

Recenzja rozprawy doktorskiej

Łukasz Borchmanna,

zatytułowanej:

Span Identification and Key Information Extraction Beyond Sequence Labeling Paradigm

1. Problem badawczy i jego znaczenie

Przedstawiona rozprawa składa się efektywnie z cyklu 8 publikacji, których spis został przedstawiony na stronie 7. Chociaż na stronie 4 można znaleźć „The thesis consists of seven papers [...]”. Rozbieżność ta jest spowodowana faktem, że praca numer 6 w zestawieniu została opublikowana jako nierecenzowany preprint na serwisie arXiv. Jednak można odnaleźć w sieci informację, że ta sama praca została zaakceptowana do prezentacji podczas warsztatów *Subset Selection in Machine Learning: From Theory to Applications* odbywających się w ramach konferencji międzynarodowej ICML 2021 o najwyższym światowym rankingu. Same warsztaty również podlegały rygorystycznej procedurze recenzji i selekcji. Podsumowując, można uznać, że cykl składa się z 8 prac, które zostały opublikowane lub recenzowane z zachowaniem wysokich standardów naukowych.

Przedstawiony cykl prac charakteryzuje dobra spójność tematyczna i przedstawiają cykl badań, w którym kolejne etapy jasno nawiązują do rezultatów badań z wcześniejszych etapów. Dobrze to podkreśla rozdział 1 rozprawy, który w syntetyczny sposób przedstawia kierunki zrealizowanych badań, powiązania pomiędzy poszczególnymi pracami i zaprezentowane w nich wyniki. Rozdział 1 jest bardzo cenną częścią rozprawy.

Jaki jest najważniejszy problem rozważany w rozprawie?

Główną osią badań w cyklu publikacji jest opozycja pomiędzy paradygmatem etykietowania sekwencji (ang. *sequence labelling*) sygnałów językowych (najczęściej elementów tekstu) przy pomocy metod uczenia maszynowego, a podejściami, których do obiektów językowych (w tym dokumentów tekstowych) jako całości są przypisywane etykiety-metadane charakteryzujące ich treść – schemat nazwany w rozprawie *wydobywanie kluczowych informacji* (ang. *Key Information Extraction*) – oraz metodami rozpoznawania fragmentów komunikatu językowego, potencjalnie wielomodalnego (nazwane w pracy *Span Identification*) jako odpowiadających jednej z predefiniowanych kategorii semantycznych lub wzorcowi. W przypadku paradygmatu etykietowania sekwencji Doktorant odnosi się do bardzo często stosowanego schematu, w którym sekwencjom elementów w danych językowych (np. sekwencjom wyrazów) przypisywane są ręcznie metadane określające ich wzorcową interpretację, a następnie na tej podstawie są trenowane modele klasyfikacji przy pomocy różnych algorytmów uczenia maszynowego.

Doktorant we wstępie wskazuje na liczne jego zdaniem ograniczenia etykietowania sekwencji, takie jak złożoność przygotowania danych treningowych, lokalność przetwarzania, pewna wrażliwość algorytmów na błędy językowe lub konieczność agregowania wydobytych informacji na poziomie całych obiektów. W oparciu o te obserwacje Doktorant argumentuje konieczność poszukiwania schematów rozwoju metod analizy obiektów językowych, w których dane treningowe stanowią obiekty

z przypisanymi metadanymi (wydobywanie kluczowych informacji) lub wzorcowe obiekty opisane kategoriami (rozpoznawanie fragmentów). Ponadto postuluje konieczność formułowania całościowych zadań przetwarzania (ang. *end-to-end*) oraz metod przetwarzania dla nich, tzn. takich schematów przetwarzania, w których wejściem jest treść zadania i obiekt językowy, a wyjściem wydobyte lub przypisane metadane. Schematy te nie powinny zakładać dekompozycji na pośrednie etapy przetwarzania i w całości powinny być realizowane przy pomocy metod neuronowych. Doktorant idzie nawet krok dalej i postuluje konkretny typ architektury neuronowej, tzw. transformery (połączenie neuronowego kodera i dekodera), jako metodę analizy, która umożliwi efektywną realizację schematu wydobywania kluczowych informacji.

