilościowej, jest dobrym wyborem. Analiza jakościowa pomaga zrozumieć naturę i kontekst problemu. Z kolei, analiza ilościowa umożliwia pomiar wpływu rozwiązań, zapewniając, że wyniki są oparte na dowodach empirycznych.

Wykorzystanie modelu CRISP-DM jest szczególnie adekwatne w przypadku projektu obejmującego analizę danych i rozwiązania oparte na SI. CRISP-DM dostarcza ustrukturyzowanej metodologii do rozwiązywania problemów związanych z danymi, obejmującej etapy takie jak: zrozumienie biznesu, zrozumienie danych, modelowanie i ewaluacja. Wszystkie te kluczowe elementy modelu znalazły się na przedstawionym schemacie realizacji pracy i mają odzwierciedlenie w samej strukturze dysertacji.

Chociaż CRISP-DM stanowi solidny fundament dla badań opartych na danych, Autor mógł szerzej i dokładniej omówić przydatność tego modelu do zadań wykraczających poza wydobywanie danych (*data mining*), co ma miejsce w recenzowanej pracy. Metody badawcze są dobrze dobrane i odpowiednie dla celów rozprawy. Charakter badań stosowanych oraz sekwencyjne podejście jakościowo-ilościowe współgrają z praktycznym kontekstem problemu i wdrożeniowym charakterem dysertacji.

#### 5 Ocena przeglądu literatury i podstaw teoretycznych

Przegląd literatury i podstaw teoretycznych związanych z określonym przez Doktoranta problemem badawczym znajduje się w rozdziale drugim rozprawy. Na ponad 60 stronach Pan mgr inż. Maciej Niemir szczegółowo i wnikliwie przeprowadza czytelnika przez podstawowe pojęcia związane z tematyką dysertacji.

Kolejne podrozdziały są logicznie uporządkowane, Autor zaczyna od ogólnych pojęć jakości (2.1), następnie zawęża swoje rozważania najpierw do problemów z jakością specyficznych dla danych (2.2–2.3), a później do danych produktowych (2.4). W dalszej kolejności Doktorant skupia się na praktycznych wyzwaniach i dostępnych rozwiązaniach w wybranym obszarze badań (sekcje 2.5–2.12).

Przegląd literatury i rozważania teoretyczne noszą silne znamiona praktycznego ukierunkowania zainteresowań Doktoranta. Sporo miejsca poświęcone jest przedstawieniu metod pozyskiwania danych (2.5) czy potencjalnych źródeł błędów (2.7). W ostatnich podrozdziałach (2.10–2.12) Autor kładzie nacisk na walidację danych, bezpośrednio wspierając cel badawczy.

Zawartość tego rozdziału wskazuje na kompleksowe podejście do zrozumienia podstaw teoretycznych i literatury dotyczącej jakości danych produktowych. Uwzględnienie wielu aspektów jakości danych i walidacji oraz liczne i adekwatne odniesienia do międzynarodowej literatury świadczą o dokładnym zbadaniu tego tematu. Zestawienie podstawowych informacji o kluczowych badaniach w tym zakresie zarówno od strony teoretycznej, jak i praktycznej w zwartej, tabelarycznej formie (Tabela 8 i 9) oceniam wysoko. Taka prezentacja porządkuje i systematyzuje zgromadzony bogaty materiał bibliograficzny, pozwala lepiej zorientować się czytelnikowi w omawianych badaniach i ułatwia zrozumienie kolejnych rozdziałów dysertacji.

## 6 Ocena badań własnych Doktoranta

Badania własne Pana mgr inż. Macieja Niemira zostały zaprezentowane w rozdziałach trzecim, czwartym i piątym dysertacji.

Pierwsze dwa badania mają charakter jakościowo diagnostyczny. Badanie trzecie odnosi się do problemu jakości danych produktowych w Global Registry Platform (Rejestr GS1) i jest związane z wdrożeniowym charakterem dysertacji.

### 6.1 Badanie wykorzystania standardów danych produktowych w praktyce

W rozdziale trzecim Doktorant przedstawił badania, w których przeprowadzono szczegółową i rzetelną analizę praktyk stosowanych w sektorze e-commerce w zakresie definiowania i wykorzystywania atrybutów danych produktowych. Badanie skupiło się na identyfikacji różnic w interpretacji i implementacji uznanych standardów oraz zasad dotyczących danych o produktach przez różne platformy handlu elektronicznego.

