

Uniwersytet im. Adama Mickiewicza w Poznaniu
Wydział Nauk o Sztuce

Wojciech Krzyżanowski

***Procesy ewolucji kulturowej muzyki w środowisku
technologii cyfrowych***

ROZPRAWA DOKTORSKA

DZIEDZINA NAUK HUMANISTYCZNYCH
DYSCYPLINA NAUKI O SZTUCE

Praca doktorska napisana pod kierunkiem:

Prof. UAM dr hab. Piotra Podlipniaka



Poznań, 2025

Tylko ta muzyka
tylko ona to sprawiła,
mnie tu wymyśliła...

-Ważka G21

Spis treści

Spis treści.....	3
Wstęp.....	4
1.1 Metodologia.....	20
1.2 Dodatkowe definicje.....	23
1.3. Ewolucyjne wątki i badania w muzykologii.....	25
1.4. Memy internetowe.....	34
1.5.1. Dancing Baby.....	43
1.5.2. Rickrolling.....	46
1.5.3. Keyboard Cat.....	49
1.5.4. Darude – Sandstorm.....	54
1. 5. 5. MLG Airhorn.....	59
2.1. Wprowadzenie memetyki.....	64
2.2. Istotne aspekty memetyki.....	66
2.3. Odpowiedzi na zarzuty.....	70
2.4. Hipotezy pomocnicze.....	85
2.5. 20 tez memetyki muzyki Stevena Jana.....	98
2.6. Kontynuator Jana – Andrew Hawzett.....	136
3.1. Autorska koncepcja systemu muzycznego.....	140
3.2.1. Stadium proto-earworms.....	142
3.2.2. Przejście od kultury oralnej do kultury piśmiennej w muzyce Zachodu.....	153
3.2.3. Teoria muzyki jako presja selekcyjna.....	166
3.2.4.1. Wprowadzenie ANT: specjacja muzyki akademickiej.....	168
3.2.4.2. Sprawcze instytucje od kościoła po MTV.....	176
3.2.4.3. MCA, ANT, VSM.....	183
3.2.5. Systemy Rekomendacji Muzyki.....	194
3.2.6.1. Muzyczne AI – wprowadzenie, rys historyczny.....	221
3.2.6.2. Muzyczne AI - case studies.....	233
3.2.6.3. Muzyczne AI - przegląd technologiczny.....	242
3.2.6.4. Muzyczne AI - perspektywa agentowa.....	254
4.1. Semantyka dystrybucyjna.....	268
4.2. Vibe composing.....	295
Zakończenie.....	302
Bibliografia:.....	305

Wstęp

Memetyczna idea, że muzyka rozprzestrzenia się jak wirus była szczególnie inspirująca w szczycie pandemii COVID-19, kiedy rozpoczynałem studia doktoranckie. Temat replikacji i mutacji wirusa dominował w publicznym dyskursie - w tym ciekawie przejawiał się w internetowych memach. Niepokojąca muzyka pomagała w rozprzestrzenianiu się materiałów audiowizualnych prezentujących teorie spiskowe na TikToku (Rovic.keeve2001, 2023). Jeszcze w 2020 roku badaczki zaproponowały, że w pełnym modelu pandemii należy uwzględniać także *infopatogeny*, czyli zaraźliwe informacje kulturowe, które rozprzestrzeniają się w sieci i w znacznym stopniu wpływają na losy samego wirusa (Magarey, 2020). Pandemia miała też niebagatelny wpływ na tryb i tok moich studiów: praktycznie wszystkie zajęcia odbywałem zdalnie, a zamiast chodzić do biblioteki, wszelkich źródeł szukałem w internecie.

Poza tym obszar na pograniczu muzyki, memów i sztucznej inteligencji ma dla mnie duże znaczenie. To temat mojej pracy badawczej, ale także praktyki artystycznej, pasji oraz pracy zarobkowej. Od pięciu lat prowadzę media społecznościowe Polskiego Radia Chopin i na bieżąco obserwuję oraz zabiegam o skuteczne rozprzestrzenianie się jego muzycznych memów w internecie. To wzbogacające doświadczenie uwrażliwiło mnie na charakterystyczne cechy skutecznych memów w tej specyficznej niszy kulturowej. Widzę, jak mocno nośność muzycznych utworów i idei jest modulowana przez czynniki pozornie nieistotne lub niezwiązane z głównym tematem. Obserwuję, jak implementacja na bieżąco algorytmów sztucznej inteligencji wpływa na rozszerzanie i blokowanie zasięgów konkretnych treści. Dostrzegam też wzrost aktywności kont jawnie manipulujących opinią publiczną w szczególnie istotnych społecznie okolicznościach.

W dużej mierze odrębnym polem zainteresowania jest dla mnie muzyczna działalność artystyczna, w której także memy i sztuczna inteligencja towarzyszą mi od lat. Jeszcze czym innym jest nie tylko przeglądanie, ale też udostępnianie memów w ramach pasji. Od 2013 roku w ramach prowadzonych stron udostępniłem na Facebooku tysiące

memów internetowych, przy okazji przeglądając ich znacznie więcej i wyszukując, śledząc te szczególnie interesujące nisze. To właśnie z tego wieloletniego zainteresowania kulturą memów internetowych wyniknęło późniejsze zainteresowanie memową teorią, a także powiązanie tej sfery z muzykologią.

Niniejsza praca jest poświęcona zagadnieniom ewolucji kulturowej muzyki. Szczególną wagę przywiązałem do opisu aspektów związanych z technologiami cyfrowymi, w tym najważniejsze miejsce zajmuje sztuczna inteligencja. Za to główny teoretyczny punkt wyjścia stanowi memetyka. Te pozornie rozbieżne tematy splatają się w jedną spójną narrację. Jedną z tez pracy jest to, że memetyczna ewolucja muzyki doprowadziła do powstania muzycznej sztucznej inteligencji, która już teraz w postaci systemów rekomendacji w sposób nienadzorowany kieruje globalnym systemem konsumpcji muzyki. Ponadto, zgodnie z własną syntezą programu badawczego memetyki muzyki stawiam tezę, że muzyczne memy, czy też produkty, warianty muzycznej kultury ewoluują w kierunku rosnącej sprawczości, autonomii w oddziaływaniu na ludzi i szersze środowisko biokulturowe w taki sposób, aby coraz skuteczniej prowokować wytwarzanie własnych kopii. To znaczy, funkcjonalnie, zapośredniczona w technologii kultura muzyczna wykazuje się cechami coraz bardziej przypominającymi intencjonalne zabieganie o własny byt.

Memetyka, która umożliwia przeprowadzenie takiego rozumowania, od czasu przełomu wieków uchodzi za dziedzinę martwą. Żeby zatem móc racjonalnie się nią posłużyć, musiałem zidentyfikować i odeprzeć ciężące na niej zarzuty. Uważam, że wykorzystanie paradygmatu memetycznego do interpretacji zjawisk kulturowych zwłaszcza w domenie rzeczywistości cyfrowej jest jedną ze szczególnie przydatnych strategii eksplanacyjnych. Rosyjskie propagandowe internetowe memy zalewają nie tylko polski internet i są istotnym elementem wojny hybrydowej. Przywrócenie regularnej, racjonalnej praktyki badawczej poświęconej tym zagadnieniom jest zatem nie tylko jednym z możliwych sposobów interpretacji zjawisk zachodzących w środowisku kulturowym ludzi, ale także realnym narzędziem utrzymywania bezpieczeństwa

narodowego. Skoro w 2006 roku na United States Marine Corps School of Advanced Warfighting na Marine Corps University przyjęta i obroniona została praca magisterska poświęcona memetyce w amerykańskich operacjach wojennych (Prosser, 2006), to prawdopodobnie także w polskich realiach 19 lat później podejście memetyczne może okazać się równie istotne.

Przechodząc od memetyki ogólnej do memetyki muzyki, proponowana refleksja również może pomóc w wyjaśnieniu i naświetleniu globalnych problemów w sercu kultury muzycznej. W 2020 roku według Amerykańskiego Stowarzyszenia Przemysłu Nagraniowego (RIAA) dochody z muzycznych streamingów stanowiły 83% dochodów całej globalnej branży muzycznej (Friedlander, 2020). Ta liczba zapewne nie uległa znacznym zmianom do 2023 roku, kiedy po wprowadzeniu nowych zasad monetyzacji (wciąż obowiązujących), aż 86% utworów obecnych na Spotify przyniosło samym artystom dosłownie zerowe zyski (Tencer, 2024). Zarabia zatem sama korporacja, która przy tym poważnie inwestuje w AI (Spotify, 2023) i formalnie raportuje, że jeśli ich (oparty na AI) system muzycznych rekomendacji przestałby działać, dalsze prowadzenie rentownej działalności zostanie zagrożone (Spotify technology, 2024). Taki stojący w centrum wiodącej muzycznej korporacji system rekomendacji idealnie nadaje się do analizy memetycznej z kilku przyczyn. Po pierwsze, powstał w wyniku memetycznej ewolucji i jego działanie jest aktem emulacji tejże memetycznej ewolucji. Po drugie, kształtuje on ewoluujące trajektorie popularności utworów, playlist i gatunków muzycznych. Po trzecie, wywołuje ukierunkowaną ewolucję preferencji muzycznych swoich słuchaczy w takim kierunku, aby konsumowali jak największą liczbę nagrań muzycznych.