Czy ma on charakter naukowy?

Problem sformułowany przez Doktoranta, wpisuje się dobrze w rozwój metod inżynierii języka naturalnego (ang. *Natural Language Processing*) w ciągu ostatnich lat oraz dążenia do lepszego wykorzystania pojawiających się nowych architektur neuronowych w jego ramach, jak również efektywniejszego budowania i wykorzystywania różnych zasobów językowych. Niewątpliwie główna oś badań zdefiniowana przez Doktoranta dotyczy ważnych problemów naukowych. Co więcej główne problemy zidentyfikowane przez Doktoranta są w dużej mierze też zbieżne z ważnym nurtem poszukiwania metod, które wymagają mniejszej ilości danych treningowych lub danych treningowych, które są bardziej ogólne i mniej ściśle powiązane ze szczegółowymi zadaniami przetwarzania języka, np. ang. *a few shot learning* i *zero shot learning*.

Natomiast nie można się zgodzić z Doktorantem, że paradygmat etykietowania sekwencji powinien zostać zastąpiony oboma schematami omawianymi w pracy, jako w pewnym sensie mniej efektywny lub wręcz przestarzały, cytując:

„propozycjach zastąpienia wspomnianego paradygmatu bardziej adekwatnymi metodami, w zastosowaniach, gdzie referencyjnym rozwiązaniem byłyby modele wykorzystujące znakowanie ciągu.”

W bardzo wielu przypadkach, jeżeli nie większości, Doktorant koncentruje się na zadaniach przetwarzania, w których w naturalny sposób całościowe przetwarzanie obiektów językowych lub dopasowanie fragmentów do wzorców jest bardziej zgodne z naturą zadania. Wykorzystanie architektury transformera nie wyklucza zastosowania schematu etykietowania sekwencji w przetwarzaniu, a wręcz przeciwnie, architektura transformera jest obecnie powszechnie wykorzystywana w takich podejściach. Przedstawiona argumentacja za przejściem na schemat *wydobywania kluczowych informacji* trochę za bardzo koncentruje uwagę na zastosowaniach, w których wyznaczanie wystąpień w tekście jest mniej istotne. Nie zawsze też zbiór „własności” (w pracy ang. *properties*) jest ograniczony i z góry zadany. Argumentacja miejscami wręcz ociera się o pewien dogmatyzm, np. „to using encoder-decoder models, which potentially solve all of the mentioned problems.” – to tylko jedna z grup metod, zresztą dość niejednorodna i ewoluująca. Dogmatyzm ogranicza kreatywność.

Czy ma on znaczenie praktyczne?

Prace składające się na cykl w większości przypadków dotyczą zagadnień mających bezpośrednie przełożenie na praktyczne i bardzo innowacyjne zastosowania. Jedna z prac jest poświęcona zdefiniowaniu cennego zasobu językowego do oceny i rozwoju nowych podejść do przetwarzania. Ten aspekt pojawia się też w kilku innych pracach. To bardzo ważny kierunek rozwoju współczesnej inżynierii języka naturalnego – podejście danocentryczne.

Problemy rozważane w poszczególnych pracach składających się na cykl.

W ramach prac składających się na cykl warto wyróżnić kilka szczegółowych problemów badawczych, które są istotne ze względu na realizację przyjętego programu badań i rozwój dziedziny.

