

prof. dr hab. Andrzej Drzewiński
Instytut Fizyki
Uniwersytet Zielonogórski

Zielona Góra, 31 października 2017 r.

Recenzja rozprawy doktorskiej mgra Andrzeja Jarynowskiego
zatytułowanej „Exploring the dynamics and the complexity of human behavior
using nonlinear physics methods”

Niniejszą opinię przygotowałem na zlecenie Rady Wydziału Fizyki, Astronomii i Informatyki Stosowanej Uniwersytetu Jagiellońskiego w Krakowie. Mgr Andrzej Jarynowski swoją rozprawę doktorską wykonał pod kierunkiem dra hab. Andrzeja Grabowskiego, prof. CIOP-PIB. Obecnie oceniana praca doktorska została poprawiona przez autora zgodnie z decyzją Komisji do Spraw Przeprowadzenia Publicznej Obrony pracy doktorskiej mgra Andrzeja Jarynowskiego działającej pod przewodnictwem dra hab. Pawła F. Góry.

Rozprawa ma charakter interdyscyplinarny, gdzie obszar badań stanowią szeroko pojęte zjawiska społeczne analizowane z wykorzystaniem metod wywodzących się z fizyki układów złożonych. Głównym celem rozprawy było wykazanie przydatności, a nawet niezbędności zastosowania metod obliczeniowych typowych dla nauk fizycznych do konkretnych problemów z zakresu nauk społecznych czy medycznych (epidemiologii). Co więcej, autor zamierzał pokazać, że takie badania interdyscyplinarne nie tylko wnoszą istotny wkład do nauk społecznych, ale dodatkowo służą rozwijaniu samych metod.

Rozprawa mgra Andrzeja Jarynowskiego, mająca postać spójnego tematycznie zbioru siedmiu artykułów, została mi przedstawiona wraz z rozbudowanym przewodnikiem. Całości towarzyszą, wymagane przepisami, oświadczenia współautorów z których wynika, że mgr Andrzej Jarynowski był wiodącym autorem, którego wkład w każdej z prac był nie mniejszy niż 50%.

Przewodnik otwiera słowo wstępne autora oraz pełny spis jego dorobku publikacyjnego oraz wystąpień konferencyjnych. Następnie we *wstępie* autor szkicuje genezę socjofizyki, nauki na pograniczu nauk fizycznych i społecznych, analizującej zachowania społeczne za pomocą narzędzi i pojęć z fizyki statystycznej. Krótki przegląd historyczny służy przybliżeniu czytelnikowi nie tylko najważniejszych postaci, ale także ich motywacji do wkraczania na nowe pola badawcze. Następnie zostały przedstawione podstawowe metody badań złożonych układów społecznych, opartych zarówno o równania różniczkowe, jak i – ostatnio dominujące - symulacje numeryczne oraz analizy statystyczne. W dalszej kolejności autor, tak nieco w gawędziarskim stylu, wprowadza nas w szczegóły metodologiczne swoich badań. To nieco nieortodoksyjne podejście można zaakceptować, a zarazem pokazuje ono swobodę z jaką autor porusza się w swojej rozległej dziedzinie badań.

Całość zamykają refleksje na temat modelowania matematycznego w naukach społecznych i

ekonomicznych uzupełnione o zwięzły, ale w zasadzie wystarczający opis używanego oprogramowania. Umieszczenie tych ostatnich informacji to dobry pomysł, gdyż oprogramowanie to nie jest szeroko stosowane, czego przykładem jest NetLogo będący zarówno językiem programowania, jak i zintegrowanym środowiskiem. Szczególnie dobrze nadaje się ono do modelowania złożonych systemów ewoluujących w czasie a opartych o system agentów.

Podział głównej części przewodnika na dwa rozdziały odzwierciedla pogrupowanie publikacji wchodzących w skład dysertacji na dwa bloki: modele epidemiologiczne oraz analiza sieciowa i modelowanie danych.