Uzyskane wyniki wykazały znaczną różnorodność w sposobie definiowania i zarządzania atrybutami produktów w różnych systemach e-commerce. Analizy pokazały też istotne problemy z jakością danych produktowych, które często były niepoprawne, niespójne i nie posiadały jednoznacznych identyfikatorów (numeru *Global Trade Item Number* – GTIN).

W celu usprawnienia handlu elektronicznego Doktorant sugeruje globalną standaryzację atrybutów produktowych i przeciwdziałanie powstawaniu błędów poprzez hybrydowe podejście łączące reguły walidacyjne z algorytmami SI. Uzyskane wyniki tych badań Doktorant opublikował wraz z Promotorką (Niemir i Mrugalska 2021, 2022b).

Przedstawione badanie jest cenne ze względu na diagnostyczny charakter i szeroki zakres analizowanych praktyk biznesowych dużych podmiotów gospodarczych w zakresie jakości danych produktowych. Jednak Doktorant mógłby swoje analizy rozszerzyć o próbę głębszej i pełniejszej oceny potencjału SI w obszarze walidacji danych. Skupiając się na korzyściach płynących z SI, pomija kluczowe wyzwania związane z wdrożeniem tych systemów, takie jak wymagania dotyczące dużych, wysokiej jakości zbiorów danych, ryzyko stronniczości algorytmów, czy też konieczność ciągłego monitorowania i aktualizacji modeli. Uwzględnienie tych aspektów wzbogaciłoby wnioski płynące z badania o bardziej kompleksowy obraz możliwości oraz ograniczeń technologii SI w przedstawianym kontekście.

## 6.2 Badanie problemów jakościowych w danych produktowych

Czwarty rozdział zawiera opis badania obejmującego zidentyfikowanie i analizę problemów związanych z jakością danych produktów, zwłaszcza nazw i kodów produktów GTIN. Doktorant przeanalizował dane z różnych źródeł, takich jak sklepy internetowe, katalogi produktów czy z platform typu *marketplace*. Uzyskane rezultaty pokazały, że dane produktowe zawierają wiele błędów. Autor słusznie zauważa, że liczne błędy w danych produktowych, zwłaszcza w zakresie nazewnictwa i zgodności z obrazami, prowadzą do poważnych konsekwencji, takich jak dezorientacja klientów i utrudnienia w zarządzaniu asortymentem.

Problemy z jakością danych okazały się złożone i, zdaniem Doktoranta, wynikają z braku standardów oraz skutecznych narzędzi walidacji. Do najczęstszych należały niepoprawne kody produktów, niejednoznaczne nazwy i różnice w opisie tego samego produktu w różnych miejscach, rozbieżności między opisami tekstowymi a wizualnymi reprezentacjami produktów. Błędy te wpływają na jakość zakupów poprzez utrudnianie klientom znalezienie właściwych produktów, co może prowadzić do niezadowolenia.

Autor dysertacji zauważył brak podstawowych mechanizmów walidacji, jednakże nie zbadał szczegółowo przyczyn tego stanu rzeczy. Oprócz wskazanych przez Doktoranta braku standardów oraz skutecznych narzędzi walidacji, można było się pokusić o szerszą analizę uwzględniającą takie czynniki jak koszty implementacji, brak świadomości o znaczeniu walidacji, ograniczenia technologiczne. Zrozumienie tych barier jest kluczowe dla opracowania skutecznych strategii poprawy jakości danych w systemach e-commerce.

W celu uniknięcia potencjalnych negatywnych skutków Autor ponownie sugeruje opracowanie jednolitych i powszechnie stosowanych standardów określających jasne zasady opisywania produktów oraz opracowanie lepszych narzędzi i systemów do sprawdzania poprawności danych. Zdaniem Doktoranta, z którym się zgadzam, z analiz wynika, że najlepiej mogą się sprawdzić rozwiązania hybrydowe łączące proste klasyczne techniki z zaawansowanymi technologiami wykorzystującymi SI.

Pełne wyniki tych badań Doktorant opublikował wraz z Promotorką: (Niemir i Mrugalska, 2022a).