Ale w internetowej kulturze muzyka to nie tylko samodzielne muzyczne utwory w dedykowanych aplikacjach do odtwarzania muzyki. Istnieje ona także jako integralna część materiałów multimedialnych, stanowiąc muzyczną warstwę funkcjonującą w różnych kontekstach w mediach społecznościowych, od których coraz trudniej się oderwać. Z mediów społecznościowych w 2024 roku korzystało ponad pięć miliardów ludzi, a dla każdego z nich średni dzienny czas konsumpcji wynosił ponad 2 godziny

(Chaffey, 2025; Kemp, 2024). Każda z tych osób przegląda memy internetowe, ponieważ stanowią one treść każdej platformy społecznościowej. Jeszcze kilka lat temu miały one przede wszystkim formę obrazkową, dziś znacznie częściej są to krótkie filmiki z muzyką i dźwiękiem. Przeglądamy memy, zamiast pracować, uczyć się, wypoczywać. Przeglądamy memy, przechodząc przez jezdnię, prowadząc samochód, opiekując się dzieckiem. W jakim celu to robimy? Nauki ewolucyjne głównego nurtu mówią, że jeżeli w kulturze pojawia i utrzymuje się jakieś zachowanie, to dlatego, że jest ono dla ludzi korzystne i pozytywne (Boyd, 1995). Co zatem takiego wartościowego otrzymujemy w nagrodę za scrollowanie, jeśli jesteśmy gotowi poświęcić w jego imię czas, siły, zdrowie i życie? Kompulsywne przeglądanie memów może dawać pewne korzyści adaptacyjne. Dzięki temu pozostajemy na bieżąco z informacjami o otoczeniu społecznym i kulturowym, co sprzyja orientacji w grupie i ułatwia dostosowanie się do zmieniających się norm. W ten sposób poznajemy bieżące przeboje muzyczne i uczymy się najnowszych zwrotów językowych. Wspólne przeżywanie treści memów może też chwilowo wzmacniać więzi społeczne. Jednak te wszystkie, często ulotne i powierzchowne korzyści biologiczne nie przeważają szali w zestawieniu z kosztami nadmiernej konsumpcji treści. Konsumpcja Tiktoków i YouTube shorts, obecnie często rozpoczynana już w najmłodszych latach życia, degraduje zdolność skupienia uwagi, przyczynia się do rozmaitych problemów psychicznych (Tereshchenko, 2023) i najczęściej jest po prostu stratą czasu i energii. To nie jest użyteczna zabawa, ale uzależnienie (Tereshchenko, 2023; De, 2025). Ten temat jest w kręgu zainteresowań muzykologii, ponieważ w takich wciągających krótkich filmikach muzyka i dźwięk odgrywa integralną, często kluczową rolę.

Memetyka to teoria, według której kultura podlega ewolucji (Dawkins, 1976). Ludzie mają wrodzoną skłonność do imitacji, kopiowania zauważonych zachowań (Merker, 2005), co prowokuje zmienność szczególnie uwidoczną w grze w głuchy telefon. Przy każdym procesie komunikat może nabrać nowych cech – nie wszystko zostaje idealnie przekazane, nie wszystko zostaje idealnie usłyszane. Pomimo szczyrych chęci kopia kopii kopii znacznie różni się od oryginału. W świetle uniwersalnego darwinizmu (Dawkins, 1983) można dostrzec, że kopiowanie na drodze imitacji wywołuje nie tylko

zmienność, ale ewolucję. To dlatego, że spełnione są trzy wymogi procesu ewolucyjnego (Darwin, 1859; Campbell, 1960). Po pierwsze, zachodzi dziedziczność, czyli informacja jest przekazywana dalej. Po drugie, zachodzą zmiany, mutacje, które trwale zmieniają informację podlegającą doborowi. Po trzecie, tylko niektórym kopiom udaje się przetrwać. A zatem kultura składa się z replikatorów - jednostek informacyjnych, które podlegają nieidealnemu procesowi kopiowania. Ludzki ograniczony umysł i pamięć stanowią dla memów główne środowisko replikacji (Dawkins, 1976). Na drodze selekcji te przedmioty kultury, które replikują się gorzej, znikają, a te, które replikują się lepiej, pozostają i mają większe oddziaływanie na przyszły kształt kultury. Poprzez długie i złożone działanie procesu ewolucji w końcu powstają produkty kultury tak skutecznie prowokujące własną replikację, że ich ewolucyjnie nabytą strategię można użytecznie postrzegać jako niemal inteligentną lub intencjonalną (Dennett, 2017).

W tym świetle w problematycznej konsumpcji krótkich filmików nie trzeba doszukiwać się korzyści dla człowieka. Ewolucyjną korzyść odnosi sama kultura – firmy Apple, Samsung, Huawei, a także platformy takie jak TikTok, Youtube, Facebook, Instagram. Nie poszczególni pracownicy, ale firma, jako skomplikowany, autonomiczny system prowokujący własne przetrwanie. W mniejszej skali ewolucyjną korzyść odnoszą także poszczególne filmiki – nie dlatego, że jakoś dynamicznie i celowo o to zabiegały, ale dlatego, że są owocami procesu memetycznej ewolucji, a ich nadrzędną strategią, jest właśnie sprowokowanie własnej replikacji, obecność na maksymalnej liczbie telefonów. Ta strategia, poprzez ewolucyjny design, jest w nich materialnie ucieleśniona. Tak samo jak drzewo nie ma rozumianej na ludzki sposób intencji celowego zabiegania o światło słoneczne, tak samo piosenka nie ma intencji zabiegania o słuchacza. A jednak w obu przypadkach ewolucja doprowadziła do powstania takich systemów, którym udaje się owe strategie realizować. Oczywiście filmy jak i piosenki można rozumieć jako celowe wytwory człowieka. Ale człowiek powstał na drodze ewolucji i zgodnie z memetyką Daniela Dennetta, w szczególności biorąc pod uwagę jego koncepcję Darwinowskich przestrzeni (Dennett, 2017), wszelką twórczość człowieka również można - w pewnym

zakresie - wyjaśnić w kategoriach ewolucyjnych. Memetyka muzyki nie odmawia zatem kompozytorom twórczej woli, ale sytuuje ją w szerszym ewolucyjnym kontekście.

Memy internetowe stanowią doskonały punkt wyjścia do naświetlenia i zrozumienia tego, jak działają memy w ogóle. Choć guzik z kciukiem w górę spopularyzował się dopiero dzięki Facebookowi, sam gest jest znacznie starszy. Udostępnianie, remiksowanie i komentowanie napotkanych treści nie zaczęło się w internecie. Kultura, w tym kultura muzyczna, od zawsze składała się z memów.

Taka perspektywa nie jest jednak popularnie akceptowana w muzykologii. Stan badań należy określić tak, że dyscyplina memetyki muzyki nie jest szeroko rozpowszechniona, a artykuły z tego zakresu regularnie publikuje przede wszystkim jeden badacz - Steven Jan. W 2007 powstała książka *The Memetics of music: neo-darwinian view of musical structure and culture* (Jan, 2007), która wciąż pozostaje najcenniejszym źródłem i jedyną tak szeroką syntezą. Jej autor Steven Jan napisał jeszcze wiele innych artykułów (Jan 2024, 2016, 2015a, 2015b, 2015c, 2012, 2011, 2010, 2004) z dziedziny memetyki muzyki, ale nie pobudziły one szerszego naukowego dyskursu. Jan w 2022 roku wydał również obszerną publikację dotyczącą muzyki i ewolucji (Jan, 2022), która, jak sam twierdzi, miała dopełniać jego wizję memetyki (Jan, 2022). W roku 2022 ukazał się również artykuł Marisy Hoeschele i Tecumseh Fitcha (Hoeschele, 2022), który komentował badania nad ewolucją kulturową muzyki przeprowadzone przez Patricka Savage'a oraz współpracowników (Savage, 2022). Sens tego komentarza był krótki: określił badania Savage'a jako badania z zakresu memetyki muzyki (Hoeschele, 2022). Właśnie na tym polu można szukać badań najbardziej pokrewnych i dla memetyki muzyki bezpośrednio istotnych. W ostatnich dekadach ukazało się wiele badań empirycznych nad ewolucją kulturową muzyki (Ravignani, 2016, Savage, 2017, Anglada-Tort, 2023, Hamilton, 2024). Ci autorzy nie powołują się jednak na memetykę muzyki, co może być działaniem celowym, ponieważ po prostu mogą jej programu badawczego nie akceptować. Tym niemniej w niniejszej pracy postaram się wykazać, że program badawczy memetyki można, a nawet powinno się łączyć z głównym nurtem badań na temat ewolucji kulturowej muzyki. Odróżnia je moim zdaniem przede