W rozdziale 1 autor charakteryzuje dotychczasowe podejścia do modelowania przebiegu epidemii, skoncentrowane bardziej na samym modelu, niż na szczegółowym odtwarzaniu realnego procesu. Skrótowno, ale wystarczająco na potrzeby przewodnika, omawia model SIR opisujący rozprzestrzenianie się choroby zakaźnej poprzez bezpośredni kontakt, gdzie cała populacja jest dzielona na osoby podatne (S), zainfekowane (I) oraz wyzdrowiałe, a w każdym razie już niezakażające (R). Liczebność każdej z tych grup podczas epidemii zmienia się z czasem, a ewolucją zmian mogą rządzić, albo równania różniczkowe (ujęcie deterministyczne), albo procesy stochastyczne. Także omówiony został – zbliżony charakterem – model propagacji informacji, np. plotki czy opinii. W tym drugim przypadku, przy prezentacji dynamiki Glaubera autor przywołuje pojęcie temperatury społecznej układu. Jak rozumiem, może być ona interpretowana jako miara przypadkowości w wyborach jednostki, ale *prosiłbym autora, aby rozwinął ten temat podczas obrony. Między innymi, ciekawi mnie związek temperatury z postawą jednostki, bardziej altruistyczną, bądź bardziej samolubną.*

W drugiej części tego rozdziału zatytułowanej „Aktualny stan wiedzy modelowania epidemiologicznego” autor koncentruje się na zastosowaniu matematycznych modeli do kontroli, przewidywania oraz zapobiegania skutkom epidemii. Z jednej strony mamy doprecyzowanie elementów i interakcji w ramach modelu SIR, z drugiej spojrzenie od strony służby zdrowia, także pod kątem kosztów i efektów leczenia. Również sporo miejsca zostało poświęcone modelowaniu proliferacji zakażeń, zarówno typu afrykańskiego pomoru świń, jak i tych roznoszonych drogą płciową. Zawarte tutaj informacje pozwalają lepiej ocenić trudności jakie napotykamy podczas konstrukcji oraz stosowania modelu rozprzestrzeniania się patogenu. Tym samym lepiej możemy ocenić wiarygodność wniosków stąd płynących. Szczególnie dobrze zostało to przedstawione przy omówieniu zakażeń, które występują w związku z udzielaniem świadczeń zdrowotnych, a w szczególności podczas pobytu w szpitalu. Takie zakażenia, ze względu na zły stan zdrowia pacjentów mają często charakter oportunistyczny, co pociąga za sobą ich dużą dynamikę.

Celem rozdziału 2, jak stoi *expressis verbis*, jest przedstawienie starannie wyselekcjonowanych metod oraz narzędzi służących analizie sieci złożonych, tak aby czytelnik

mógł dogłębnie prześledzić różne, wzajemnie się uzupełniające, wątki badawcze autora. Mimo pewnej „patchworkowości” formy, w zasadzie autor zrealizował swój zamysł. Podstawowe pojęcia oraz procesy ilustrowane są licznymi przykładami zaczerpniętymi z własnych publikacji, choć czasem autor mógłby się bardziej posiłkować innymi pracami.

W rozdziale zapoznajemy się z problematyką modelowania sieci złożonych, które będąc sposobem przedstawiania różnych zjawisk i obiektów oraz relacji pomiędzy nimi, mogą znaleźć zastosowanie przy opisie i badaniu bardzo różnych systemów. W przypadku sieci społecznych mogą być one tworzone przez poszczególnych ludzi, grupy ludzi, firmy czy nawet państwa.

W dalszej kolejności zostały zaprezentowane bezskalowe sieci Barabasiego-Alberta oraz sieci spełniające reguły tzw. małego świata Watts-Strogatza oraz uwypuklono różnice pomiędzy sieciami społecznymi i pozostałymi. Autor także przedstawił proces formowania się sieci, a także omówił wpływ struktury sieci na procesy w niej zachodzące. Zwrócił także uwagę na koncepcję „progu” kluczową dla wielu modeli dotyczących zjawisk kolektywnych.

W **konkluzjach** autor dokonał podsumowania wyników, także w postaci analizy SWOT. Uważam to za udany pomysł pozwalający na syntetyczną ocenę wyników rozprawy stanowiąc zarazem krytyczną ocenę stosowanego warsztatu naukowego. Autor słusznie akcentuje problem niedostatecznych i niejednorodnych informacji bazowych dla budowy modeli, trudności z oceną odpowiedniości i trafności modeli czy zbyt małą nośność otrzymanywnych konkluzji.

