



Prof. dr hab. Włodzisław Duch  
Katedra Informatyki Stosowanej WFAiIS,  
i Laboratorium Neurokognitywne, ICNT,  
Uniwersytet Mikołaja Kopernika, Toruń



Recenzja rozprawy doktorskiej mgr. Wojciecha Krzyżanowskiego,  
„Procesy ewolucji kulturowej muzyki w środowisku technologii cyfrowych.”

Strona formalna: niniejsza recenzja została napisana na podstawie uchwały Rady Naukowej Dyscypliny Nauk o Sztuce Uniwersytetu im. Adama Mickiewicza w Poznaniu, z dnia 5.11.2025 roku.

Mgr Wojciech Krzyżanowski przygotował rozprawę doktorską pracując pod opieką dr hab. Piotra Podlipniaka, profesora UAM z Instytutu Muzykologii Wydziału Nauk o Sztuce. Tematyka ewolucji kulturowej muzyki w czasach sztucznej inteligencji jest niezwykle aktualna. Autor ma unikalne doświadczenia, które pozwalają mu zgłębić ten temat z wielu perspektyw: prowadzącego media społecznościowe Polskiego Radia Chopin, śledzenia memów internetowych, roli systemów rekomendacyjnych opartych na sztucznej inteligencji w rozprzestrzenianiu się memów muzycznych, własnej praktyki artystycznej i pracy badawczej muzykologa w ramach studiów doktoranckich. Jest też członkiem międzynarodowej grupy badawczej muzykologów „Many voices”. Z konieczności wiele poruszonych w tej pracy aspektów potraktowano tylko szkicowo, pomimo tego, że całość liczy 378 stron.

Rozprawa składa się z wstępu, czterech obszernych rozdziałów i zakończenia. Jej główne tezy, przedstawione we wstępie, to pokazanie jak memetyka pomaga zrozumieć ewolucję muzyki i technologii muzycznych prowadząc do powstania systemów o rosnącej sprawczości i autonomii, zachęcających ludzi do tworzenia kolejnych wersji utworów. Można było tu przywołać ideę trzech światów Poppera, która przypisuje emergentne własności i autonomię światu kultury, wyłaniającemu się z procesów mentalnych zbudowanych na fizycznych procesach zachodzących w mózgach wielu ludzi. Memy to jednostki informacyjne zdolne do własnej replikacji. Chociaż nie ma tu intencjonalnego dążenia do replikacji sprzyjają jej pewne cechy świata kultury, społeczne uwarunkowania procesów psychologicznych, superweningentnych na zachodzących w mózgach procesach neurobiologicznych, prowadzące do optymalizacji replikacji, często ograniczając przy tym pole decyzyjne człowieka. Już Darwin zauważył, że przetrwanie pewnych słów czy fraz ma charakter doboru naturalnego. Intuicyjne zrozu-

mienie tego procesu pojawia się nawet w piosenkach: „Bo nie ludzie słowa, ale słowa ludzi niosą” (Na językach, Kayah). Muzyka, która przetrwała w różnych okresach historycznych, miała tendencję do replikacji, „niosła ludzi”, dzięki czemu mogła się rozpowszechniać. Przyjmowało się to, co miało charakter memów, które ewoluowały w kierunku rosnącej sprawczości.

Fakt, że tak trudno jest przewidzieć co się stanie muzycznym przebojem czy memem pokazuje nasz brak zrozumienia procesów ewolucji memów. Jest to trudne, bo te procesy działają nie na jednym poziomie, ale we wszystkich trzy światach: od procesów pamięci i uczenia się mózgow przez naśladownictwo, po interpretację stanów mentalnych na poziomie psychologicznym i społeczne milieu sprzyjające rozprzestrzenianiu i utrwalaniu się pewnych procesów.