wszystkim jedna najważniejsza koncepcja, której opisu nie odnalazłem w zastanej literaturze. Mianowicie, memetyka określa muzykę jako replikator, a zatem strukturę informacyjną, która prowokuje tworzenie własnych kopii. Centralną kategorią analizy muzyki jako memu powinna zatem moim zdaniem być jego sprawczość. Jeżeli muzyka podlega ewolucji memetycznej, to znaczy, że istnieje konkurencja pomiędzy tą muzyką, która dobrze się replikuje i tą, która robi to gorzej. Wygrywa ta, która dzięki jakiejś strategii wywołuje taki efekt w środowisku zewnętrznym, że zostaje skopiowana, a kolejne pokolenia dalej rozwijają tę cechę. Stąd przedstawiam swoją tezę, wokół której osnuta jest narracja dwóch głównych części pracy (2 i 3), że muzyka memetycznie ewoluuje w kierunku rosnącej sprawczości. W zastanej literaturze nie odnalazłem także interpretacji najnowszych technologii muzycznych jako wyniku ewolucji memetycznej. Ten obszar spaja się z wcześniej wymienionym, ponieważ najnowsze systemy sztucznej inteligencji, w tym te o zastosowaniach muzycznych, wykazują wzrastającą autonomię i ograniczają pole decyzyjności i sprawczości człowieka (Morris, 2015; Carnovalini, 2020; Paolizzo, 2020, Endsley, 1995). Staram się przeprowadzić również ewolucyjną narrację historyczną, żeby wykazać, że takie cechy najnowszych muzycznych technologii głęboko wynikają z procesów dziejowych, są z nimi spójne – a może nawet były możliwe do przewidzenia. Jako że niniejsza praca dotyczy owych technologii, należy wspomnieć również o dwóch istotnych publikacjach w tej dziedzinie. W 2023 roku ukazała się książka *Zarażeni dźwiękiem. Rynek muzyczny w czasach sztucznej inteligencji* Stanisława Trzcieskiego (Trzcieski, 2023), a w roku 2024 wydana została praca *Niech się stanie muzyka! O muzycznej generatywności* Marcina Strzeleckiego (Strzelecki, 2024). Obie książki opisują odmienne zagadnienia związane z muzyką i sztuczną inteligencją, co jest wartościową nowością w polskojęzycznej literaturze.

Niniejsza praca dotyczy procesu memetycznej ewolucji muzyki od prehistorii do AI i stanowi próbę wyjaśnienia go w kategoriach naturalistycznych. Podstawowe założenie metodologiczne stanowi przyjęcie perspektywy ewolucyjnej, zgodnie z którą zjawiska muzyczne można rozumieć jako przejawy mutacji, replikacji i selekcji informacji kulturowej. Analiza prowadzona jest zatem w duchu uniwersalnego darwinizmu

(Dawkins, 1983), gdzie procesy kulturowe traktowane są analogicznie do biologicznych, a ich dynamika opisywana jest przez te same algorytmiczne zasady doboru.

Praca ma także charakter muzykologii posthumanistycznej, to znaczy nie opisuje człowieka jako jedyne go twórcę muzyki. Kulturę muzyczną analizuję jako efekt współdziałania wielu aktorów, zarówno ludzkich, jak i pozaludzkich, w sieci rozproszonej sprawczości. W tym sensie przyjęta metodologia wpisuje się w badania nad kulturą w duchu teorii aktora-sieci (Latour, 2005), gdzie kluczowe znaczenie ma śledzenie dynamiki przepływu informacji bez uprzywilejowania podmiotu ludzkiego.

Według Michaela Levina sam proces ewolucji można rozumieć jako proto-poznawczą siłę sprawczą, wytwarzającą struktury zdolne do adaptacji i kumulowania informacji (Levin, 2023). Z kolei Richard Dawkins twierdzi, że ewolucji podlegają zarówno geny, jak i memy (Dawkins, 1976), a zatem zarówno organizmy żywe, jak i kultura. Daniel Dennett natomiast uważa, że człowiek stanowi istotę hybrydyczną złożoną z genów oraz memów (Dennett, 2017). Wszystkie te idee są kluczowe dla rozważań podjętych w niniejszej pracy. Wzorując się na tych propozycjach teoretycznych i aplikując je na grunt badań muzykologicznych, staram się zauważać muzykę i kulturę muzyczną we wszelkich jej przejawach oraz przy tym akcentować, że heterogeniczne osoby, które ją tworzą, wspólnie łączą się w złożoną sieć relacji o rozproszonej sprawczości.

Jednym ze sposobów poszukiwania argumentacji przemawiającej za proponowanymi hipotezami jest krytyczny przegląd literatury. Obejmuje on zarówno teksty muzykologiczne, jak i prace z zakresu teorii ewolucji, memetyki, filozofii umysłu, kognitywistyki oraz sztucznej inteligencji. Celem tego przeglądu nie jest jedynie zebranie istniejących ujęć, lecz ich zestawienie, konfrontacja i reinterpretacja w świetle problematyki ewolucji muzyki. Krytyczny charakter analizy polega na wskazywaniu ograniczeń poszczególnych teorii, ich potencjalnej komplementarności oraz na wyprowadzaniu nowych ujęć poprzez twórczą syntezę. Ze względu na duży poziom ogólności niektórych zagadnień podejmowanych w pracy, jej interdyscyplinary charakter, a także opisywanie najświeższych, dynamicznie zmiennych zjawisk kulturowych,

powołuję się na różne typy źródeł. Przy opracowaniu niektórych zagadnień muszę korzystać z doniesień prasowych i blogów, za to w innych fragmentach powołuję się na recenzowane artykuły naukowe. Nie tworzy to sprzeczności, a dopełnienie.

Niniejsza praca oprócz wstępu i zakończenia składa się z czterech głównych części. Pierwsza z nich rozpoczyna się wprowadzeniem metodologicznym, w którym dodatkowo komentuję ten aspekt niniejszej pracy. W części pierwszej opisuję także skrótowo znane mi dotychczasowe badania muzykologiczne związane z ewolucją. Stanowi to wprowadzenie do wyjaśnienia roli paradygmatu memetycznego w muzykologii. Przyznając zarazem, że status memetyki jako nauki jest kontrowersyjny, dalej w pierwszej części wskazuję, że memetyka jako teoria naukowa istnieje w dużej mierze odrębnie od internetowej kultury memów. Popularnie jako mem rozumiane są treści internetowe z mediów społecznościowych i innych serwisów, co wyodrębniam jako kategorię memów internetowych. Moim zdaniem memy internetowe stanowią świetny przykład memów, a zatem przegląd memów internetowych służy jako wprowadzenie do tego tematu. Po krótkim opisie kultury memów internetowych wyodrębniam także kategorię internetowych memów muzycznych, czyli takich memów internetowych, w których muzyka pełni istotną rolę. Prezentuję listę 5 moim zdaniem najważniejszych muzycznych memów internetowych, a są to: *Dancing Baby*, *Keyboard Cat*, *Darude - Sandstorm*, *Rickroll*, oraz *MLG Airhorn*. Na podstawie tych różnorodnych internetowych memów muzycznych przedstawiam kręte i zawiłe ścieżki ewolucji internetowej kultury muzycznej. Wybrane memy oprócz znacznej popularności w określonym czasie były również wpływowe na szerszą, nie tylko muzyczną i nie tylko internetową kulturę. Muzyka pełni w nich różne, zaskakujące role, ale zawsze ma niezwykłe znaczenie. Ponadto, przykłady memów internetowych demonstrują złożony charakter kultury internetowej oraz dokumentują nieoczywiste role muzyki, która nie sprowadza się tylko do plików dźwiękowych z piosenkami.

W drugiej części pracy wprowadzam najważniejsze hipotezy i pojęcia teoretyczne. Prezentuję fundamentalne założenia memetyki oraz przytaczam moim zdaniem najmocniejszą wobec niej krytykę, na którą następnie odpowiadam. Jako że centralną

tezą całej pracy jest to, że memetyka stanowi odpowiednią teorię do opisu kultury muzycznej, a jednocześnie memetyka ma opinię zdyskredytowanej pseudonauki, przejmując na siebie ciężar wykazania, że najcięższe zarzuty można obalić, a konstruktywną krytykę można wykorzystać do ulepszenia programu badawczego memetyki. Między innymi przedstawiam zarzut, że memy nie istnieją, a zatem memetyka to pusta metafora. Używając odwołań do różnych badań empirycznych oraz modeli teoretycznych staram się przekonać, że ta krytyka jest bezzasadna. Wykazuję istnienie innych replikatorów niż geny oraz zachodzenie w kulturze procesu ewolucji. Innym argumentem przeciwko memetyce jest to, że nie ma ona skutecznych powiązań z innymi dyscyplinami nauki. Ten zarzut odpieram, demonstrując spójność programu badawczego memetyki z modelem komunikacji Claude'a Shannona (Shannon, 1948), zasadą wolnej energii Karla Fristona (Friston, 2010), konstruktorową teorią informacji Davida Deutscha (Deutsch, 2015), teorią złożzeń (Sharma, 2023), koncepcją poznania 4E (Schiavio, 2024), koncepcją informacjożerców (Biedermann, 2006), a także informatyką, a w szczególności dziedzinami AI oraz algorytmiki ewolucyjnej (Neri, 2012). Następnie omawiam bliżej założenia teoretyczne memetyki muzyki, krytycznie analizując 20 tez memetyki muzyki sformułowanych przez Stevena Jana w jego pracy *Memetics of Music: Neo-darwinian view of musical structure and culture* (Jan, 2007). Omawiając memetykę Jana, stawiam ten sam zarzut, co wobec innych memetyków z okresu przełomu wieków, których twórczość zbiorczo nazywam 'pierwszą falą memetyki': próbują oni sformalizować memetykę jako kompletną, dojrzałą naukę. Według mnie memetykę nadal należy traktować albo jako użyteczny program badawczy, który można spójnie połączyć z innymi, dojrzałszymi naukami, albo podchodzić do memetyki jako oddzielnej nauki, która jest niedojrzała i nie ustaliła jeszcze zamkniętego paradygmatu w rozumieniu Kuhna (Kuhn, 1962), na którym można by oprzeć memetyczną naukę normalną.