Ostatni rozdział zawiera **zestawienie najważniejszych osiągnięć** wchodzących w skład prac stanowiących rozprawę. Całość zamyka rozbudowana bibliografia oraz zestawienia tabel i rysunków.

Obecnie przejdę do omówienia publikacji wchodzących w skład rozprawy doktorskiej mgra Andrzeja Jarynowskiego.

W **pierwszej z prac** (I.1) stanowiących podstawę dysertacji, autor wykorzystując symulacje komputerowe, analizuje ekspansję zakażenia drogą płciową wirusem brodawczaka ludzkiego (HPV) w Polsce, a następnie – w oparciu o uzyskane wyniki – proponuje optymalną strategię przeciwdziałania adresowaną dla służby zdrowia. Trudno o przecenienie wagi tych badań, gdyż HPV szeroko występuje w naszym kraju a zarazem posiada wysoki potencjał onkogenny prowadząc do raka szyjki macicy.

Metoda bazuje na symulacjach agentowych opartych o model SIR, które obejmują okres od 1989 do 2039 roku. Ponieważ stan nosicielstwa może trwać wiele lat, ważne było uwzględnienie przez autora przemian demograficznych oraz długookresowych zmian zachowań społecznych, w szczególności takich jak aktywność seksualna. W artykule widać, jak ważne było oparcie się o rzeczywiste dane jakościowe oraz ilościowe dla właściwego skalibrowania parametrów modelu (takich jak prawdopodobieństwo przekazania patogenu, czy przejścia przewlekłego zakażenia w

postać nowotworową). Warto podkreślić, że w licznych wcześniejszych pracach innych autorów ten aspekt był zaniewany, co znacznie umniejszało praktyczne znaczenie otrzymywanych wniosków.

Na podstawie analizy wyników symulacji najlepszą strategią dla zminimalizowania liczby przyszłych zachorowań jest polityka powszechnych szczepień dziewcząt przed czasem ich inicjacji seksualnej połączona z regularnymi badaniami przesiewowymi. W świetle starzenia się naszego społeczeństwa oraz wzrostu aktywności seksualnej potrzeba takich działań jest bardzo wysoka.

W drugiej pracy (I.2) zostało przeanalizowane rozprzestrzenianie się w warunkach szpitalnych odpornej formy bakterii gronkowca złocistego (szczepu MRSA), który w krańcowym przypadku może wywołać śmiertelną posocznicę. Parametry modelowe zostały ustalone w oparciu o bogatą bazę danych sztokholmskich szpitali dotyczącą pobytów wszystkich pacjentów ze zdiagnozowaną infekcją MRSA w latach 1999-2006. Do symulacji zakażeń szpitalnych został zaimplementowany model kwazi Monte Carlo dla łańcuchów Markowa, dzięki któremu można symulować epidemie na empirycznym zbiorze danych. Sam model należał do modeli epidemiologicznych typu *SI*, gdyż MRSA były odporne na leczenie (raz zarażony pacjent może infekować do końca życia). Prawdopodobieństwo przejścia ze stanu *S* do *I* dla danego agenta w chwili *t* jest określone poprzez formułę wykorzystującą macierz kontaktów między zarażającymi na poziomie sal, macierz kontaktów na poziomie klinik oraz dodatkową macierz uwzględniającą wszystkich potencjalnych pacjentów. Warto podkreślić, że w pierw, używając modelu regresji logistycznej (0-podatny, 1-zainfekowany), autor pokazał brak statystycznie znaczącej zgodności pomiędzy czasem pobytu w szpitalu a prawdopodobieństwem zakażenia (hipoteza zerowa). Ponieważ w takiej sytuacji można jedynie albo odrzucić hipotezę zerową, albo uznać, że wyniki doświadczenia jej nie przeczą, autor wybrał drugą opcję przechodząc do bardziej zaawansowanych metod. *W tym miejscu poproszę autora o krótkie omówienie podczas obrony kryteriów stosowanych przy odrzucaniu „standardowej metodologii statystycznej” na rzecz „narzędzi stosowanych dla układów złożonych”.*

Metoda pozwoliła nie tylko zrekonstruować mechanizm epidemii, ale również wyznaczyć ścieżki zakażeń. Co więcej, została podjęta próba odnalezienia i oszacowania roli potencjalnych „super-roznosicieli” w danym szpitalu. Dodatkowo autor przeanalizował ograniczenia stosowanego modelu oraz możliwości – ale też i koszty – jego rozszerzenia.