Takie podejście otwiera szerokie pole do badań nad ewolucją kultury. Nie każdemu się to spodoba, bo wymaga szerokiej wiedzy, wykraczającej poza tradycyjne dyscypliny. To jedna z przyczyn, dla których memetyka napotyka trudności w swoim rozwoju. John Onians, z University of East Anglia, napisał książkę „**Neuroarthistory: From Aristotle and Pliny to Baxandall and Zeki**”. Yale, 2007, ale neurohistorii muzyki jeszcze nie mamy. Podzielałam przekonanie mgr Krzyżanowskiego, że interdyscyplinarne podejście do bio-info-muzykologii będzie odgrywać coraz ważniejszą rolę w rozumieniu zarówno historii jak i współczesnego rozwoju muzyki.

### **Wprowadzenie i metodologia**

Rozprawa doktorska mgr Krzyżanowskiego oparta jest na interdyscyplinarnym ujęciu ewolucji muzyki, łącząc podejście **biomuzykologii obliczeniowej**, dyscypliny w początkowym stadium rozwoju, z ogólną posthumanistyczną refleksją nad kulturą muzyczną. U podstaw tej refleksji leży memetyka, która łączy się z biologią ewolucyjną, kognitywistyką, psychologią ewolucyjną, kulturoznawstwem, matematycznymi modelami ewolucji, teorią informacji, neuronauką i teoriami procesów mentalnych, nauką o systemach, cybernetyką, oraz sztuczną inteligencją i uczeniem maszynowym oraz ich zastosowaniami do przetwarzania języka naturalnego i systemów rekomendacyjnych. To próba powiązania nauk humanistycznych, przyrodniczych i informatyki, którą można określić jako bio-info-muzykologiczną. Opisując muzykę jako dynamiczny system przepływu informacji Autor odwołuje się do czterech znanych modeli: **uniwersalnego darwinizmu Dawkinsa**, teorii aktora-sieci (ANT Lotoura), cyber-

netycznego modelu systemu wiabilnego (VSM Beera) oraz architektury wieloskalowych kompetencji (MCA) Michaela Levina. Nazwa „system wiabilny” nie jest powszechnie stosowana i lepszym tłumaczeniem na polski byłby „system żywotny”. W biologii ewolucyjnej używamy też nazwy „system autopojetyczny”. W tym rozdziale znajdujemy sporo rozważań metodologicznych dotyczących ogólnie humanistyki, filozofii nauki jak i robocze definicje kluczowych pojęć: informacji, informacji kulturowej, ewolucji, memów, oraz sprawczości, którą Autor nazywa też „agencjalnością”. To częściej używana forma niż alternatywna nazwa „agencyjność”, kojarzona z agentem, mającym siłę sprawczą.

Przypisujemy sobie rolę agenta, chociaż wiemy, że na nasze decyzje wpływ ma bardzo wiele czynników. Ewolucja systemów biologicznych jak i kulturowych prowadzi do ich **sprawczości i autonomii**. Udane projekty są powielane, a nieudane zapomniane, nie trafione w gust odbiorców. Mamy obecnie bardzo liczne subkultury, więc ewolucja muzyki prowadzi do powstania niezwykle zróżnicowanych form, które trafiają w gust wybranych odbiorców. Ludzie umożliwiają byt i ewolucję pewnych form sztuki, ale w niewielkim stopniu sprawują nad ich rozpowszechnianiem wolicjonalną kontrolę. Śpiew ptaków czy walenii wskazuje na siłę i inteligencję, ewoluuje w stronę większej sprawczości zdobywania potencjalnego partnera. Dźwięki mowy i muzyki niosą ze sobą informacje o stanie mentalnym, od percepcji (np. audiowizualny efekt kiki-buba) po emocje. Projekt Cantometrix Alana Lomaxa zmierzał do powiązania struktur tradycyjnej muzyki wokalne (używając 37 kryteriów do jej opisu) z cechami organizacji społecznych różnych subkultur. Muzyka umacnia społeczne struktury a jednocześnie jest ich wyrazem. Okazało się, że zbierane przez etnomuzykologów pieśni były w większości kultur wariantami nie więcej niż dziesięciu prototypów. To wskazuje na ewolucję memów muzycznych.