W mojej krytyce memetyki muzyki Jana sygnalizuję już role trzech koncepcji naukowych, których integrację z memetyką muzyki na przykładach rozwinę dalej w części trzeciej. Są to teoria aktora-sieci (Latour, 2005), wieloskalowa architektura kompetencji Levina (Levin, 2022) oraz model systemu wiabilnego (Beer, 1972).

Według mnie program naukowy memetyki może być wartościowym poznawczo narzędziem dla muzykologii. Przede wszystkim, za Dawkinsem (1976), proponuję przyjęcie memetycznej perspektywy, która pozwala opisać znane zjawiska przy użyciu nowych formalizmów. Poza tym uważam, że zjawiska dzisiejszej kultury internetowej oraz w szczególności systemy AI działające na polu muzyki można w pewnych aspektach wyjaśnić wyłącznie dzięki perspektywie memetycznej.

Trzecia część pracy zawiera opis tej wyżej wspomnianej perspektywy memetycznej w odniesieniu do kultury muzycznej. Rozpaczynam od przedstawienia replikatorowej, muzycznej perspektywy na początki muzyki. Przytaczam empiryczne badania, które pokazują powstawanie replikatorów w warunkach laboratoryjnych (Aguera y Arcas, 2024). Omawiam również zjawisko earworms – zaraźliwych melodii, które utrzymują się w umyśle bez intencji ludzkiego nosiciela (Jakubowski, 2017). Następnie przedstawiam koncepcję tego, jak powstałe w skomplikowanej kulturze muzyczne instytucje działają jako mempleksy (Jan, 2007). Odwołuję się w tej części do trzech wspomnianych wcześniej koncepcji. Model systemu wiabilnego (Beer, 1972) pozwala opisać uniwersalne systemowe mechanizmy sprawnie działającej instytucji, takie jak mechanizmy planowania, nadzoru czy samoregulacji, podzielone na dalsze podsystemy. Teoria aktora sieci (Latour, 2005) pozwala dostrzec rozproszoną sprawczość w sieciach ludzi oraz obiektów takich jak instrumenty, partytury czy teorie muzyczne. Wreszcie wieloskalowa architektura kompetencji (Levin, 2023) zapewnia podstawę do opisu powstałych na drodze ewolucji systemów złożonych z zagnieżdżonych struktur, gdzie rozproszona sprawczość staje się kluczem do wyjaśnienia inteligentnego funkcjonowania mempleksów takich jak systemy rekomendacji muzycznej. Dalej przyglądam się systemom rekomendacji ogólnie, oraz analizuję platformy muzyczne: Spotify, Youtube i TikTok. Choć YouTube i TikTok są serwisami z różnorodną zawartością audiowideo, wykazuję, że materiały muzyczne są dla obu z nich kluczowe. Omawiam także skróconą historię różnych technologii sztucznej inteligencji do zastosowań muzycznych. Są to w mojej koncepcji najbardziej wyraziste przykłady tego, że memetyczna ewolucja muzyki prowadzi do wzrostu sprawczości memów muzycznych. Na koniec tej części omawiam tzw. agencjalne

systemy sztucznej inteligencji (Russel, 2023), które już obecne wykazują się znaczną autonomią działania i bardzo dynamicznie się rozwijają.

W ostatniej, czwartej części prezentuję efekty testów działania współczesnej AI w zakresie muzyki, które mają na celu zobrazowanie stanu zaawansowania tych technologii. Na tych empirycznych przykładach, które stanowią swojego rodzaju zapis dziennika z badań terenowych, nie opieram ciężaru dowodzenia w niniejszej dysertacji. Prezentuję możliwości badania przy pomocy technik AI diachronicznej zmienności pola semantycznego słowa muzyka, oraz obecne kompetencje popularnych modeli AI do tworzenia aplikacji muzycznych.

Podziękowania

Memetyką muzyki po raz pierwszy zainteresowałem się w 2018 roku, kiedy na studiach magisterskich napisałem pod kierunkiem Prof. Justyny Humięckiej-Jakubowskiej pracę zaliczeniową poświęconą tej tematyce. Wtedy też przeczytałem książkę *The Memetics Of Music: Neo-Darwinian view of musical Structure and Culture* (Jan, 2007), która ukształtowała moje myślenie na ten temat. Już w 2018 wpadłem na pomysł sformułowania krytyki 20 tez memetyki Stevena Jana wymienionych w tejże książce, a prawdziwego przeprowadzenia tej krytyki dokonałem dopiero w niniejszej pracy. Składam zatem gorące podziękowania w kierunku Pani Profesor za pomoc w pierwszych krokach na tym polu, które później stało się przedmiotem mojej dysertacji doktorskiej.

Wielkie podziękowania składam także dla mojego promotora prof. Piotra Podlipniaka, który kształtował i inspirował moje badania związane z ewolucją i muzyką przez ostatnie pięć lat. Ekspercka wiedza Profesora w znacznej mierze kształtowała moje horyzonty oraz warsztat badawczy. Mam szczerą nadzieję i głębokie przecucie, że to jeszcze nie koniec i będę miał szansę pod kierunkiem Profesora jeszcze się wiele nauczyć i zebrać

wiele ciekawych doświadczeń. Pomimo mojego tragicznego bezładu, chaotycznego trybu pracy i długotrwałych braków postępów zawsze odczuwałem ze strony Profesora gotowość do pomocy, profesjonalne wsparcie, otrzymywałem naukowe i praktyczne porady i wskazówki. Najważniejsza była możliwość regularnego uczenia się od profesjonalnego, czynnego, cenionego w światowym środowisku naukowca, który zawsze z pełną życzliwością i otwartością dzielił się swoim cennym warszatem.

Podziękowania należą się także mojej rodzinie, która zawsze motywowała mnie do pracy i z uzasadnioną obawą pytała, jak mi idzie. Tego dalej nie wiem, ale ta praca wygląda na ukończoną. W szczególności dziękuję mojej żonie, Mai, ponieważ moje pisanie kosztowało ją dużo wyrzeczeń. Oczywiście dziękuję także Rozalci za szczególną inspirację. W trakcie zgłębiania meandrów replikacji memetycznej, sam dokonałem replikacji genetycznej i jestem pod ciągłym wrażeniem jej niebywałej tajemnicy.

Dziękuję również Kamilowi, dzięki którego przepastnej wiedzy w dziedzinie programowania w ogóle mogłem zainteresować się polem AI, ponieważ chętnie tłumaczył mi najciekawsze nowości przynajmniej od czasów debiutu AlphaGo w 2016. Kamil pokazał mi także wykład *If brains are computers, who designs the software?* Daniela Dennetta, który później obejrzałem pewnie jeszcze z 20 razy. To ten wykład sprawił, że zafascynowałem się memetyką teoretyczną i pamiętam dokładnie, że to po jego obejrzeniu postanowiłem, że tematem mojego doktoratu będzie właśnie memetyka muzyki. Mój wtedy jeszcze promotor pracy magisterskiej Prof. Podlipniak był początkowo wobec tego pomysłu zdrowo sceptyczny, ale zaufał mojej intuicji i pozwolił mi na zgłębianie tego tematu, za co jestem bardzo wdzięczny.

Niniejsza dysertacja to efekt przynajmniej 6 lat pracy, choć może na podstawie niektórych fragmentów ciężko w to uwierzyć. To w zasadzie czwarte przeformułowanie tej pracy. Najpierw chciałem napisać pracę głównie rozszerzającą memetyczne koncepcje Stevena Jana z jego książki z 2007, ale po tym jak już zacząłem, w 2022 Jan

opublikował ponad 800-stronicową książkę (Jan, 2022), w której sam rozszerzył swoje koncepcje z taką erudycją, że kontynuowanie mojej pracy było kompletnie pozbawione sensu. Postanowiłem zatem skupić się na tym, czego badacz nie poruszał, w tym głównie na opisie roli sztucznej inteligencji. Pod koniec 2022 roku publicznie ukazał się także ChatGPT, a w roku 2023 stan zaawansowania generatywnych systemów AI zmieniał się tak dynamicznie, że moje wcześniejsze podejście do AI także całkowicie się zdezaktualizowało. Następnie napisałem ponad sto stron notatek, z czego nie wykorzystałem dosłownie nic, ale za to zrozumiałem, że najlepszą główną tezą i zarazem osią spajającą całą moją pracę będzie opis ewolucji memetycznej sprawczości.

Zdaję sobie jednak sprawę z tego, że praca w obecnym kształcie to tylko niepełny i nieidealny wstęp do pełniejszego opracowania tematu. Zakres niniejszej pracy jest bardzo szeroki, ale za to wiele zjawisk jest opisanych na poziomie ogólnym, elementarnym, tylko w takim zakresie, który jest istotny dla przeprowadzenia rozumowania. W związku z tym pomięte jest wiele aspektów memetyki muzyki, w tym na przykład memetyczny opis instrumentów muzycznych. Uważam, że ten szeroki temat zasługuje na osobne opracowanie. Poza tym, w części trzeciej opisuję kluczowe z punktu widzenia prowadzonej narracji momenty tranzycji w historii muzyki. Nie może to być jednak rozumiane jako kompletny model historii muzyki. Już na etapie prehistorii skupiam się na zaraźliwych melodiach, ale nie opisuję szerzej rytualnej roli muzyki oraz jej prawdopodobnych ról na wczesnych etapach formowania społeczeństw przez naszych przodków.