Trzecia z prac (I.3) zajmuje się propagacją plotki w sieci powiązań społecznych i szuka odpowiedzi, jak istotna jest tutaj ludzka aktywność typowa dla danej społeczności. Do konstrukcji modelu zostały wykorzystane badania ankietowe przeprowadzone przez Narodowy Instytut Zdrowia Publicznego oraz Państwowy Zakład Higieny określające charakter sieci kontaktów międzyludzkich w Polsce. Dobór 1012 uczestników ankiet, którzy zarejestrowali dane o 16501

kontaktach starał się uwzględnić strukturę demograficzną Polski. Każdy z uczestników otrzymał dzienniczek w którym odnotowywał informacje na temat bezpośrednich kontaktów interpersonalnych, które miały miejsce jednego, losowo wybranego dnia. Ochotnicy odnotowywali takie informacje jak całkowity czas kontaktu z daną osobą w ciągu dnia oraz częstość tych kontaktów. Przeprowadzona analiza pokazała, że sieć kontaktów może być podzielona na kontakty codzienne (72.4% wszystkich spotkań), kontakty rzadkie (16.3%) oraz przypadkowe (11.3%). Z kolei wagę danego kontaktu powiązano z czasem jego trwania, który może wpadać w których z pięciu przedziałów czasowych, gdzie najkrótsze kontakty trwają poniżej 5 minut, a najdłuższe powyżej 4 godzin.

Symulacje zostały przeprowadzone na układzie 100 000 jednostek, zarówno dla sieci statycznej, jak i dynamicznej oraz dla kontaktów o jednakowych, bądź zróżnicowanych wagach, co pozwoliło wygenerować cztery różne modele sieci społecznych. Liczba kontaktów wszystkich typów jest generowana w oparciu o rozkłady prawdopodobieństwa bazujące na danych ankietowych. Dla sieci statycznej kontakty są generowane jednorazowo, na początku symulacji, w oparciu o rozkład prawdopodobieństwa dla wszystkich typów kontaktów. Struktura sieci dynamicznej ewoluuje podczas symulacji za sprawą kreowania na nowo sieci kontaktów rzadkich i przypadkowych w każdym kroku czasowym. Kiedy uwzględniane także są wagi kontaktów, generuje się je w oparciu o procentowy udział kontaktów z danego przedziału czasowego.

W algorytmie kreującym kontakty interpersonalne uwzględniono także tendencję do klastrowania (gronowania), typową dla rzeczywistych sieci społecznych. Opiera się ona na obserwacji, że jeżeli osoby A i B są silnie związane, to w gronie osób silnie związanych z A będzie wiele tych samych osób, co w gronie osób silnie związanych z B . Z kolei inna cecha charakterystyczna dla sieci społecznych, niewielka wartość średniej najkrótszej drogi, przy użytym algorytmie, również była zagwarantowana.

Sam proces rozchodzenia się plotek bazuje na standardowym modelu wprowadzonym przez Daley'a i Kendalla, gdzie każda z jednostek sieci może być w jednym ze stanów: ignorant, plotkarz, albo milczek (wie, ale nie powie). Stan początkowy to jeden, losowo zlokalizowany plotkarz oraz pozostali będący ignorantami. W wyniku interakcji ignoranta oraz plotkarza z prawdopodobieństwem βw_{ij} otrzymujemy dwóch plotkarzy, gdzie w_{ij} oznacza wagę, zaś β atrakcyjność plotki. Prawdopodobieństwo zamiany dwóch plotkarzy (plotkarza i milczka) w dwóch milczków opisuje γw_{ij} . Efektywność propagacji plotki określa liczba osób, którzy plotkę poznali. Jest ona wyznaczana na podstawie tysiąca niezależnych symulacji, osobno dla każdego typu sieci społecznej (stycznej albo dynamicznej, z wagami albo bez wag).

Jak zostało pokazane, proces propagacji plotki zachodzi odmienne niż w przypadku rozprzestrzeniania się epidemii, gdzie dynamiczne podejście przyspiesza propagację patogenu, zaś

dla plotki efekt jest odwrotny.