Krytyczne uwagi na temat cytatu Curta Sachsa „najwcześniejszą muzykę znajdziemy u najbardziej prymitywnych ludów” wydają mi się przesadne. W pierwszej połowie XX wieku oczywiście dominowało europocentryczne i rasistowskie spojrzenie, ale nie zmienia to faktu, że tempo ewolucji plemion Afryki Centralnej, gdzie panowały stabilne warunki przez tysiąclecia, było znacznie wolniejsze niż w większej części świata. Mieliśmy tego liczne przykłady na konferencjach EvoLang o wczesnej ewolucji języka, na których było wiele referatów na temat ewolucji muzyki.

W pierwszym rozdziale mamy też dość szczegółowe omówienie 5 wzorców memów internetowych. Memy muzyczne (np. *Rickrolling*, *Keyboard Cat*, *Darude – Sandstorm*) ewoluują w cyfrowym środowisku, zmieniając swoje znaczenia i funkcje, np. dźwięk **MLG Airhorn**, który ewoluował od sygnału sportowych kibiców, przez ironiczny trolling, aż do komunikacji politycznej. Sukces replikacyjny muzyki w sieci zależy od jej podatności na remiksy i transformacje. Czy jest jakaś statystyka memów internetowych pokazująca jakie wzorce są podstawą największej liczby memów? Dlaczego akurat te wzorce zostały wybrane?

Drugi rozdział to **wprowadzenie do memetyki**, historia rozwoju i odpowiedzi na krytykę tej dyscypliny. Na początku tego rozdziału pojawia się pytanie: czy istnieje stochastyczny dryf czy wszystko jest efektem dynamicznych sił selekcji wyznaczających jakiś kierunek? W dłuższym okresie czasu w układach biologicznych selekcja dominuje, ale to wymaga ciągłego nacisku nadającego kierunek zmianom, a ewolucja działa powoli i wpada w ślepe zaułki kiedy warunki się zmieniają. Nie mamy dobrej teorii określającej jak szybko memy mogą się dostosować do zmieniających się warunków, wiele z nich zanika po krótkim czasie. Konkurencja nowych utworów muzycznych, procesów społecznych, wpływ warunków pogodowych czy okresów wzmożonej podróży na rozprzestrzenianie się wirusów to stochastyczne siły selekcji powodujące dryf procesów ewolucji. Czy w przypadku piosenek można mówić o częściowo ślepej generacji rozlicznych wariantów? To dotyczy raczej całych muzycznych stylów, np. pojawienia się muzyki disco, trans czy rapu. Wielokrotnie cytowany w tej rozprawie Michael Levin podkreśla, że procesy zachodzą jednocześnie na różnych poziomach, we wszystkich trzech światach i ich warstwach.

Autor broni memetyki przed zarzutami o nienaukowość, wykazując jej spójność z modelem komunikacji Shannona, zasadą wolnej energii Karla Fristona oraz informatyką i algorytmiką ewolucyjną. W szczególności w 2000 roku w przeglądowej pracy Aunger napisał, że „nikt nie wykazał jeszcze, że [memy] istnieją”. To dość dziwny zarzut, bo to samo można powiedzieć o wszystkich konstruktach kulturowych i psychologicznych, a nawet obiektach fizycznych, takich jak kwarki, których nie da się zobaczyć. W przypadku memów jest to łatwiejsze niż dla takich pojęć jak świadomość czy inteligencja, bo ich nośnikiem są ślady w sieciach neuronowych mózgu, atraktory w modelach pamięci skojarzeniowej (np. cytowana w rozprawie moja praca „*Memetics and Neural Models of Conspiracy Theories*”. *Patterns* 2021). Takie ślady pamięci stanowią część siatki pojęciowej, która decyduje o tym, co przyswoimy i co bę-