W ramach prac [1] (rozdz. 2) i [2] (rozdz. 3) cyklu podjęto problemy rozpoznawania i klasyfikacji wystąpień nazw własnych w polskich tekstach oraz wykrywania i klasyfikacji wystąpień technik językowych charakterystycznych dla języka propagandy w tekstach angielskich. Obywa problemy sformułowano jako wykrywanie i klasyfikacja występowania sekwencji wyrażen językowych, czyli w ramach paradygmatu etykietowania sekwencji

Prace [3] (rodz. 4), [7] (rozdz. 8) oraz [8] (rozdz. 9) dotyczą problemów specyfikowania nowych, trudnych zadań przetwarzania obiektów językowych (przede wszystkim dokumentów tekstowych). W tym praca [7] formułuje nowy problem wydobywania wielu *kluczowych informacji* dla jednego obiektu, a praca [8] kompleksowy problem multimodalnego wydobywania różnorodnych informacji z dokumentów. Obydwie te prace dotyczą również budowy zbiorów danych językowych w sposób pozwalający na wiarygodną ocenę algorytmów przetwarzania, w tym istotność praktyczną uzyskiwanych wyników oceny.

Prace [3] (rozdz. 4), [4] (rozdz. 5) i [5] (rozdz. 6) dotyczą problemu wykrywania w obiektach tekstowych wystąpień wyrażen językowych odnoszących się do kategorii semantycznych zdefiniowanych przy pomocy ograniczonej liczby wzorcowych przypadków. W pracy [3] problem zostaje sformułowany jako problem wykrywania najbliższych sąsiadów, w pracy [4] proponowane jest unikatowe ujęcie jako problemu wykrywania zgodności temporalnych sekwencji oraz praca [5] koncentruje się na problemie efektywności wykorzystania architektury neuronowych transformerów.

W pracy [7] (rozdz. 8) poza nowym zbiorem danych rozważany jest problem wykorzystania architektury transformera do wydobywania wielu *kluczowych informacji* z obiektów tekstowych. Praca [6] podnosi problem wykorzystania informacji o układzie graficznym dokumentu w ramach zadania wydobywania *kluczowych informacji* przy pomocy architektury transformer.

Podsumowując, problemy szczegółowe podjęte w pracach cyklu pokazują na konsekwentną i szeroko zakrojoną realizację programu badawczego przyjętego dla całej rozprawy, a także dotyczą zagadnień bardzo istotnych dla rozwoju całej dziedziny.

2. Wkład autora

Jaki jest najważniejszy wkład prac opisywanych w rozprawie?

Ponieważ rozprawa składa się z cyklu prac wieloautorskich, co jest całkowicie zrozumiałe biorąc pod uwagę wagę podejmowanych problemów i zakres zrealizowanych prac, to najpierw zostanie omówiony i oceniony wkład wnoszony przez przedstawione prace, a później spróbujemy odpowiedzieć na pytanie o osobisty wkład Doktoranta w osiągnięty wynik.

Ze względu na rozległy zakres wykonanych prac, zarówno pod względem tematycznym jak i osiągniętych rezultatów, to poniżej zostaną podkreślone jedynie najistotniejsze elementy wkładu wniesionego w rozwój dziedziny.

Ważnym rezultatem pracy [1] (rozdz. 2) były nowe metody do rozpoznawania i klasyfikacji wystąpień nazw własnych w polskich tekstach, które były jednym z pierwszych dla języka polskiego zastosowań kontekstowych modeli wektorów osadzeń (ang. *contextual embeddings*) w połączeniu z rekurencyjnymi sieciami neuronowymi. Zaproponowane metody uzyskały ówczasie najlepsze wyniki dla języka polskiego w tym zadaniu.

W pracy [2] (rozdz. 3) zaproponowano ciekawe rozwiązanie do rozpoznawania w tekście angielskim wystąpień językowych technik propagandowych polegające na połączeniu kontekstowych osadzeń RoBERTa z dodatkową warstwą klasyfikacji sekwencji (CRF). W zadaniu klasyfikacji wystąpień pod

względem typów technik zaproponowano innowacyjne rozwiązanie polegające na kombinacji trzech modeli transformerowych w postaci złożonej architektury neuronowej i algorytmu ich łącznego dotrenowywania. Zaproponowane rozwiązanie osiągnęło jedne z najlepszych wyników w konkursie SemEval 2020, a sam artykuł nagrodę *Best Paper*.