### 6.3 Analiza problemu jakości danych produktowych w Rejestrze GS1

Rozdział piąty poświęcony jest analizie jakości danych bazy produktów Rejestru GS1. Badanie jest dobrze ugruntowane teoretycznie i metodologicznie. Autor przeprowadził wszechstronną analizę uwzględniając różne aspekty jakości danych, od błędów składniowych po naruszenia zasad biznesowych. Wykorzystał do tego celu ugruntowane metody badawcze, takie jak eksperckie wywiady półstrukturyzowane czy *data profiling* obejmujący, między innymi, dokładną analizę struktury i schematu bazy. Doktorantowi udało się zidentyfikować wiele różnych typów błędów w danych, co świadczy o dokładności analizy. Doprowadziło to nie tylko do opracowania katalogu problemów jakościowych (Tabela 24), ale także pozwoliło Autorowi zaproponować potencjalne rozwiązania wykrytych problemów (Tabela 26).

W przedstawionej analizie doceniono umiejętność rozpoznawania zarówno prostych, jak i złożonych problemów związanych z walidacją danych, co świadczy o głębokim zrozumieniu tego zagadnienia. Propozycje rozwiązań dotyczą zastosowania zarówno klasycznych reguł i technik, jak i wykorzystania algorytmów do automatycznej walidacji danych, również tych korzystających ze SI. Doktorant zwrócił uwagę na ryzyko fałszywych alarmów generowanych przez algorytmy walidacyjne oraz podkreślił potrzebę ostrożności przy automatycznej korekcie danych.

Pomimo wysokiej jakości przeprowadzonego badania i wykonanych analiz, kilka aspektów mogło zostać bardziej dopracowanych. Przykładowo, Autor wspomina o analizie danych odstających i wzorcach, ale brakuje szczegółowych informacji na temat zastosowanych metod statystycznych w odniesieniu do analizy bazy danych. Z racji tego, że przedstawione analizy dotyczą jednego konkretnego zbioru danych, wydaje się że Doktorant mógł przeprowadzić szerszą dyskusję jak ten fakt może wpłynąć na możliwości uogólniania uzyskanych rezultatów i wniosków z nich płynących.

Chociaż Pan mgr inż. Maciej Niemir zidentyfikował wiele problemów, warto byłoby dokładniej zbadać i omówić przyczyny powstawania poszczególnych błędów, co mogłoby pomóc w opracowaniu bardziej skutecznych rozwiązań. Dodatkowo, Doktorant wprowadził przedstawił różne propozycje rozwiązań, jednak nie przeprowadził analizy kosztów i korzyści związanych z ich wdrożeniem. Oprócz tego, warto byłoby poświęcić więcej uwagi na konsekwencje niskiej jakości danych dla użytkowników systemu oraz dla organizacji jako całości.

Podsumowując przedstawione wyniki badania stanowią wartościowy wkład w obszarze jakości danych produktowych i są solidną podstawą do dalszych badań. Doktorant przeprowadził rzetelną analizę i zidentyfikował wiele ważnych problemów. Przedstawione wyniki są istotne zarówno dla badaczy, jak i praktyków zajmujących się zarządzaniem danymi. Jednakże, aby zwiększyć wartość naukową i praktyczną badania, warto rozważyć przedstawione wyżej rekomendacje.

## 6.4 Sumaryczna ocena badań własnych Doktoranta

Wszystkie trzy badania charakteryzują się wysokim poziomem merytorycznym i metodologicznym i są dobrze przemyślane. Przeprowadzone analizy są szczegółowe i wyczerpujące. Zgromadzone dane zostały dokładnie przeanalizowane, a uzyskane wyniki mają duże praktyczne znaczenie dla opracowywania modelu wspomagającego walidację danych.

Skupienie się na poprawie jakości i spójności danych o produktach bezpośrednio odpowiada na krytyczną potrzebę w nowoczesnym handlu elektronicznym. Wyniki wszystkich trzech badań skutecznie uwypuklają konieczność poprawy jakości danych produktowych w systemach e-commerce, wskazując na szereg problemów, takich jak brak standaryzacji, niespójność danych, błędy w identyfikatorach oraz trudności w automatycznej walidacji danych.