cji, od krótkich motywów po całe utwory i muzyczne instytucje. Jan dostrzega procesy wzajemnych oddziaływań na różnych poziomach, ale warto na to popatrzeć z punktu widzenia opisu za pomocą aproksymacji na różnym poziomie szczegółowości. Opis symboliczny też może być przydatny, chociaż nie będzie doskonały. Widać to w analizie ponad 100 kwartetów smyczkowych dokonanej w pracy doktorskiej Andrew Hawketta. Mgr Krzyżanowski krytykuje zbyt jego zdaniem uproszczone podejście definiowania memu jako ciągu melodycznego, ale to jest konkretny, dobry punkt wyjścia do kolejnych, bardziej szczegółowych rozważań, chociaż wybór 3 do 11 nut w tej pracy słabo uzasadniony.

Mamy tu ciekawy problem: czy proces tworzenia muzyki to wyraz głębszej idei, archetypu narzucającego strukturę, czy raczej kombinacji elementarnych memów łączących się ze sobą w nowy sposób? Widzimy to obecnie w dużych modelach językowych (Large Language Models), opisanych w czwartej części rozprawy, które działają na oba sposoby, w procesie dyfuzji tworzące całe obrazy lub w procesie autoregresyjnym tworzące kolejne fragmenty (słowa, tokeny) doklejane do sekwencji. Prawdopodobnie obydwie procesy zachodzą w naszych mózgach.

W rozdziale trzecim znajdujemy **autorską koncepcję systemu muzycznego**. To próba zdefiniowania podejścia systemowego, w którym mamy ewoluujący, sprawczy i dynamiczny układ przepływu informacji. Uczestniczą w nim zarówno ludzie, instrumenty i technologie muzyczne jak i obiekty kultury. Jest to rozproszony system, który wykazuje cechy autonomii w rozumieniu cybernetycznej teorii Mariana Mazura, czyli posiada zdolność regulacji, adaptacji i utrzymania swojej struktury niezależnie od intencji poszczególnych osób, które go współtworzą. Oferuje ludziom odciążanie poznawcze (cognitive offloading) dzięki automatyzacji wielu działań, przenosząc je z ich pamięci na zewnętrzne nośniki i instytucje. Przytoczone w rozprawie badania symulacyjne wykazują, że w interaktywnych środowiskach informacyjnych replikatory powstają niemal nieuchronnie, gwałtownie dominując w populacji. Ewolucja muzyki rozpoczęła się prawdopodobnie od powstania pierwszych muzycznych replikatorów u naszych przaprzodków. Przyczyniło się do tego naśladownictwo, imitacja dźwięków i pamięć słuchowa. Wspomniane wcześniej zjawisko natrętnych melodii powracających w myślach może być tego pozostałością. W teorii powstawania świadomości Juliana Jaynesa prawa półkula mózgu tworzy halucynacje, które rozpoznaje lewa półkula. Początki ekspresji muzycznej są częstym tematem referatów Julian Jaynes Society.

Z tej perspektywy można było prześledzić ewolucję muzyki, od niedoskonałej imitacji po powstanie coraz bardziej precyzyjnej notacji symbolicznej, umożliwiającej długotrwałą precyzyjną pamięć, kumulację złożoności i powstawanie wielkich form muzycznych. Pozwoliło to też na sformułowanie abstrakcyjnych zasad kształtujących praktykę muzyczną. Procesy te są dość dokładnie opisane na wielu stronach rozprawy mgr Krzyżanowskiego. Można się oczywiście spierać co do szczegółów, ale są to z pewnością interesująca przemyślenia wartę szerszej dyskusji.

Współczesne instytucje muzyczne autor analizuje przez pryzmat trzech znanych teorii, znanych z nauk społecznych. Teoria Aktora-Sieci (ANT) jest teorią socjologiczną, zakładającą rozproszoną sprawczość ludzi i aktorów pozaludzkich (instrumentów, tekstów, technologii, teorii), w sieciach podtrzymujących się wzajemnie powiązań. Wiele takich sieci organizowanych jest przez instytucje kultury, podtrzymujące różne style muzyczne, lub skupiające się wokół konkretnych technologii. Rozprawa omawia je w perspektywie historycznej jak i obecnego rozwoju. Autor wiele stron poświęcił w tej części koncepcjom Michaela Levina, który podkreśla wieloskalową architekturę kompetencji (MCA) w biologii, od komórek do całych organizmów aż po społeczne organizacje i całą naturę, z których każda posiada własne cele i kompetencje.