Cennym rezultatem pracy [3] jest nowy stosunkowo duży zbiór danych – dokumentów prawnych w języku angielskim – definiujący zadanie rozpoznawania wystąpień klauzul prawnych 21 typów. Przeprowadzono również cenne badania nad połączeniem reprezentacji przy pomocy różnych neuronowych modeli językowych z różnymi technikami agregacji i rzutowania wektorów na potrzeby dopasowywania fragmentów tekstu do wzorców w danych treningowych. Wypracowane rozwiązania stanowią wartościowy, trudny do pobicia poziom odniesienia.

Unikatowe rozwiązanie problemu wykrywania fragmentów komunikatu językowego zostało zaproponowane w pracy [4], gdzie w oparciu o ideę dopasowania podsekwencji został zaproponowany algorytm dopasowujący sekwencje w tekście jako odpowiadające sekwencjom w słowniku w oparciu o reprezentację przy pomocy bezkontekstowych wektorów osadzeń (tj. pochodzących z neuronalnych modeli językowych). Kluczowym elementem jest zaproponowany algorytm *Dynamic Boundary Time Warping*, który wyznacza w efektywny sposób semantyczne dopasowanie sekwencji w tekście do sekwencji wzorcowych w bazie reprezentujących poszczególne kategorie, np. typy klauzul czy też typy wyrażen temporalnych. Cenną zaletą zaproponowanego podejście jest możliwość działania przy ograniczonej liczbie przykładów wzorcowych (ang. *a few shot learning*).

Bardzo cenne rezultaty zostały również osiągnięte w pracy [5] (rozd. 6), w której zaproponowano szereg mechanizmów poprawiających efektywność neuronowej architektury transformera. W efekcie osiągnięto nie tylko poprawę efektywności działania, ale, co jeszcze ciekawsze i ważniejsze, możliwość konstruowania modeli językowych wykazujących własności selektywnej reprezentacji dłuższych tekstów. Zaproponowano między innymi wzorce selektywnej uwagi i hierarchicznego kodowania z dostrajaniem procesem selekcji.

Przechodząc do prac z zakresu „wydobycia kluczowych informacji”, w pracy [6] (rozd. 7) zaprezentowano rozszerzenie modelu transformera do reprezentacji tekstu o kontekstową informację o układzie graficznym. Osiągnięto poprawę wyników w wydobyciu informacji z multimodalnych zbiorów danych (obraz i tekst).

W pracy [7] zostało zdefiniowane zadanie wydobywania wielu własności (ang. *multi-property extraction*) z tego samego obiektu językowego (np. dokumentu). W odróżnieniu od typowego zadania wydobywania informacji, zbiór własności jest przypisany do całego obiektu bez wskazywania w anotacji wyrażen językowych reprezentujących poszczególne własności. Jest to ciekawe nowe ujęcie tego problemu, zilustrowane otwartym zbiorem danych. Zaproponowane zostało kilka rozwiązań bazowych oraz nowa, interesująca architektura tzw. transformera o podwójnym wejściu, w której na wejściu podawane są treść obiektu oraz nazwy własności, a na wyjściu generowane są wartości (niekoniecznie występujące w tekście). Zaproponowany algorytm osiągnął najlepszy wynik w testach.

W ramach trzech prac z cyklu ([3],[6] i [8]) zaproponowano nowe, dobrze zaprojektowane zbiory danych językowych definiujące ciekawe i trudne zadania językowe. Zbudowane zbiory danych są otwarte i przyczynią się do rozwoju dziedziny poprzez stymulowanie rozwoju nowych algorytmów.