Moje badania prowadzone na użytek memetyki muzyki na polach biologii rozwojowej czy sztucznej inteligencji także są bardzo uproszczone. Zachowałem starań, aby zapoznać się z taką ilością literatury, która pozwoliła mi pewne fundamentalne zagadnienia zrozumieć na tyle, żeby stworzyć użyteczne dla narracji mojej pracy modele.

Pozwolę sobie zadedykować tę pracę potrójnie: po pierwsze Rozalci, której dedykuję wszystko, co robię. Po drugie dla dwóch Profesorów, którzy mocno na mnie wpłynęli i niestety odeszli w trakcie pisania tej pracy. Pierwszy to wspomniany memetyk Daniel Dennett, którego refleksja jest dla mnie ogromną inspiracją i podstawą badawczą dla tej pracy. Po drugie Profesor Rafał Zapała, z którym miałem przyjemność wielokrotnie o muzyce dyskutować. Rafał jako pierwszy przekonał mnie do wartości posthumanizmu, który jest pewnym punktem wyjścia w tej dysertacji. Poza tym pamiętam, że kiedy powiedziałem mu, że będę pisał o memetyce, zapytał „Ale co Ty chcesz jeszcze o memetyce pisać, przecież wszyscy czytaliśmy *Samolubny Gen?*” i nie miałem dobrej odpowiedzi, ponieważ faktycznie w książce *Samolubny Gen* Richard Dawkins opisał wszystko, co najważniejsze, a reszta to już tylko dalsze empiryczne potwierdzenia i poszerzanie szczegółów. Miałem jednak nadzieję, że przyniosę Rafałowi moją ukończoną pracę i przekonam go do tego, że w tym temacie można jeszcze wiele powiedzieć. To niestety mi się nie udało.

Zgodnie z zasadami wykorzystania AI w pisaniu prac dyplomowych na UAM, opiszę tutaj moją praktykę w tym zakresie. Przede wszystkim po pierwsze muszę zaznaczyć, że systemy AI są jednym z głównych przedmiotów badawczych niniejszej pracy i korzystałem z wielu z nich w celu przeprowadzenia badań empirycznych. Taką pracę empiryczną z wykorzystaniem AI opisuję w czwartej części tej rozprawy. Poza tym AI przysłużyła mi się na wielu nieoczywistych polach. Na przykład system rekomendacji YouTube polecił mi wykład, który dr Patrick Savage przeprowadził online w ramach konferencji ESCOM. Na koniec tego wykładu Savage powiedział o tym, że rozpoczyna nowy międzynarodowy projekt badawczy związany z muzyką i ewolucją. Zgłosiłem się i do dziś jestem członkiem zespołu Many Voices, w ramach którego nie tylko wydaliśmy artykuł w piśmie Science Advances i pracujemy nad kolejnymi, ale także zapoznałem się z wieloma muzykologami z całego świata i poszerzyłem swoją wiedzę w zakresie muzyki i ewolucji. Poza tym, szczególnie w zakresie artykułów prasowych dotyczących kultury internetowej oraz najnowszych artykułów technicznych dotyczących AI, korzystałem z AI jako wyszukiwarki tekstów źródłowych. Mam pełną świadomość tego, że nawet najbardziej zaawansowane systemy jeszcze dziś czasami mają halucynacje i prowadzą do nieistniejących tekstów. Dlatego podchodziłem do tego zagadnienia z dużą ostrożnością i sprawdzałem rekomendowane lektury na każdym kroku. Dzięki tej praktyce natrafiłem na dziesiątki bardzo ciekawych artykułów, które często pozwalały mi pozytywnie lub negatywnie zweryfikować moje przypuszczenia.

1.1 Metodologia

Znaczna część niniejszej pracy jest poświęcona refleksji nad ogólną teorią ewolucji muzyki. Ponadto, kluczowe znaczenie mają zagadnienia związane z szeroko rozumianą ewolucją oraz wybranymi aspektami życia jako dynamicznego procesu w sferze materialnej. Tych zagadnień nie sposób zrozumieć bez opierania się na fachowej wiedzy m.in. z biologii ewolucyjnej. Ponadto omawiana ewolucja muzyki zachodzi również w sferze technologii cyfrowych, które to wymuszają stosowanie precyzyjnych pojęć i logiki nauki o informacji.

Można zatem powiedzieć, że niniejsza praca mieści się na styku muzykologii oraz bioinformatyki – lub biomuzykologii i muzykologii obliczeniowej. Stąd można przypisać niniejszą pracę do rodzącego się nowego obszaru biomuzykologii obliczeniowej – choć chcąc zachować jeszcze większą precyzję, taka nazwa szczególnie dobrze charakteryzuje badania (ilościowe eksperymenty empiryczne) opisane w czwartej części rozprawy. Zagadnienia teoretyczne przedstawione w trzech wcześniejszych częściach pracy można natomiast określić mianem filozofii biomuzykologii obliczeniowej.

Interdyscyplinarność tej pracy polega na łączeniu metod i zakresów badawczych z następujących dyscyplin czy programów badawczych: memetyka (Dennett, 2017), biologia ewolucyjna (Dawkins, 1976), biologia kognitywna (Kováč, 2000), metabiologia (Chaitin, 2012), psychologia ewolucyjna (Kurzban, 2005), neuronauka (Duch, 2021), kognitywistyka (Friston, 2010), informatyka (Shannon, 1948), kulturoznawstwo (Watkins, 2005), software studies (Manning, 2012), nauka o systemach (Beer, 1984), cybernetyka (Mazur, 1966), a także sztuczna inteligencja ze szczególnym uwzględnieniem takich dziedzin jak uczenie maszynowe (Whiteson, 2006), przetwarzanie języka naturalnego (Hamilton, 2016), systemy rekomendacyjne (Ricci, 2022), sztuczne sieci neuronowe (Eck, 2002).

Odpowiednia dla niniejszej pracy jest także post-popperowska koncepcja programów badawczych Imre Lakatosa (Lakatos, 1978). Opisywane pojęcia i tezy wspólnie pracują na bioinformuzykologiczny program badawczy, a jednym z jego produktów i narzędzi jest paradygmat memetyczny. Oprócz generalnej zasady podążania od zagadnień najbardziej ogólnych do problemów szczegółowych, w trakcie pracy pojawiają się również problemy ‘poboczne’, ale związane z szeroko zakrojoną tematyką pracy.

Ostatecznie najpoważniejszym wyzwaniem metodologicznym jest powiązanie nauki humanistycznej, jaką jest muzykologia z biologią ewolucyjną, czyli nauką przyrodniczą, w jeden spójny program badawczy. W uproszczeniu przyjmuję jednak, że to nauki przyrodnicze rządzą się bardziej rygorystycznymi wymogami – jak empiryczna falsyfikowalność – czego nie da się powiedzieć o filozofującej humanistyce.

Daniel Dennett, memetyk, którego refleksja była największą inspiracją do powstania tej pracy, działał nieustannie na granicy nauk, czy też może na granicy nauki i filozofii. Sam jednak widział się bardziej jako filozofa niż naukowca i, jak mówił, szukał raczej ciekawych pytań niż odpowiedzi. Takie pytania następnie trafiały do inżynierów, programistów, informatyków zajmujących się praktycznymi problemami, którzy wyszukiwali techniczne rozwiązania dla ciekawie i twórczo postawionych filozoficznych problemów. Taka refleksja ma ciekawe implikacje dla metodologii nauk. Zgodnie z propozycją taksonomii nauk podziałem Rickerta (Staiti, 2021) nauki humanistyczne i przyrodnicze nie różnią się przedmiotem, a właśnie metodologią ich uprawiania.

Humanistyka, zdominowana przez metody jakościowe, może być tutaj widziana jako pewien „pierwszy front” nauki: wyszukiwanie ciekawych pytań i proponowanie szerokich hipotez uogólniających, a także projektowanie eksperymentów. Następnie w tak wstępnie przygotowany teren badawczy można wkroczyć z badaniami ilościowymi, obliczeniowymi, falsyfikowalnymi zgodnie z metodami nauk przyrodniczych. W ten sposób uzyskujemy pełny obraz i skuteczną metodę poszerzania wiedzy naukowej w dotąd niezbadanych obszarach (Grobler, 2006).

Przy doborze części materiału badawczego, tj. memów internetowych, trudno opierać się wyłącznie na źródłach akademickich. Najnowsze treści kultury nie zostały jeszcze dostatecznie systematycznie opracowane i same dane na temat ich rozprzestrzeniania się i recepcji trzeba z konieczności oprzeć na doniesieniach prasowych, artykułach internetowych itd. Tym niemniej z tego typu artykułów nie zostaną wyciągnięte żadne nieuprawnione wnioski istotne do weryfikacji hipotez. Natomiast owe źródła są sprawdzane krytycznie.