Autor w swoich analizach kładzie nacisk zarówno na poprawność techniczną (np. walidacja GTIN), jak i poprawność semantyczną (np. weryfikacja adekwatności treści w polach tekstowych i obrazach). Doktorant wskazuje na wysoki potencjał metod opartych na SI w zakresie automatycznego wykrywania i korygowania błędów w jakości danych produktowych. Jednocześnie proponuje zrównoważone i kompleksowe podejście do rozwiązywania problemów w tym zakresie poprzez połączenie prostych reguł walidacyjnych z zaawansowanymi algorytmami SI. Takie podejście, wynikające z przeprowadzonych badań i analiz, jest nowatorskie, pragmatyczne i perspektywiczne. Należy jednak zauważyć, że oprócz zalet, wykorzystanie SI wiąże się również z ograniczeniami. Przykładowo, problemy związane ze skalowalnością i wydajnością obliczeniową metod opartych o SI mogą mieć znaczny wpływ na wdrożenie proponowanych rozwiązań w praktyce gospodarczej.

Pomimo tego, że badania są generalnie na wysokim poziomie, w tekście miejscami brakuje niektórych szczegółów na temat metodologii badań. Dostarczenie tych informacji wzmocniłoby wiarygodność i odtwarzalność ustaleń.

Ogólnie, badania własne Doktoranta umożliwiły (i) pozyskanie szerokiej wiedzy na temat praktyk biznesowych w zakresie danych produktowych, (ii) przeprowadzenie dogłębnej analizy problemów związanych z jakością danych w handlu elektronicznym oraz (iii) zostały wykorzystane do zaproponowania innowacyjnych hybrydowych mechanizmów walidacji wykorzystujących rozwiązania oparte na SI.

## 7 Ocena zaproponowanego modelu jego wdrożenia i weryfikacji

### 7.1 Ocena zaproponowanego modelu

Pan mgr inż. Maciej Niemir, w oparciu przeprowadzoną analizę literaturową oraz serię badań własnych, zaproponował w rozdziale szóstym autorski model poprawy jakości danych produktowych, który wykorzystuje SI. Zasadniczym celem tego modelu jest wspomaganie procesu weryfikacji danych oraz dostarczenie użytkownikom szybkiej informacji zwrotnej o potencjalnych błędach.

Proponowany model walidacji danych produktowych obejmuje sekwencyjnie etapy, które mają na celu poprawę jakości danych poprzez ich uporządkowanie, sprawdzenie poprawności i sugerowanie poprawek. Formalnie składa się on z trzech następujących poziomów: Normalizacja danych, Walidacja techniczno-logiczna, Sugestie poprawy jakości. Doktorant w sposób schematyczny przedstawił ten model na Rysunku 16.

Celem normalizacji danych jest uporządkowanie i ujednoczenie danych. Obejmuje on czynności takie, jak: usuwanie zbędnych znaków czy spacji, przekształcanie danych do standardowego formatu, ujednoczanie formatu danych (np. numerów GTIN) oraz używanie spójnej terminologii.

Walidacja techniczno-logiczna ma za zadanie sprawdzenie poprawności danych pod kątem technicznym i logicznym. W ramach tego poziomu sprawdza się czy dane są zgodne ze standardami (np. poprawny format daty, zakres wartości numerycznych) oraz czy dane są ze sobą logicznie zgodne (np. czy jednostka miary odpowiada kategorii produktu).

Ostatni, najbardziej zaawansowany, poziom służy do generowania przez system propozycji wprowadzenia potencjalnych poprawek, nawet jeśli nie ma pewności czy błąd występuje. Działania te mogą dotyczyć: zdefiniowania progów, powyżej których dane są odrzucane (np. zbyt małe zdjęcie), przewidywania wartości (np. klasyfikacja produktu) lub oceny jakości danych na podstawie różnych innych kryteriów (np. czy nazwa jest zrozumiała).

Model uwzględnia automatyzację tego procesu poprzez zastosowanie metod opartych o SI, co pozwala na szybszą analizę dużych ilości informacji. W ramach modelu dostępna jest semantyczna analiza tekstu wraz z generowaniem propozycji konkretnych rozwiązań. Wykorzystanie w modelu technologii SI w trybie dostarczania sugestii zamiast automatycznego wprowadzania zmian danych pozwala uniknąć modyfikacji, które mogłyby prowadzić do błędów. Dzięki automatyzacji w odniesieniu do prostych reguł i zastosowaniu zaawansowanej analizy danych i obrazów, wykorzystanie tego modelu ma doprowadzić do zwiększenia dokładności i spójności przechowywanych danych produktowych. Dodatkowo, Doktorant zakłada, że stosując ten model zwiększy się efektywność weryfikacji danych przy jednoczesnej minimalizacji ryzyka pomyłek popełnianych przez ludzi. W konsekwencji, wykorzystanie proponowanych rozwiązań ma przełożyć się na przyspieszenie procesów biznesowych.