Szczególnie istotna jest w tym zakresie praca [8] (rozd. 9), w która dotyczy problemu rozszerzonego, wymagającego zbioru danych multimodalnych (tekst i wybrane elementy reprezentacji graficznej), nazwanego *Document Understanding Evaluation*. Nazwa może być trochę przerysowana, ale zbiór reprezentuje szeroko zakrojone zadania wydobywania informacji z multimodalnych dokumentów tekstowych, łączy i rozszerza szereg zbiorów danych i jest bardzo istotną propozycją.

Podsumowując, cykl prac składający się na rozprawę wnosi bardzo istotny, na wysokim poziomie naukowym i w wielu aspektach unikatowy wkład w rozwój metod inżynierii języka naturalnego oraz zasobów językowych.

Jaki jest najważniejszy wkład autora opisywany w rozprawie?

Wkład osobisty Doktoranta w prace omówione w rozprawie z pewnością jest znaczący, a w zakresie elementów twórczych bardzo znaczący. Jednak precyzyjna ocena wkładu osobistego Doktoranta w prace omówione w rozprawie jest trudna, ponieważ deklaracje określające udział współautorów w wielu przypadkach nie rozgraniczają precyzyjnie wkładu, np. w przypadku 6 prac z 8 za „conceptualization and methodology”, które wydają się być kluczowe, odpowiedzialna była poza Doktorantem przynajmniej jeszcze jedna osoba, a w przypadku jednej pracy (gdzie taki aspekt nie jest wyróżniony), to kluczowe, kreatywne działania również są przypisane do więcej niż jednej osoby. Aby dobrze ocenić wkład twórczy samego Doktoranta, konieczne byłoby określenie, za które części merytoryczne metod i planu zasobów był on odpowiedzialny. Niemniej, osiągnięcia zaprezentowane w pracach cyklu są na tyle znaczące, że nawet podzielenie ich na pół sprawia, że w rezultacie otrzymujemy znaczący wkład.

3. Poprawność

Czy stwierdzenia zawarte w rozprawie są godne zaufania?

Wyniki wszystkich prac włączonych do cyklu zostały poddane odpowiedniej weryfikacji naukowej. Poza samym procesem recenzowania, na bardzo dobrych konferencjach i w czasopiśmie, w każdej pracy zaprezentowano wiarygodne wyniki różnorodnych testów rozszerzone o analizę. Również zbudowane zasoby zostały poddane ocenie pod względem jakości.

Czy uzasadnienia są poprawne?

Wszystkie uzasadnienia oraz procesy weryfikacji i oceny zostały przeprowadzone zgodnie z obowiązującymi w dziedzinie standardami. To nie znaczy, że uważny czytelnik nie dostrzeże potencjalnych ograniczeń, czy nawet źródeł problemów. Nie ma rozwiązań idealnych, a pisanie publikacji na najwyższym poziomie światowym jest sztuką ewaluacji pod akceptację recenzentów.