Problemem narzucającym się od razu przy takiej mnogości dyscyplin jest rozumiana w sensie filozofii nauki Thomasa Kuhna (Kuhn, 1962) niewspółmierność pojęć. Z uwagi na to, że różne nauki w ramach własnych paradygmatów mogą przyjmować rozbieżne znaczenia dla niektórych pojęć, zwracam szczególną uwagę na precyzyjne ich zdefiniowanie.

1.2 Dodatkowe definicje

Na potrzeby niniejszej rozprawy muszę ustalić robocze definicje dla kilku kluczowych pojęć. Rozprawa dotyczy ewolucji składających się z informacji kulturowych memów muzycznych ku wzrastającej sprawczości. Najważniejsze pojęcia rozumiem w sposób następujący:

Informacja to rozróżnialny stan fizyczny, który może być określony jako redukcja niepewności w sygnałach. Z perspektywy teorii konstruktorów informacja posiada funkcjonalno-kontrfaktyczną sprawczość, czyli wyznacza, jakie zmiany są wykonalne i powtarzalne na jej nośnikach (Marletto, 2015). W systemach ewoluujących (biologicznych lub kulturowych) informacja jest tym, co podlega kopiowaniu ze zmiennością i selekcją, tj. jest to dziedziczny wzorzec, który utrzymuje się, ponieważ zwiększa żywotność lub wydajność systemu (Shannon, 1948; Lewontin, 1970; Kolchinsky, 2018; Adami, 2012).

Informacja kulturowa to informacja nabywana społecznie (Jablonka, 2005). Rozpowszechnia się ona w populacji poprzez imitację oraz inne formy uczenia się, a nawet wymianę substancji chemicznych (Jablonka, 2005). Obejmuje reprezentacje umysłowe oraz ich zewnętrzne przejawy, takie jak umiejętności, normy, symbole, teksty, artefakty czy instytucje. Informacja kulturowa w sposób przyczynowy wpływa na zachowanie. Utrzymuje się ona dzięki transmisji społecznej, dzięki czemu może się zmieniać i ewoluować w czasie (Boyd, 2009; Mesoudi, 2018; Richerson, 2005; Whiten, 2011).

Ewolucja to fundamentalny proces, którego istotą jest zmiana częstości występowania cech w populacjach. Zachodzi on w każdym systemie, w którym spełnione są trzy warunki konieczne: istnieje różnorodność, selekcja oraz dziedziczenie (Lewontin, 1970). Warunki te są zarazem czynnikami, które w cyklicznej interakcji tworzą mechanizm

napędzający ewolucję. Innymi słowy muszą istnieć replikatory, czyli takie struktury informacyjne, które prowokują powstawanie własnych kopii. Zbiór replikatorów musi zawierać różne ich warianty, czyli muszą podlegać one dziedzicznej zmienności. Co więcej, te replikatory muszą (z wewnętrznych lub zewnętrznych przyczyn) wykazywać się różnym sukcesem w tworzeniu kopii. Niektóre tworzą ich więcej, inne mniej, a jeszcze innym w ogóle się to nie udaje i znikają z populacji. Tym samym ewolucja jest uniwersalnym algorytmem odkrywczym, który może zachodzić w dowolnym medium spełniającym powyższe warunki. W jej ramach może zachodzić zarówno dobór naturalny jak i dobór sztuczny, które stanowią ekstrema całego spektrum *przestrzeni darwinowskiej*. Ewolucja zachodzi zarówno w systemach biologicznych jak i kulturowych.

Mem to replikator przekazywany kulturowo, czyli jednostka dziedzicznej informacji kulturowej, która poprzez swoją strukturę wpływa na zewnętrzne środowisko w taki sposób, że zostaje kopiowana. To kopiowanie, wraz ze zmiennością, zachodzi głównie, kiedy jednostki imitują nawzajem swoje zachowanie (Merker, 2005). Memy istnieją zarówno w ludzkich umysłach jak i w przedmiotach kultury. Rozprzestrzeniają się ze zmiennością, dziedzicznością i zróżnicowanym sukcesem, dlatego mogą być podstawą kumulatywnej ewolucji kulturowej. W minimalnym, darwinowskim sensie, memy wykazują funkcjonalną sprawczość. Warianty, które skuteczniej stwarzają warunki dla własnej replikacji, przetrwają i rozprzestrzeniają się (Dawkins, 1976, 1982; Hull, 1980; Dennett, 1995, 2017; Mesoudi, 2011; Henrich, 2015).

Sprawczość, agencjalność (od łac. *agere* – działać, sprawiać, poruszać) to zdolność systemu ludzkiego lub nie-ludzkiego do wywierania skutków w świecie poprzez inicjowanie lub modyfikowanie zdarzeń. Nie wymaga ona świadomości ani intencjonalności, lecz polega na relacyjnej możliwości generowania różnicy w sieciach przyczynowych i społecznych (Giddens, 1984; Emirbayer, 1998; Latour, 2005; Dennett, 2017).

1.3. Ewolucyjne wątki i badania w muzykologii

Myślenie o muzyce w kategoriach ewolucyjnych nie jest niczym nowym. Wręcz przeciwnie – refleksje na temat ewolucji i muzyki zawierał w swoich dziełach już sam Karol Darwin (1871), a niektóre z jego poglądów wciąż zyskują nowe interpretacje (Frank, 2024) i nie zostały do końca sfalsyfikowane. Bogatą historię stosowania ewolucji jako metafory w muzykologii opisuje Steven Jan (Jan, 2022). Jan bogato wskazuje jak tradycja jeszcze XIX-wieczna stopniowo przerodziła się we współczesną muzykologię ewolucyjną oraz biomuzykologię. Tymczasem dla niniejszej rozprawy najważniejszy punkt odniesienia poza memetyką muzyki stanowi kilka głównych pozycji z tego zakresu. Pierwszym i najważniejszym z nich jest refleksja muzyczna samego Karola Darwina. Uczony pisał na przykład:

Rozmyślając o doborze płciowym, przekonamy się, że człowiek pierwotny, lub raczej pewien wczesny przodek człowieka, prawdopodobnie najpierw używał swojego głosu, aby produkować prawdziwe muzyczne kadencje, jak w śpiewie, tak samo jak czynią dziś niektóre gibony; z szeroko zakrojonej analogii możemy konkludować, że ta zdolność była szczególnie wykorzystywana podczas zalotów, wyrażała różne emocje, takie jak miłość, zazdrość, tryumf i służyła za wyzwanie dla rywali. Jest zatem prawdopodobne, że imitacja muzycznych okrzyków przez precyzyjne dźwięki mogła dać początek słowom wyrażającym różne złożone emocje.

When we treat of sexual selection we shall see that primeval man, or rather some early progenitor of man, probably first used his voice in producing true musical cadences, that is in singing, as do some of the gibbon-apes at the present day; and we may conclude from a widely-spread analogy, that this power would have been especially exerted during the courtship of the sexes,—would have expressed various emotions, such as love, jealousy, triumph,—and would have served as a challenge to rivals. It is, therefore, probable that the imitation of musical cries by articulate sounds may have given rise to words expressive of various complex emotions. - Darwin, 1871

Powyższy cytat demonstruje poglądy Darwina na pochodzenie muzyki oraz jej ewolucyjną funkcję. Jego zdaniem była ona przez naszych dalekich przodków wykorzystywana do wyrażania przeróżnych afektów oraz stanowiła istotne narzędzie przy zabieganiu o partnera do reprodukcji.

Przekonanie Darwina o tym, że jeszcze u przodków człowieka jakaś forma muzyki pełniła ważną funkcję, ma istotne implikacje dla muzycznej ontologii i aksjologii. Dla kontrastu, jeszcze w XX wieku, a nawet dziś, w myśli o muzyce funkcjonuje przekonanie, że muzyka to pewien niekonieczny, ozdobny, rozrywkowy dodatek w życiu człowieka. Najślynniejsza jego formalizacja znana jest jako hipoteza sernika słuchowego Stevena Pinkera (Pinker, 1997; Shintel, 2021). Dla moich późniejszych rozważań nad memetyką warto podkreślić, że Darwin mówi tu już o imitacji muzycznych okrzyków przez precyzyjne dźwięki, co dało początek słowom, zatem przychyła się do hipotezy, że istnienie muzyki poprzedzało istnienie języka, ale przede wszystkim, wskazuje on na kluczową rolę imitacji, która jest fundamentalna dla memetyki. Co więcej, skoro Darwin uważa muzykę i język za ściśle połączone, znaczącą wagę dla niniejszej pracy ma też inny jego pogląd:

Przetrawanie pewnych faworyzowanych słów w walce o byt to dobór naturalny.

The survival of certain favoured words in the struggle for existence is natural selection (Darwin, 1871).

Warto, z intencją zachowania oryginalnego znaczenia, dokonać pewnej szerszej interpretacji znaczenia tych słów. Można dziś rozumieć je tak: z pomocą człowieka niektóre słowa wygrywają walkę o byt z innymi słowami i ten proces ma charakter doboru naturalnego.