Model ten stanowi oryginalne, zaawansowane, i innowacyjne rozwiązanie dla poprawy jakości danych produktowych. Łączy on w sobie tradycyjne metody walidacji z nowoczesnymi technikami SI. Jego kompleksowe podejście, wykorzystanie SI oraz orientacja na użytkownika są głównymi zaletami tego modelu. Dzięki wykorzystaniu SI, oferuje on wysoką precyzję i efektywność w wykrywaniu i korygowaniu błędów. Model potencjalnie charakteryzuje się znaczną elastycznością i prawdopodobnie może być dostosowany do różnych rodzajów danych i wymagań biznesowych. Jednakże, aby w pełni ocenić jego skuteczność, uniwersalność i potencjał, konieczne jest przeprowadzenie dalszych badań i prawdopodobnie wprowadzenie udoskonaleń.



## 7.2 Ocena weryfikacji modelu

W rozdziale siódmym, Doktorant opisał proces wdrożenia proponowanego modelu i wyniki jego weryfikacji na rzeczywistej bazie danych produktowych. Przedstawiona weryfikacja modelu jest szczegółowa, dobrze udokumentowana i kompleksowa – dotyczy wszystkich poziomów modelu. Autor rozprawy wykazał się głęboką znajomością tematu i umiejętnością zastosowania odpowiednich narzędzi i technik. Ocena skuteczności modelu przeprowadził poprzez ewaluację znacznej liczby przygotowanych walidatorów. Pan mgr inż. Maciej Niemir zastosował w tym celu następujące metryki: czułość, swoistość, dokładność, precyzję, miarę F1 i ujemną wartość predykcyjną (*Negative Predictive Value* – NPV). Wszystkie metryki zostały dokładnie omówione w podrozdziale 6.3.

Wśród walidatorów znalazły się klasyczne, proste techniki oraz zaawansowane modele językowe wraz z innymi technologiami wykorzystującymi SI, co świadczy o holistycznym podejściu do problemu. Dokładnie zbadano wpływ poszczególnych walidatorów na jakość danych, co pozwala na ocenę efektywności całego modelu.

Autor słusznie podkreślił znaczenie normalizacji danych jako pierwszego kroku w procesie walidacji. Użycie wyrażen regularnych oraz słowników znaków Unicode było poprawnym wyborem. Zastosowane walidatory są dobrze przemyślane i obejmują szeroki zakres potencjalnych błędów. Wyjątkowo szczegółowo Doktorant opisał i ocenił walidatory wykorzystane na trzecim poziomie proponowanego modelu, które w znacznej części korzystały z technologii SI. Wykorzystanie SI do generowania sugestii jest niewątpliwie innowacyjne i pozwala na bardziej zaawansowaną analizę danych.

Przedstawiona weryfikacja modelu jest przekonująca i dostarcza solidnych dowodów na jego skuteczność.

## 7.3 Podsumowanie oceny modelu i jego weryfikacji

Przedstawiony model wielopoziomowej walidacji danych produktowych jest dobrze zaprojektowany i przetestowany. Jego zastosowanie może przyczynić się do znacznej poprawy jakości danych w systemach handlu elektronicznego.

Wprawdzie Doktorant omawia różnice pomiędzy zaproponowanym modelem a innymi podejściami, ale porównanie takie mogłoby mieć charakter bardziej obszerny i ustrukturalizowany. Autor mógł także się pokusić o szerszą dyskusję ograniczenia tej weryfikacji do jednego konkretnego zbioru danych, co może wpływać na możliwość uogólnienia wyników. Ponadto, przeprowadzenie analizy kosztów związanych z implementacją i utrzymaniem modelu wzbogaciłoby ocenianą dysertację.

W opisie Autor pokrótce wspominał o *user experience*, ale nie zagłębił się w aspekty związane z doświadczeniem użytkownika w odniesieniu do zaprojektowanego systemu. Zbadanie, w jaki sposób użytkownicy postrzegają systemy walidacji i wchodzi z nimi w interakcje oraz jak wpływa to na ich efektywność pracy, może dostarczyć ciekawych informacji na temat poprawy użyteczności tych rozwiązań. Te dodatkowe wskazówki mogłyby zwiększyć zakres i praktyczną przydatność modelu i opartego o niego systemu informatycznego.