Jeśli to możliwe, czy podczas obrony autor mógłby przedstawić ewolucję czasową populacji ignorantów, plotkarzy oraz milczków dla różnych typów sieci w zależności od stosunku parametrów β oraz γ .

Kolejna, **czwarta praca** (II.4) zajmująca się różnymi relacjami związanymi z gwiazdami globalnego rynku fonograficznego (najlepiej sprzedającymi się w latach 2003-2011) wpisuje się w zastosowania analizy sieci społecznych. Jej celem było poznanie zasad funkcjonowania tego specyficznego rynku obejmującego zarówno producentów (firmy fonograficzne), produkty (artystów) i konsumentów, powiązanych zarówno pomiędzy sobą, jak i wewnątrz poszczególnych grup. Do analizy powiązań autor wytypował trzydziestkę najpopularniejszych wykonawców, którzy w tych latach sprzedali co najmniej 11 milionów płyt (na dowolnych nośnikach). Zarazem autor pragnął dowiedzieć się, co wpływa na decyzje klientów przy zakupie utworów.

Szukając korelacji pomiędzy artystami, autor rozważał trzy aspekty: podobieństwo w strukturze sprzedaży ich utworów oparte o sprzedaż płyt, podobieństwo okresów popularności na podstawie pozycji na listach przebojów oraz podobieństwo wynikające z oceny eksperckiej. Badając korelacje pomiędzy sprzedażą utworów dla par wykonawców w zsynchronizowanych tygodniowych okresach czasu autor oparł się o macierz współczynników korelacji liniowej Pearsona, co pozwoliło mu zdefiniować miarę odległości w grupie, a tym samym zrekonstruować jej hierarchiczną strukturę. Autor wskazał, że np. w porównaniu z korelacjami pomiędzy cenami różnych akcji na rynku finansowym, korelacje w sprzedaży płyt mają odmienny charakter.

Ogólny wniosek jest taki, że brak jest istotnych korelacji pomiędzy siecią relacji pomiędzy artystami opartą o podobieństwo sprzedaży utworów, siecią relacji opartą o wzajemne umiejscowienie na listach przebojów a siecią relacji opartej o reprezentowany typ muzyki. *W tym ostatnim przypadku może budzić niepokój dominujący udział wiedzy eksperckiej, co wydaje się być wadą modelu stochastycznego. Dlatego poproszę autora o ustosunkowanie się do tego podczas obrony.*

Kolejnej **piąta praca** (II.5) poświęcona jest modelowaniu dynamiki sprzedaży płyt gwiazd muzyki popularnej w oparciu o dyskretne procesy stochastyczne, gdzie bazę empiryczną stanowią tygodniowe zestawienia najlepiej sprzedających się płyt w latach 2003-2013. Autor wskazuje na znaczne podobieństwa analizowanych ciągów czasowych na rynku fonograficznym, z tymi zachodzącymi na rynku energetycznym oraz dla rejestrowanych trzęsień ziemi. Szczegółowa analiza obejmuje zarówno gwałtowny wzrost sprzedaży płyt danego muzyka po premierze (charakterystyczny „kolczasty” pik) oraz wpływ singli poprzedzających premierową płytę, jak i sezonowość sprzedaży. Co więcej, analiza ciągów czasowych przez pryzmat wykładnika Hursta pokazuje, że na rynku fonograficznym i energetycznym procesy „pokonują mniejszą drogę” niż

przy błędzeniu losowym $H < 0.5$ (po wzroście szybko zaczyna maleć), zaś dla trzęsień ziemi zazwyczaj $H > 0.5$. Naturalnie w przeciwieństwie do skoków cen na rynku energetycznym oraz trzęsień ziemi, decyzje o terminie premiery płyty podejmowane są świadomie. W efekcie, przykładowo, czas pomiędzy kolejnymi premierami danego artysty zazwyczaj wynosi dwa lata.