To stwierdzenie interpretuję dwojako w kontekście memetyki. Po pierwsze określenie 'favoured words' jednoznacznie wskazuje na rolę sprawczości człowieka w tym procesie, ale jednak mówi o doborze naturalnym zachodzącym wśród słów. Po drugie Darwin używa określenia 'dobór naturalny', a nie 'dobór sztuczny' i należy to uznać za

zabieg jak najbardziej świadomy i celowy. Skoro między słowami nie zachodzi dobór sztuczny, to znaczy, że zdaniem Darwina ludzie nie podejmują świadomych decyzji i działań w celu zapewnienia przetrwania określonym słowom. Proces ten jest więc niejako automatyczny – słowa rozmnażają się i mutują, ewoluują, niezależnie od ludzkiej woli, po prostu niektóre wyewoluowane słowa lepiej trafiają w ludzkie gusta, a inne gorzej, poprzez swoje cechy. Jeśli nie zależą one od ludzkiej woli, to muszą wynikać z połączenia ludzkiego działania niecelowego, a także zależności społecznych, środowiskowych oraz losowych. Ludzie zatem umożliwiają byt i ewolucję słów, ale nie sprawują nad tym żadnej wolicjonalnej kontroli. Podobnie jak drzewa umożliwiają byt i ewolucję człowieka, ale nie sprawują nad nimi wolicjonalnej kontroli.

Takie podejście dzisiaj moglibyśmy niemalże utożsamić z radykalną memetyką, według której rola wolnej woli i kreatywności człowieka jest zerowa, a kulturę kształtują wyłącznie zależności dynamiki procesów memetycznych (np. Blackmore, 1999). To podejście jednak budzi bardzo poważne wątpliwości. Badacze uważają, że mechanizmy ewolucji kulturowej nie wykluczają ludzkiej kreatywności (Kronenfelder, 2011; Chatterjee, 2022). Poza tym przebieg samej ewolucji kulturowej również opisuje się w bardziej złożony sposób. W informatyce wyabstrahowany algorytm ewolucji memetycznej łączy darwinowski mechanizm selekcji naturalnej z lamarckowskim mechanizmem uczenia, w którym jednostki mogą modyfikować swoją strukturę w trakcie życia, a następnie przekazywać te zmiany dalej (Moscato, 2003). Ponadto, współcześnie badania nad rolą doboru płciowego w ewolucji muzyki sugerują na przykład, że muzyka wyewoluowała jako kosztowny sygnał sprawności, wskazujący na inteligencję, kreatywność i dobre geny potencjalnego partnera, podobnie jak ozdobne cechy u innych gatunków (Miller, 2000). Częściowego empirycznego wsparcia dla tej hipotezy dostarczają np. badania z 2012, w których mężczyźni pozujący do zdjęcia z gitarą zostali ocenieni jako bardziej atrakcyjni niż ci bez gitary (Tiffertet, 2012).

Szczególnie ważne dla historii muzykologii ewolucyjnej były też badania Alana Lomaxa (Lomax, 1968), których bogate opracowanie stworzył Savage (Savage, 2018). Według Savage'a (2018) projekt Cantometrics Alana Lomaxa to jedno z największych

przedsięwzięć badawczych w historii muzykologii. Lomax wraz ze swoim zespołem przez 11 lat gromadził nagrania muzyki tradycyjnej z całego świata i opisywał je precyzyjnie we własnym systemie, analizując około 1800 pieśni z 148 populacji na podstawie 36 cech klasyfikacyjnych (Savage, 2018). Wnioski, które wyciągał, dotyczyły zarówno struktury muzyki, jak i historii migracji oraz kontaktów kulturowych (Lomax, 1968). W zasadzie jego przedsięwzięcie możemy uznać za działanie aktywistyczne, za przejaw wczesnej etnomuzykologii stosowanej. Celem było zmierzenie się ze wciąż dominującym w muzykologii europocentryzmem, co sam Lomax określał jako potrzebę nowej metody komparatystycznej wolnej od wartościowania na rzecz opisu mierzalnych wzorców i regularności (Lomax, 1976; Feld, 1984). Było to bardzo poważne zadanie, gdyż chodziło o to, aby pokazać, że style muzyczne nie są hierarchiczne, lecz wzajemnie powiązane. Lomax postanowił zaproponować własną miarę opisu różnych tradycyjnych kultur muzycznych w taki sposób, aby żadnej z nich szczególnie nie wartościować, a jednocześnie umożliwić porównania.

Taka metodyka badań miała służyć do wykazania, że kultury muzyczne mogą być do siebie mniej lub bardziej podobne pod różnymi względami. W efekcie kantometryka (cantometrics) ujawniała ich pokrewieństwo – w moim rozumieniu również pokrewieństwo memetyczne. Lomax oczywiście nie używał określenia „memetyka”, ale jego myślenie zdaje się w pełni zbieżne ze współczesnymi badaniami nad dyfuzją i ewolucją wzorców kulturowych. Sam Passmore i Patrick Savage rozwinęli w praktyce badania Lomaxa. Badacze wykorzystali w tym celu cyfrową bazę The Global Jukebox, która zawiera niemal 6 tysięcy pieśni z ponad 1000 społeczeństw. Owe dane posłużyły im do tego, aby połączyć ilościowe i jakościowe analizy muzycznej różnorodności. W swoim artykule wprowadzili oni pojęcie „unusualness”, czyli nietypowości pieśni względem regionalnych wzorców opisanych w ramach kantometryki (Passmore, 2023). Zatem jedną z nowości względem badań Lomaxa i współpracowników jest to, że autorzy skupiają się na różnicach, a nie podobieństwach. Identyfikują w ten sposób np. pieśni z towarzyszeniem fletni Pana (panpipe songs) z Kurska jako największe zaburzenie wcześniejszych ogólnych wzorców. To znaczy, jest to najbardziej nietypowe lokalne muzyczne zjawisko względem lokalnego wzorca spośród wszystkich badanych

kultur świata. Metodyka Passmore'a i Savage'a pozwala testować hipotezy Lomaxa o związku stylu muzycznego z organizacją społeczną, a zarazem wskazuje na znaczenie innych czynników. Badacze dostrzegają i parametryzują np. izolację kulturową czy religię jako czynniki kształtowania repertuarów (Passmore, 2023). Najważniejszy wniosek z badań autorów jest taki, że najlepszym predyktorem nietypowości pieśni jest jej przynależność do konkretnego społeczeństwa, co statystycznie potwierdza istnienie „stylu muzycznego” jako cechy grupowej (Passmore, 2023). W swoim artykule omawiającym muzykologię ewolucyjną Savage (Savage, 2019) przytacza zaskakujący na dzisiejsze standardy cytat z klasycznego muzykologa Curta Sachsa:

(...)nagą prawdą jest, że przyśpiewki Pigmejów i Pigmoidów są nieskończenie bliższe pierwotnej muzyce niż symfoniom Beethovena i pieśniom Schuberta... jedyna dopuszczalna robocza hipoteza jest taka, że najwcześniejszą muzykę znajdziemy u najbardziej prymitywnych ludów.

(...)the plain truth that the singsong of Pygmies and Pygmoids stands infinitely closer to the beginnings of music than Beethoven's symphonies and Schubert's lieder...the only working hypothesis admissible is that the earliest music must be found among the most primitive peoples (Sachs, 1943)

To bardzo problematyczna wypowiedź. Na pierwszy rzut oka widać w niej pewne europocentryczne i rasistowskie wartościowanie - muzyka pigmejów jest nieskończenie bliższa początkom muzyki niż pieśni Schuberta. Co ma oznaczać owa nieskończoność? Czy chodzi o to, że według miary klasycznych teoretycznomuzycznych metod analizy muzyka Schuberta jest bardziej skomplikowana, a zatem bardziej wartościowa niż muzyka Pigmejów? Takie błędne myślenie może być konsekwencją mylnego pojmowania ewolucji. Rasistowskie jest także założenie, że cała muzyka Pigmejów w ogóle może zostać porównana do jednego gatunku jednego europejskiego kompozytora.

Z drugiej strony stawiamy tu bardzo ważny dla memetyki problem – jak duże mogą być różnice tempa memetycznej ewolucji? Sachs stawia roboczą hipotezę, że najwcześniejszą muzykę znaleźć można u najbardziej „prymitywnych“ ludów. Czy jest

możliwa sytuacja, aby kultura „stała w miejscu”? Paradoksalnie, aby uzyskać długą trwałość czy pozorną niezmiennosc memów należy wynaleźć system zapisu tychże memów w zewnętrznym nośniku w sposób skwantyzowany. Na przykład zapis nutowy, rolka z pianoli czy format mp3. Tego typu materialny zapis jest narażony na zniszczenie z uwagi na czynniki środowiskowe, tzn. materialny rozkład nośnika zapisu. Tym niemniej wiemy, że partytury nutowe mogą przetrwać w niezmienionym kształcie nawet kilkaset lat.

Co innego dotyczy memów w ich najbardziej elementarnej formie – czyli pewnej zorganizowanej aktywności neuronalnej żywego organizmu. Memy zapisane w ludzkim umyśle ulegają nieustannym zmianom (Dennett, 2017) i każde ich przekazanie z definicji wiąże się z transformacją, ewolucją. Mając to na względzie, trudno wyobrazić sobie, żeby oralnie przekazywana kultura nie podlegała zmianom.

Tym niemniej z ewolucji genetycznej znamy przykłady organizmów, które zdają się nie podlegać ewolucji przez więcej niż setki tysięcy lat (Kenny, 2016). W pewnym okresie osiągnęły tak doskonałe formy w swoim środowisku, że jeśli tylko to środowisko nie ulegało znacznym zmianom, były faworyzowane przez dobór. Można powiedzieć, że w tym wypadku presja spowodowana konkurencją była bardzo niewielka.