Wydaje się również, że uzupełnienie rozważań Autora o analizę wpływu ekonomicznego ulepszonej walidacji danych na operacje e-commerce byłaby ciekawa. Warto też pamiętać, że stosowanie SI wiąże się też z pewnymi ograniczeniami oraz potencjalnymi problemami i zagrożeniami.

Podsumowując, pomimo powyższych uwag, model ten stanowi wartościowy wkład w obszarze zarządzania jakością danych i potencjalnie może znaleźć zastosowanie w praktyce do poprawy jakości danych produktowych w różnych systemach i branżach.

## 8 Ocena podsumowania

Doktorant, w ostatnim rozdziale rozprawy, przygotował syntetyczne podsumowanie zrealizowanych prac. W systematyczny sposób odniósł się do zdefiniowanych w początkowym rozdziale dysertacji zadań opisując, jak każde z nich zostało wykonane. Następnie, odpowiedział zwięźle, ale treściwie na postawione wcześniej pytania badawcze. Podsumowując wykonane zadania i odpowiadając na pytania badawcze, Autor jeszcze raz nakreślił przebieg całego procesu badawczego uwypuklając jego systematyczny charakter. Pozytywnie oceniam również wyraźne powiązanie celu badawczego, zadań i ich realizacji. Zapewnia to przejrzystość i pozwala czytelnikom lepiej zrozumieć co zostało osiągnięte w pracy i w jaki sposób.

W ramach podsumowania Doktorant zasygnalizował kierunki dalszych badań. Omówienie obszarów przyszłych dociekań nie tylko określa potencjał kontynuowania pracy, ale także wskazuje na świadomość ograniczeń lub luk, których nie udało Autorowi się w pełni wypełnić. Jednak zabrakło mi tutaj odniesienia się Doktoranta wprost do ograniczeń stosowanych metod i uzyskanych rezultatów. Taka dyskusja jest typowa dla prac badawczych i sygnalizuje czytelnikom, że wyniki badań należy interpretować z ostrożnością oraz że wymagane są dalsze badania.

Ostatni podrozdział podsumowania poświęcony jest aspektom wdrożeniowym przeprowadzonych badań i opracowanego modelu. Autor w taki sposób podkreśla praktyczne zastosowanie przedstawionych w dysertacji osiągnięć i ich wpływ na rzeczywistość. Zgadzam się z Doktorantem, że wyniki jego badań mają bezpośrednie zastosowanie praktyczne w obszarze e-commerce. Opracowany model walidacji danych może być wdrożony w celu poprawy jakości danych produktowych w sklepach internetowych, platformach *marketplace* oraz innych systemach zarządzania produktami. Skutkiem tego może być zwiększenie zaufania klientów, redukcja kosztów związanych z błędami w danych, takich jak zwroty towarów, reklamacje i utrata klientów, a także poprawa efektywności procesów biznesowych, związanych przykładowo z zarządzaniem magazynem, marketingiem czy obsługą klienta.

Z wdrożeniem wiążą się też wyzwania dotyczące przykładowo kosztów zastosowania w praktyce modelu obejmujących np. inwestycje w infrastrukturę IT i integrację z istniejącymi systemami. Dodatkowo te składniki modelu, które korzystają z technologii SI, mogą wymagać dużych ilości wysokiej jakości danych do skutecznego działania, a sam model pewnie będzie wymagał regularnej aktualizacji i dostosowania do zmieniających się warunków rynkowych.

## 9 Ocena aspektów formalnych maszynopisu

### 9.1 Ocena układu rozprawy

Recenzowana rozprawa doktorska Pana mgr inż. Macieja Niemira składa się z 255 stron maszynopisu, obejmując tradycyjnie uporządkowane elementy takie, jak streszczenia, wprowadzenie, sześć rozdziałów tematycznych, podsumowanie wraz z kierunkami przyszłych badań, bibliografię i zestawienia rysunków i tabel.

Struktura pracy jest przejrzysta i logiczna, co ułatwia czytelnikowi orientację. Być może niewielkie modyfikacje struktury, polegające na zebraniu badań własnych Autora w jednym, nadrzędnym rozdziale i wydzielenie osobnego rozdziału na dyskusję, w którym znalazłyby się rozważania dotyczące ograniczeń badań i perspektyw na przyszłość, mogłoby jeszcze bardziej zwiększyć czytelność pracy.