Cybernetycy studiowali warunki potrzebne do przetrwania stabilnych systemów w model Systemu Wiabilnego (VSM), który zapewnia przetrwanie i zdolność do adaptacji na 5 poziomach: jednostek wykonawczych, koordynacji i komunikacji między nimi, integracji całości, strategicznego planowania działań w określonym środowisku, i utrzymaniu tożsamości w długofalowej perspektywie. Jako przykład podano tu filharmonie i duże wytwórnie płytowe. Muzyka akademicka stała się intelektualnie wyrafinowana, ale poza wąskim kręgiem specjalistów memetycznie „bezpłodna”. Wytwórnie płytowe stały się sprawczymi silnikami memetycznym, odkrywającymi i tworzącymi nowe przeboje stosując mechanizmy doboru sztucznego ukierunkowanego na zysk. Odejście od odgórnie kształtowanej mody stało się możliwe dzięki systemom rekomendacji, takim jak Spotify, YouTube czy TikTok, które bez kontroli przez ludzi kierują globalną konsumpcją muzyki. O sukcesie utworu decyduje jego popularność. Prowadzi to do presji selekcyjnej na artystów, którzy zaczynają projektować utwory pod wymogi algorytmów. TikTok pozwala na aktywną prosumpcję, umożliwiając błyskawiczną ewolucję memów muzycznych poprzez masowe remiksowanie. Filtrowanie kolaboracyjne

dziemy rozsyłać dalej. To fizyczny substrat dla replikatorów. Moc sprawczą nie ma sama Informacja tylko jej fizyczna ekspresja, która na nas wpływa. Może do nas dotrzeć przez zmysły, ale też na skutek aktywacji pamięci, jak to widzimy w przypadku natrętnych wspomnień, nazywanych popularnie earworms, lub fachowo INMI (Involuntary Musical Imagery), omawianych dokładniej w dalszej części rozprawy. Tu warto było zacytować „Muzokofilę” Olivera Sacksa. Percepcja jest w znacznej mierze automatyczna, więc pasywna, ale zrozumienie i reakcje są wynikiem działania aktywnego układu poznawczego.

Takie podejście pokazuje również użyteczność predykcyjną pojęcia memów, które tworząc element w siatce pojęciowej pamięci umożliwiają nowe skojarzenia i miejsce na zmodyfikowane wersje memów, które się do tej siatki mogą łatwo przyłączyć w wyniku wewnętrznych procesów lub zewnętrznej stymulacji. To rozróżnienie na **i-memy** (wewnętrzne wzorce w umyśle) oraz **e-memy** (zewnętrzne manifestacje, jak nagrania czy nuty) zrobił McNamara w cytowanej pracy z 2009 roku. Memy muzyczne wykazują cechy przypominające intencjonalne zabieganie o własny byt – wygrywają te utwory i wzorce, które skuteczniej prowokują ludzi do tworzenia własnych kopii. Mamy wiele modeli matematycznych rozprzestrzeniania się informacji w sieciach, takich jak przywołany w tym rozdziale model Niezależnych Kaskad (Independent Cascade) oraz Model Progu Liniowego (Linear Threshold Model), ale też modele oparte na sieciach Isinga, Hopfielda, socjofizyce i ekonofizyce. Mamy wiele czasopism technicznych na temat algorytmów ewolucyjnych a także memetycznych. W opisie czasopiśmie „Memetic computing” czytamy „Ewolucja kulturowa ucieleśnia memy jako podstawowe jednostki informacji przekazywanej w ramach zdolności umysłowych i propagowanej wśród populacji poprzez naśladownictwo i adaptację.” Informatycy nie przejęli się krytyką filozofów i kulturoznawców, rozwijając matematyczne modele memetyki. Mam wrażenie że krytyka była kolejnym przykładem działania memetycznej ewolucji na poziomie mentalnym. Powielamy argumenty, które są dla nas wygodne, to typowy błąd potwierdzenia i niechęć do zmian, utrwalana przez memy w tradycyjnych dyscyplinach. To dodatkowe argumenty do tych opartych na rozważaniach Dennetta i kulturoznawców.