Dla osiągnięcia postawionego sobie celu autor buduje dwa modele stochastyczne: stacjonarny oraz niestacjonarny. Ten pierwszy, aby odtworzyć ową kolczastą strukturę ciągu czasowego, gdzie ponad zaszumiony poziom bazowy wyskakują piki związane z promocją nowej płyty artysty (poziom promocyjny), wykorzystuje kwasygeometryczne błędzenie losowe z powrotem do długoterminowej wartości średniej. Dla właściwego zachowania związanego z obecnością poziomu promocyjnego autor wprowadza dodatkowe elementy związane ze specyfiką czasowych trajektorii kolektywnych zachowań konsumentów muzyki. Z kolei w modelu niestacjonarnym, autor modyfikując parametry modelu, bądź wprowadzając nowe, może uwzględnić związek pomiędzy datami promocji a okresem roku oraz wpływ sprzedaży poprzedniego albumu na dany.

Z pewnością dla tak rozbudowanego modelu, duży problem stanowi ustalenie właściwych wartości parametrów, czemu autor poświęca sporo miejsca. Między innymi, wspomina o funkcji ryzyka, która zazwyczaj opisuje prawdopodobieństwo, że obiekt, który spełnia wymagania w danej chwili czasu, po jakimś, ustalonym czasie przestanie je spełniać. *Czy autor mógłby to dokładniej opisać dla analizowanego przez siebie zagadnienia? Jaki to ma związek z rozkładem Weibulla, gdzie jako podstawowe parametry mamy skalę i kształt a czasami dodatkowo użyty jest parametr przesunięcia.*

Generalnie, zestawienie czasowych trajektorii modelowych i empirycznych dla zachowań kolektywnych konsumentów muzyki pokazuje, że autorowi udało się odtworzyć podstawowe zachowania. Co więcej, model może służyć wytwórniom płytowym np. do lepszego wyznaczenia dnia premiery dla płyty danego wykonawcy. Bardzo ciekawa jest przy tym dodatkowa analiza zakupów w okresie Bożego Narodzenia, która wymyka się zaproponowanym modelom. Zgadzam się z wnioskiem autora, że jest ona pochodną istnienia dwóch kategorii kupujących – świadomych fanów muzyki (zwracających uwagę na jakość produktu i przywiązanych do konkretnych gatunków muzycznych) oraz klienteli „świętecznej” podążającej za trendami.

Szósta z prac (II.6) poświęcona jest tej części rynku fonograficznego, która obejmuje tzw. single promujące nowe albumy oraz ich wykonawców. Autor przeanalizował 17 000 singli, które w latach 1966-2015 lokowały się na listach przebojów w dwunastu krajach europejskich, pokazując jak ich popularność migrowała z jednego kraju do drugiego. Do tego procesu rozprzestrzeniania się opinii, traktowanego jako cykl życia produktu, autor podchodzi jak do opisu ekspansji epidemii. Jednocześnie, analizując popularność poszczególnych utworów, autor wyciąga wnioski o

wzajemnych powiązaniach rynków fonograficznych w różnych krajach europejskich.

Autor uznaje kraj za „zainfekowany” danym utworem, jeśli osiągnie on szczyt listy przebojów w tym kraju, a podstawowym parametrem jest liczba tygodni po którym to następuje, licząc od dnia światowej premiery. Na tej podstawie można wyznaczyć minimalne drzewo rozpinające dla każdego singla ze stu szczegółowo analizowanych. Ponieważ jednak, bazując wyłącznie na cotygodniowych listach przebojów, można je skonstruować na wiele sposobów, autor w sposób losowy konstruuje za każdym razem 1000 takich grafów. Prawdopodobieństwo przyłączenia się konkretnego węzła do danego jest odwrotnie proporcjonalne do różnicy czasów potrzebnych do ich „zainfekowania” oraz proporcjonalne do wielkości rynku fonograficznego w pierwszym z krajów. Takim sposobem autor potrafił zidentyfikować najbardziej prawdopodobne ścieżki propagacji popularności utworów, przykładowo wskazując, że najczęściej Wielka Brytania jest źródłem ekspansji, Holandia najłatwiej adoptuje trendy, a pozostałe kraje można podzielić na dwie grupy – wcześniej, bądź później ulegające wpływom.

Zawarte w publikacji wyniki nie tylko służą wyznaczeniu bliskości kulturowej poszczególnych krajów, specyfikują zachodzące w Europie procesy przenoszenia wzorców i trendów, ale mogą mieć wymiar mocno praktyczny. Przykładowo pozwalają przewidując, kiedy utwór stanie się popularny w danym kraju, lepiej przygotować kampanię promocyjną.