### 9.2 Ocena poprawności redakcyjnej

Od strony redakcyjnej rozprawa doktorska została przygotowana z dużą starannością. Autor wykazał się doskonałą znajomością zasad redakcyjnych, co przejawia się w poprawnym opisie tabel, rysunków oraz w rzetelnie przygotowanej bibliografii. Mimo szczegółowej analizy, zauważono jedynie drobne niedociągnięcia stylistyczne, takie jak zbędne kropki po numerach rozdziałów i nieuzasadnione stosowanie numerowanej listy w niektórych przypadkach (np. wymieniając obszary przyszłych badań na str. 237).

### 9.3 Ocena języka i stylu

Język i styl pracy są poprawne i zrozumiałe z marginalnymi niedociągnięciami. Sporadycznie pojawiają się drobne uchybienia stylistyczne, takie jak nieprecyzyjne użycie niektórych wyrażzeń. Przykładowo, na stronie 120 występuje błąd w liczbie rzeczownika (jest „problemu”, powinno być „problemy”).

### 9.4 Ocena wykazu pozycji literaturowych

Doktorant przeprowadził obszerny przegląd literatury, cytując 233 publikacje, w tym artykuły naukowe, monografie i materiały konferencyjne. Dobór źródeł jest adekwatny do tematyki pracy i zapewnia solidne podstawy teoretyczne dla przedstawianych argumentów. Autor wykazał się głęboką znajomością literatury przedmiotu, dobierając do swojej pracy wysokiej jakości publikacje naukowe. Dokładny przegląd literatury, obejmujący zarówno krajowe, jak i zagraniczne źródła, świadczy o solidnym przygotowaniu merytorycznym Doktoranta.

### 9.5 Ocena szaty graficznej

Praca zawiera bogaty zestaw elementów graficznych, w tym 37 rysunków i 61 tabel. Wszystkie są wykonane estetycznie i w wysokiej jakości, co znacznie ułatwia odbiór treści. Liczba i rodzaj zastosowanych elementów graficznych są adekwatne do charakteru przedstawianych informacji.

## 10 Podsumowanie oceny i konkluzja

Moja ogólna ocena pracy przygotowanej przez Pana mgra inż. Macieja Niemira jest jednoznacznie pozytywna. Tematyka tej rozprawy bez wątpienia mieści się w dyscyplinie naukowej nauki o zarządzaniu i jakości.

Doktorant poprawnie zidentyfikował lukę badawczą i sformułował cel, problem oraz zadania i pytania badawcze. Poprzez adekwatny i bogaty przegląd publikacji, przeprowadzenie serii badań jakościowo ilościowych, opracowanie i weryfikację modelu, udało się Autorowi odpowiedzieć na pytania badawcze, zrealizować zadania oraz osiągnąć założony cel badawczy.

Rozprawa jest kompleksowym i innowacyjnym opracowaniem, które wnosi znaczący wkład w dziedzinę zarządzania danymi produktowymi w e-commerce. Proponowany model walidacji oparty na SI jest dobrze uzasadniony i przetestowany w warunkach praktycznych. Autor, oprócz dogłębnej wiedzy teoretycznej pokazał, że potrafi zastosować zaawansowane technologie w rozwiązywaniu złożonych problemów.

Pomimo przedstawionych wątpliwości, zauważonych niedoskonałości czy wskazanych błędów, recenzowana rozprawa doktorska wywarła na mnie korzystne wrażenie i zasługuje na bardzo wysoką ocenę.

Na podstawie przeprowadzonej oceny stwierdzam, że rozprawa doktorska Pana mgra inż. Macieja Niemira spełnia wszystkie formalne wymagania zawarte w stosownych aktach prawnych niezbędne do nadania stopnia naukowego doktora. Autor wykazał się głęboką znajomością tematu, umiejętnością samodzielnego prowadzenia badań naukowych oraz oryginalnością w podejściu do poruszanego problemu. W związku z powyższym, wnioskuję o dopuszczenie Pana mgra inż. Niemira do dalszych etapów postępowania w sprawie nadania stopnia naukowego doktora w dziedzinie nauk społecznych, dyscyplinie nauki o zarządzaniu i jakości.

Ze względu na wysoką jakość dysertacji, z przyjemnością proponuję również wyróżnienie jej.

Rafał Michalski