W dalszej części tego rozdziału mamy odwołania do teorii informacji Shannona i zasady wolnej energii, rozpowszechnionej przez Karla Fristona. To mylna nazwa, bo nie chodzi tu o energię, tylko minimalizację niepewności przewidywań. Model wewnętrzny systemu poznawczego ciągle się zmienia, optymalizacji ulega struktura pozwalająca na antycypację do-

chodzącej do niego informacji w najbardziej oszczędny sposób. Mierzy to informacja algorytmiczna, powszechnie stosowana w uczeniu maszynowym, a nie informacja Shannona. Wpływ muzyki (lub szerzej, środowiska) nie wynika tylko z przekazywanej informacji, ale zachodzi dzięki interakcji ucieleśnionej (embodied), usytuowanej w określonym kontekście (embedded), rozszerzonej przez możliwości stwarzane przez środowisko (extended) i enaktywnej (enactive), czyli reakcji wywołanych przez aktywne działanie.

Ciekawe rozróżnienie na memy racjonalne i nieracjonalne wprowadził znany fizyk David Deutsch w swojej książce z 2011 roku (nie ma jej w spisie literatury), poświęcając cały rozdział memetyce. Umiejętność grania muzyki określił jako mempleks, czyli kompleks memów. Teoria złożzeń (assembly theory) cytowana jest błędnie. Uwagi na temat semantyki dystrybucyjnej, która jest podstawą analizy języka naturalnego i dużych modeli językowych, można odnieść do muzyki, ale raczej jej fragmentów, a nie całych utworów. W sumie ten podrozdział przedstawia imponująco wiele teorii z różnych dyscyplin mogących coś wniesić do opisu procesów memetycznych.

W podrozdziale 2.5 krytyce poddana została koncepcja Stevena Jana, który w 2007 roku opublikował książkę „The Memetics of Music: A Neo-Darwinian View of Musical Structure and Culture”. To poważne dzieło, które niestety nie zyskało na popularności, chociaż proponuje złożony paradygmat memetyki muzyki wzorowany na genetyce, a nawet szereg eksperymentów do testowania proponowanych rozważań. W swojej rozprawie doktorskiej mgr Krzyżanowski przeprowadził krytyczną analizę 20 najważniejszych tez książki Jana. Odróżnia się w niej memy od mnemonów, czyli takich śladów pamięci, które nie są memami. To oczywiście płynne rozróżnienie, które w ramach teorii neuronauk sieciowych można powiązać z połączeniami w pamięci skojarzeniowej. Stopień rozpowszechnienia i kopiowania memów może przyjmować wartość od zera do bardzo dużej liczby. Mogą to być prywatne memy pomiędzy parą osób lub rozpowszechniane publicznie, przyjmujące wiele wariantów. Również podobieństwo memów muzycznych trudno jest określić w skwantyzowany sposób. Jan zbyt sztywno próbował trzymać się analogii genetycznych, ograniczając się do analizy zachodniej muzyki klasycznej korzystającej z zapisu nutowego. W jego rozumieniu dzieło muzyczne można uznać za chwilowy sojusz memów. Lepszym określeniem, zaproponowanym przez Michaela Levina do opisu morfogenezy w biologii rozwojowej jest podkreślenie **wieloskalowej architektury różnych kompetencji**, rozproszonych procesów na różnych poziomach organiza-

wspomaga tworzenie licznych sieci wspólnych zainteresowań. W ogromnych sieciach społecznościowych obserwujemy wpływ odbiorców na rekomendacje pasujących do ich gustów memów, a te wpływają na artystów, którzy produkują ich więcej. Widzimy więc integrację złożonych oddziaływań na wielu poziomach systemu, pętle sprzężenia zwrotnego w których sprawczość jest rozproszona.