Poczynając od głównego przesłania Doktoranta, że paradygmat etykietowania sekwencji jest w jakimś sensie ograniczony, czy nawet przestarzały i schematy postępowania rozwijane przez Doktoranta przynoszą jakościową zmianę. W pracach z cyklu przedstawionych zaprezentowano ograniczony zestaw zadań przetwarzania języka naturalnego. Pomijając dwie prace [1] i [2], gdzie zastosowano z powodzeniem ‘klasyczne’ etykietowanie sekwencji, to w pozostałych pracach analizowano w większości zadania, które z definicji nie są predystynowane do trenowania klasyfikatorów w oparciu o anotowane ręcznie fragmenty tekstu, np. streszczanie, czy też klasyfikacja większych fragmentów tekstu. Świetnie, że w pracy zaproponowano ciekawe nowe podejścia do tych problemów oraz je lepiej zdefiniowano przy pomocy zbiorów danych. W pracy [4] zmierzono się z zadaniem rozpoznawania jednostek identyfikacyjnych (nazywanych po angielsku *named entity recognition*). Zaproponowany, bardzo ciekawy algorytm, porównano z rozwiązaniem opartym na etykietowaniu sekwencji i stwierdzono na str. 70, że „Both DBA and DBTW outperform the LSTM-CRF baseline in a few-shot setting.” oraz w wykazano istotność statystyczną różnicy. Tylko, że rzut oka w tabelę 5.2 na str. 69 pokazuje, że obydwa modele są lepsze, ale na poziomie miary F1 poniżej 0,4, czyli absolutnie niezadowolającym do żadnego z zastosowań. Co więcej testy przeprowadzono tylko na jednym zbiorze Ontonotes, gdzie jednostki identyfikacyjne są sklasyfikowane do ograniczonej liczby gruboziarnistych klas. Nie wiadomo z treści pracy, czy zbiór testowy zachowuje separację leksykalną w zakresie jednostek w stosunku do podzbioru treningowego. Podobnie nie wiadomo jak podzbiory się mają do siebie pod względem pochodzenia kontekstów występowania. Wreszcie porównano się jedynie z jednym podejściem odniesienia i to specyficznym. Z literatury przedmiotu, wiadomo, że kilka poważnych wyzwań w zakresie rozpoznawania jednostek identyfikacyjnych leży gdzie indziej i że metody oparte na etykietowaniu sekwencji osiągają bez problemu wartości miary F1 powyżej 0,9,

choć oczywiście to też są wyniki z prac opublikowanych na najlepszych gremiach, gdzie obowiązuje „sztuka ewaluacji pod recenzje”. Prawdziwe praktyczne wyniki są niższe (doświadczenie własne recenzenta), ale również wysokie.

Podsumowując ten fragment, pogłoski o śmierci paradygmatu etykietowania sekwencji są przesadzone i nie przesądzone przez wyniki rozprawy, a wszelki dogmatyzm w nauce ogranicza.

Inna kwestia, którą można zauważyć w kilku pracach, np. [3] czy [4] to uzyskanie dobrych wyników z przy pomocy metod korzystających z technik ważenia istotności termów IDF, SIF (*Smooth Inverse Frequency*), czy SVD, podczas, gdy w tych pracach trudno znaleźć informację na jakich zbiorach modele tego typu były budowane. Jeżeli były budowane na całości dostępnych danych, niestety typowa praktyka w dziedzinie, to istnieje niebezpieczeństwo, że pośrednio przenoszą wiedzę z części testowej danych do działania metody. Może to wpływać pozytywnie na wynik testów. Brakuje, np. w [3] i [4] wyjaśnienia tej kwestii, ale to częsta przypadłość prac z literatury przedmiotu.

W przypadku zaproponowanego modelu dwuzródłowego transformera w pracy [7] (rozdz. 8) można przeczytać na str. 127 niepokojące stwierdzenie, że „Both dual-source models failed on the unseen subset.”. Nie zostało to głębiej przeanalizowane, a sygnalizuje niestety potencjalne przeuczenie się modelu, co jest możliwe biorąc pod uwagę trudność zadania i rozbudowaną architekturę neuronową.

W przypadku zaprezentowanych zbiorów danych anotowanych ręcznie w pracach [7] i [8] brakuje podania podstawowego wskaźnika jakości jakim jest zgodność anotatorów oraz procesu jej zapewnienia i dochodzenia do odpowiedniego poziomu (nie jest możliwe uzyskanie wysokiej zgodności od razu przy pierwszym podejściu do zadania). Zamiast tego zbiory oceniane są na podstawie wyników działania metody, ale jest to rezultat pośredni i trudny do interpretacji, jeżeli nie wiemy jaka była trudność zadania dla człowieka. W przypadku pracy [3] zgodność anotatorów jest oceniona na podstawie miary F1, co nie daje nam informacji na ile ona jest przypadkowa.