Siódma z prac (II.7) to analiza sieci powiązań międzyludzkich pomiędzy postaciami występującymi w utworach literackich przeprowadzana równoległe drogą algorytmiczną oraz ankietową. Badania literackich sieci społecznych zogniskowane na porównaniu metod autodetekcji oraz percepcji czytelnika miały istotny ładunek nowości. W przypadku badań ankietowych opierały się one na dwóch typach odpowiedzi: zaznaczaniu interakcji podczas lektury utworu oraz wypełnieniu macierzy interakcji wartościami oceniającymi rangę relacji. Otrzymane wyniki, na tym etapie, mają moim zdaniem bardziej walor ciekawostki niż w istocie poznawczy.

Praca doktorska pana Andrzeja Jarynowskiego powstała w bardzo aktywnej grupie badawczej we współpracy z wieloma osobami. Lista referencji, oprócz siedmiu prac wchodzących w skład rozprawy, zawiera jeszcze dziesięć artykułów z tzw. listy ministerialnej, współautorstwo dwóch książek, liczne publikacje spoza listy ministerialnej, kilkanaście opracowań popularnonaukowych oraz wiele wystąpień konferencyjnych. Prace wchodzące w skład rozprawy zostały opublikowane w pismach o stosunkowo niskiej punktacji ministerialnej, lecz zważywszy że dotyczą głównie obszaru badań socjologicznych, wydaje się to zrozumiałe. Do najważniejszych wyników wchodzących w skład rozprawy zaliczam:

- wykazanie na przykładzie zakażeń wirusem brodawczaka ludzkiego realnej możliwości konstrukcji modeli o potencjale predykcyjnym opisujących rozprzestrzenianie patogenów,

pomocnych dla nadzoru epidemiologicznego,

- wykazanie, że przy dostatecznie bogatej bazie danych można wyznaczyć ścieżki propagacji infekcji w warunkach szpitalnych,

- zaproponowanie metody i przeprowadzenie oceny korelacji pomiędzy sieciami relacji pomiędzy gwiazdami globalnego rynku fonograficznego,

- odtworzenie w modelu stochastycznym zachowania krzywych sprzedaży płyt dla gwiazd muzyki popularnej,

- wyznaczenie ścieżek propagacji popularności utworów muzycznych, a tym samym pewnych wzorców kulturowych, wśród krajów Europy Zachodniej.

W rozprawie pojawiają się liczne usterki, głównie językowe, ale także edytorskie, które psują dobre wrażenie, ale nie uważam za celowe wymienianie ich tutaj.

Obecnie, kiedy mamy na co dzień do czynienia z kolosalną ilością danych, wsparcie nauk społecznych przez metody typowe dla nauk fizycznych oraz modele matematyczne wydaje się mocno zasadne. Autor, jednak słusznie zauważa, że w pracach wpisujących się w ten nurt, a zazwyczaj przygotowywanych przez fizyków, często brakowało pogłębionej analizy zjawisk społecznych, zarówno od strony danych jakościowych, jak i ilościowych. Ponieważ autor rozprawy przykłada do tego bardzo dużą wagę i – o czym świadczą liczne przykłady - realizuje w praktyce zalecenie Daniela Bernoulliego ze swojego motto, uważam że główny cel pracy został osiągnięty. Zarazem widać, że autor jest już ukształtowanym badaczem, dysponującym wszechstronnym przygotowaniem oraz umiejętnością samodzielnego prowadzenia pracy naukowej.

Zanim przejdę do ostatecznej konkluzji czuję się zobowiązany podkreślić słuszność propozycji drugiego z recenzentów, który uprzednio zaproponował preredagowanie pracy doktorskiej. W obecnej formie, przy tym doborze artykułów, umiejętności oraz wyniki badań mgra Andrzeja Jarynowskiego prezentują się zdecydowanie lepiej.

Reasumując, stwierdzam, że rozprawa spełnia wymogi ustawy o stopniach i tytule naukowym stawiane rozprawom doktorskim i wnoszę o dopuszczenie mgra Andrzeja Jarynowskiego do publicznej obrony przed Radą Wydziału Fizyki, Astronomii i Informatyki Stosowanej Uniwersytetu Jagiellońskiego.