Rozdział trzeci wieńczy analiza najnowszych technologii generatywnej AI. Omówienie systemów sprzed ponad 5 lat (MuseNet, Agents) pokazuje tylko, jak bardzo rozwinęła się technologia w ostatnich latach. Dobra tokenizacja dźwięków i obrazów jest znacznie trudniejsza niż pisanego tekstu. Wbrew powszechnym opiniom nowsze systemy nie tylko kopiują styl, ale dzięki modelom dyfuzyjnym (np. *Riffusion*) potrafią generować autentycznie nowe struktury dźwiękowe tworząc je w abstrakcyjnej przestrzeni ukrytej (latent space), zanim nie zostaną zamienione na konkretne spektrogramy i wyrażone w postaci dźwięków. Autor nazywa to halucynacjami w stylu znanego memu „secret horses”. Mieszanie różnych dźwięków tworzy nowe brzmienia, których nie potrafiły stworzyć elektroniczne syntezatory. Pojawienie się nowych generatorów muzyki (Suno, Udio, ElevenLabs, Sonauto AI, Mureka, Mubert, AIVA, Sounddraw, ACE Studio, Yue AI i wielu innych) i wiele nowinek technicznych, takich jak protokoły MCP pozwalające łączyć AI z muzycznymi aplikacjami, otwiera nowe pole dla tworzenia muzycznych memów w każdym możliwym znanym jak i całkowicie oryginalnym stylu. Powstaje też wiele wyrafinowanych utworów audiowizualnych. Agencjalność systemów rekomendacyjnych przyczynia się do dominacji streamingu w branży muzycznej, jak i coraz większego udziału muzyki generowanej przez systemy AI w ofercie streamingowych platform. Już w 2023 roku Spotify oferowało usługę osobistego asystenta (AI DJ), który wybiera nam utwory i może generować nowe w stylu, który lubimy. Nie ma wątpliwości, że wpływ systemów AI na sztukę i muzykę będzie coraz silniejszy. Wkrótce mogą się pojawić cyfrowi artyści, a wprowadzenie systemów z długotrwałą pamięcią neuronową pozwoli na rozwijanie konkretnych osobowości o specyficznych upodobaniach. Pada tu propozycja by nowe replikatory oparte na LLM-ach nazwać **lemami** (na cześć Stanisława Lema). Powstają niezależne agenty AI, potrafiące samodzielnie wnioskować, planować i używać narzędzi, ale też na podstawie obserwacji konsumpcji muzyki odkrywać upodobania ludzi. Mamy więc memetyczną ko-ewolucję, kontrolowaną do pewnego stopnia przez duże korporacje, które na tym zarabiają.

W czwartym, ostatnim rozdziale rozprawy znajdujemy eksperyment w zakresie semantyki dystrybucyjnej muzyki. Wyszukano 10 lub 30 najbliższe semantycznie słowa skorelowane ze słowem „music” w kolejnych dekadach, od 1800 do 1990 roku. Listy tych słów przedstawiono na wielu stronach. Byłoby znacznie lepiej przedstawić zmiany list słów w czasie w formie graficznej, albo w tabeli ilustrującej zmiany pozycji poszczególnych słów. Praktyczną ilustracją możliwości systemów generatywnych jest **vibe composing** – eksperyment z nową metodą tworzenia muzyki, w której człowiek wchodzi w rolę „ogrodnika” lub kuratora, sterując procesami generatywnymi AI za pomocą języka naturalnego. Eksperymenty autora z modelami ChatGPT, DeepSeek oraz Google Gemini pokazały, jak łatwo AI na podstawie promptu w języku angielskim (choć równie dobrze działają prompty w języku polskim) jest w stanie stworzyć w pełni funkcjonalne, złożone narzędzia do produkcji muzyki sterowane przez liczne parametry. W tej pracy doktorskiej bardziej by pasowały programy oparte na ewolucyjnych algorytmach, w których użytkownik pełniłby rolę kuratora, a system starał się doskonalić początkowy program. Memetyka może też dotyczyć komputerowych programów, co wiadać po licznych aplikacjach mobilnych.