Na str. 128 możemy przeczytać:

„It obtained Mean-Multi-Property- F of 62.6 on the non annotated test-A set, while achieving 77.5 on the annotated test-B. This discrepancy suggests that the linguists indeed succeeded to remove non-inferable properties.” Niestety możliwe jest też inne wyjaśnienie, że anotatorzy usunęli wszystkie trudne dla siebie przypadki. W przypadku tego zbioru danych nie została zmierzona zgodność anotatorów, ani też nie została wyspecyfikowana procedura procesu anotacji.

4. Wiedza kandydata

Które z rozdziałów (lub sekcji w rozdziałach) rozprawy omawiają istniejący stan wiedzy i dzięki temu potwierdzają ogólny stan wiedzy kandydata w zakresie Informatyki?

Biorąc pod uwagę całość rozprawy, łącznie z bardzo cennym rozdziałem wprowadzającym, Doktorant wykazał się bogatą wiedzą na temat stanu wiedzy i metod stosowanych w inżynierii języka naturalnego, jak w również w zakresie sztucznej inteligencji (np. algorytmy maszynowego uczenia), czy szerzej informatyki (np. zagadnienia złożoności obliczeniowej i optymalizacji).

Jakie obszary tych dyscyplin zostały omówione w tych rozdziałach/sekcjach? Jaka jest opinia recenzenta o jakości tych rozdziałów/sekcji?

Poza pracami z cyklu, rozprawa została opatrzona bardzo dobrym, syntetycznym rozdziałem wprowadzającym, który umiejętnie spiął prace cyklu w jedną całość. Biorąc pod uwagę również bardzo bogaty zestaw załączników, to przedstawiona rozprawa jest bliska bycia jednolitym dziełem, pomimo iż jest oparta na cyklu publikacji.

Jaka jest opinia recenzenta o bibliografii? Na ile bibliografia jest kompletna?

Każda z prac z cyklu zawiera dobry, a w przypadku dłuższych prac bardzo dobry przegląd prac z literatury. Bibliografia wydaje się być kompletna.

5. Inne uwagi¹

Ponieważ przewód doktorski był przeprowadzony na polskiej uczelni i rozprawa jest bronią na polskiej uczelni, to bardzo rozczarujące jest bardzo słaba obecność języka polskiego w tematyce badań zaprezentowanych w rozprawie. Tylko praca [1] dotyczyła języka polskiego. Żaden z zaproponowanych zbiorów danych nie opisuje języka polskiego. Jest to smutny fakt.

Rozprawa jest perfekcyjnie dopracowana od strony edycyjnej. Może być wzorem dla przedstawiania cyklu publikacji jako rezultatu przewodu doktorskiego.

6. Podsumowanie

Biorąc pod uwagę opinie zaprezentowane w poprzednich punktach i wymagania zdefiniowane przez art. 187 Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (z późniejszymi zmianami)² moja ocena rozprawy pod względem trzech podstawowych kryteriów jest następująca:

A. Czy rozprawa zawiera oryginalne rozwiązanie problemu naukowego? (wybierz jedną opcję stawiając znak **X**)

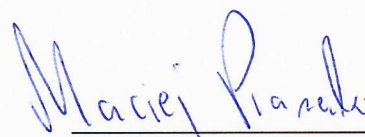
<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zdecydowanie TAK	Raczej TAK	Trudno powiedzieć	Raczej NIE	Zdecydowanie NIE

B. Czy po przeczytaniu rozprawy zgadzasz się, że kandydat posiada ogólną wiedzę teoretyczną w dyscyplinie Informatyka techniczna i telekomunikacja?

<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zdecydowanie TAK	Raczej TAK	Trudno powiedzieć	Raczej NIE	Zdecydowanie NIE

C. Czy kandydat posiada umiejętność samodzielnego prowadzenia pracy naukowej?

<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zdecydowanie TAK	Raczej TAK	Trudno powiedzieć	Raczej NIE	Zdecydowanie NIE


Podpis

¹ Opcjonalnie

² <http://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20190000276>