#### **Podsumowanie i wnioski.**

W omawianej rozprawie doktorskiej ewolucja kulturowa muzyki przedstawiona została w perspektywie programu badawczego memetyki, w której kluczowe znaczenie mają replikatorzy. Doszliśmy do momentu, w którym systemy informacyjne uniezależniają się od człowieka. Systemy rekomendacji i generatywna sztuczna inteligencja przejmują kompetencje, które dotąd należały do kompozytorów i kuratorów. Muzyka staje się systemem, który coraz bardziej przypomina intencjonalny byt zabiegający o własne przetrwanie w technologicznym ekosystemie. Zrozumienie tych procesów jest kluczowe nie tylko dla muzykologii, ale także dla bezpieczeństwa kulturowego i społecznego w erze algorytmów. Rekomendery na Spotify to potężne silniki, które bez przerwy „pompują” muzykę do naszych uszu, ucząc się, co nas najbardziej uzależnia, by kręcić się jeszcze szybciej.

Bez wątplenia w tej pracy jest sporo kontrowersyjnych stwierdzeń, wymagających dogłębnych dyskusji jak i badań empirycznych, ale takie spojrzenie na muzykologię, zarówno z perspektywy historycznej jak i obecnych szybki zmian, jest bardzo cenne. Autor opisał szereg mniej znanych wątków rozwoju nauki, które mogą się przyczynić od lepszemu zrozumieniu ewolucji muzyki. To naprawdę bardzo interdyscyplinarne obszary, które nie było łatwo

przedstawić w spójny sposób. W tej recenzji dodałem do kilka uwag i nowych wskazówek ukazujących dodatkowe kierunki warte analizy.

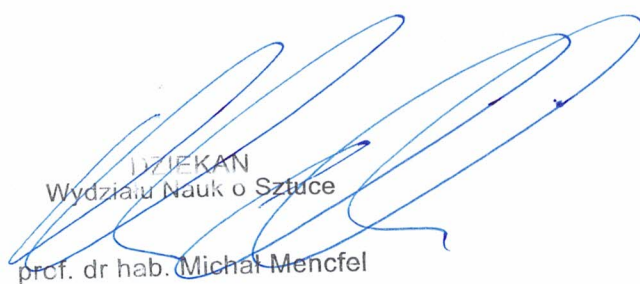
Mam też szereg uwag krytycznych. Układ pracy warto nieco uporządkować. Ewolucyjną korzyść odnosi nie sama kultura, ale infrastruktura, która jest jej nośnikiem – firmy Apple, Samsung, Huawei, a także platformy takie jak TikTok, Youtube, Facebook, Instagram. Bibliografia, chociaż obszerna, ma liczne braki, wiele cytowanych w tekście pozycji zostało w niej pominiętych, np.: Alan Lomax nie jest cytowany, zapewne chodziło o „Folk Song Style and Culture” (1968) i „Cantometrics: An Approach to the Anthropology of Music: Audiocassettes and a Handbook”. UC Berkeley, 1976. Brakuje też cytowanych książek Davida Deutscha (2011), jak i książki Stevena Jana (2016). Sharma i inni (2023) jest za to cytowany dwukrotnie, podobnie artykuł Levina (2022).

Podsumowując, warto będzie po pewnych poprawkach opublikować tę rozprawę w formie książkowej. Z pewnością spełnia ona z nawiązką formalne jak i zwyczajowe wymogi dla prac doktorskich. W mojej opinii zasługuje też na wyróżnienie.



Prof. Włodzisław Duch,

Toruń, 1/01/2026



WZIEKAN  
Wydział Nauk o Sztuce  
prof. dr hab. Michał Mencfel