Historyk Jerzy Topolski wielokrotnie podkreślał, że wszelkie rekonstrukcje przeszłości obarczone są wysokim stopniem niepewności. Wskazywał, że historycy często uzupełniają luki w materiale źródłowym własnymi domysłami, które mogą prowadzić do tworzenia narracji bardziej wyobrażonej niż empirycznie uzasadnionej (Topolski, 1983). Jego zdaniem badacz nie dysponuje bezpośrednim dostępem do faktów historycznych, a jedynie śladami i poszlakami, które podlegają interpretacji (Topolski, 1983). Dlatego wyciąganie wniosków o odległej przeszłości, zwłaszcza w przypadku braku źródeł, zawsze grozi projekcją współczesnych wyobrażeń i schematów myślowych na dawne epoki (Topolski, 1998).

Savage twierdzi, że wcześnie muzykolodzy komparatywni inspirowali się źle rozumianą ewolucją (Savage, 2019). Myślano, że ewolucja to bieg od stanów mniej skomplikowanych do bardziej skomplikowanych, od gorszych do lepszych. Koncepcja Darwinowska mówi jednak, że ewolucja to nieustanne przystosowywanie się do środowiska, a nie proces nieustannej jakościowej poprawy (Darwin, 1871). Przy tym warto wspomnieć, że niniejsza praca, tak jak współczesne nauki ewolucyjne, całkowicie odcina się od szkodliwego i fałszywego paradygmatu ewolucjonizmu społecznego. Koncepcja ta, związana przede wszystkim z myślą Herberta Spencera, była wykorzystywana do legitymizowania ideologii zachodniej wyższości i uzasadniania opresji słabszych (Hofstadter, 1955; Stocking, 1982). Współcześni badacze kulturowej ewolucji jednoznacznie odrzucają takie polityczne nadużycia teorii ewolucji, wskazując, że były one błędną interpretacją samego Darwina (Savage, 2019).

Ważne znaczenie dla niniejszej pracy ma również koncepcja biomuzykologii. Określenie bio-musicology spopularyzował w 2015 Tecumseh Fitch (2015), jednocześnie proponując pewne wytyczne dla 'przyszłej dyscypliny bio-muzykologii'. Fitch w swoim artykule o 'przyszłej dziedzinie bio-muzykologii' nie wspomina natomiast, że sam niemalże identyczny termin biomuzykologia (biomusicology) stworzył Nils Wallin w 1992 w swojej pracy *Biomusicology: neurophysiological, neuropsychological, and evolutionary perspectives on the origins and purposes of music*. Deklaratywny i regulatywny ton Fitcha sugeruje raczej, że oto wynalazca nowej dziedziny ogłasza zasady jej funkcjonowania. Tymczasem Wallin w 1992 pisał wprost:

Centralną ideą tej książki jest określenie neurofizjologicznych i ewolucyjnych wymagań do powstania i pierwotnego celu muzyki, szkicując przy tym podstawę zjednoczonego bio-socjo-kulturalnego pola teorii muzyki. W mojej opinii taka teoria jest niezbędnym jądrem gałęzi muzykologii, którą nazwałem biomuzykologią. To termin analogiczny do biofizyki, biolingwistyki itd. Jestem przekonany, że w nadchodzących latach biomuzykologia, jako połączenie z genetyką i naukami kognitywnymi, będzie odgrywać coraz ważniejszą rolę w muzykologii.

The central idea of this book is to establish the neurophysiological and evolutionary prerequisites for the origins and the primordial purpose of music, thereby sketching the foundation of a synthetic (unified), bio-socio-cultural field theory of music. In this author's opinion, such a theory is the indispensable core of the branch of musicology that I have called biomusicology, a term analogous to biophysics, biolinguistics, etc. I am convinced that in the years to come biomusicology, as a bridge to genetics and to the cognitive sciences, will play an increasingly important role in musicology. (Wallin, 1992)

A zatem, pomimo tego, że Fitch w swoim artykule opisującym biomuzykologię (2015) nie wspomina o tym ani słowem, za twórcę biomuzykologii w 1992 musimy uznać Wallina (1992). Zarazem także Fitch proponuje wiele ciekawych koncepcji. Fitch opisuje biomuzykologię jako dziedzinę, która przede wszystkim bada muzykalność – czyli zestaw zdolności i predyspozycji niezbędnych do tworzenia i postrzegania muzyki (Fitch, 2015). Muzyka w tej koncepcji jest produktem zmiennym kulturowo i historycznie, ale opartym na stabilnej, uniwersalnej fizjologii i psychologii człowieka (Fitch, 2015). Jego zdaniem biomuzykologia powinna traktować śpiew, bębnienie, synchronizację społeczną i taniec jako podstawowe elementy muzykalności, obecne we wszystkich kulturach i częściowo dzielone ze światem zwierząt (Fitch, 2015). Fitch podkreśla także znaczenie międzygatunkowych badań porównawczych. Mają one dotyczyć homologii i analogii elementów muzykalności wśród innych niż człowiek gatunków zwierząt (Fitch, 2015). Jego zdaniem w odróżnieniu od etnomuzykologii, która koncentruje się na formach i funkcjach społecznych, biomuzykologia powinna skupiać się na podstawowych i uniwersalnych zdolnościach leżących u podstaw tworzenia muzyki (Fitch, 2015). Jej celem powinno być zrozumienie muzykalności z perspektywy mechanizmów biologicznych, ontogenezy, filogenezy i funkcji ewolucyjnych, zgodnie z pluralistycznym modelem wyjaśniania zaproponowanym przez Niko Tinbergena na bazie wcześniejszych koncepcji Ernsta Mayra (Fitch, 2015).

W tym miejscu pojawia się wątek najważniejszy dla badań kulturowej ewolucji muzyki w ramach biomuzykologii. Fitch proponuje, że model Tinbergena należy rozszerzyć o wymiar zmiany kulturowej w czasie historycznym. Ten poziom plasuje się pomiędzy ontogenezą (rozwój jednostki) a filogenezą (ewolucja gatunku). W językoznawstwie ten proces nazywa się glossogenezą (Hurford, 1990). Glossogeneza opisywana w kontekście językoznawczym również przez Fitcha (Fitch, 2008) jest praktycznie tożsama z procesem ewolucji kulturowej. Co więcej, Fitch (2008) wprost powołuje się na Samolubny Gen Dawkinsa (1976), a zatem tworzy pomost między biomuzykologią a memetyką.

1.4. Memy internetowe

Pierwsze określenie treści propagowanych w internecie jako „memy” przypisuje się Mike’owi Godwinowi. 1 października 1994 roku w magazynie Wired w artykule „Meme, Counter-meme” sformułował on także zasadę znaną dziś jako prawo Godwina (Godwin’s Law). Głosi ono, że ‘im dłużej trwa internetowa dyskusja, tym większa szansa, że pojawi się w niej porównanie do Hitlera lub nazistów’ (Godwin, 1994). To sformułowanie w kontekście artykułu pozwala wyciągnąć kilka wniosków. Po pierwsze, dla Godwina naturalnym medium internetu były wypowiedzi językowe i to w nich dostrzegał memy. Po drugie, memy jako idee w dyskusjach to rozumienie bardzo bliskie dawkinsowskiemu memowi, na którego zresztą Godwin się wprost powołuje (Godwin, 1994; Dawkins, 1976). Po trzecie, memy internetowe od samego początku były mocno zakorzenione w rzeczywistości społeczno-politycznej.

Retrospektywnie jako jeden z pierwszych internetowych memów Knobel i Lankshear wskazują emotikony, takie jak :-)) zaproponowany przez Scotta Fahlmana w 1982 roku (Knobel, 2007). Dalszy rozwój technologii cyfrowych ułatwił przesyłanie bardziej złożonych treści. Kolejny kamień milowy w historii memów internetowych, pierwszy mem internetowy ‘z krwi i kości’ stanowi już *Dancing Baby* z 1996 (Knobel, 2007). Można zatem zauważyć ciekawą paralełę, że memy internetowe ‘narodziły się’ w formie tańczącego niemowlaka. Dla tej rozprawy nie bez znaczenia jest też fakt, że w tym pierwszym internetowym memie ważną rolę pełniła muzyka. Fenomen *Dancing Baby* szerzej omawiam w kolejnym rozdziale.

W akademickich środowiskach ewolucjonistów koncepcja memu wzbudza kontrowersje i wciąż nie została szeroko zaakceptowana (Chvaja, 2020). Za to memy internetowe mają się świetnie. Limor Shifman (2014) określiła obecne środowisko kulturowe jako *hipermemetyczne* - jej zdaniem niemal każde znaczące wydarzenie społeczne czy polityczne wywołuje natychmiastowy wysyp memów w internecie (Shifman, 2014). W efekcie współczesne memy internetowe stanowią wszechobecny element komunikacji,

przenikający wiele sfer życia online oraz offline. Przy okazji warto zauważyć, że coraz trudniej jest być offline. Alruthaya i współpracownicy (2021) zauważają, że pokolenie Z nie zna życia bez internetu, a Prensky już w 2001 (Prensky, 2001) zaproponował określenie *cyfrowi tubylcy* na ludzi wychowanych i ukształtowanych w świecie cyfrowych technologii (odwrotnie niż starsze pokolenia, czyli *cyfrowi imigranci*).

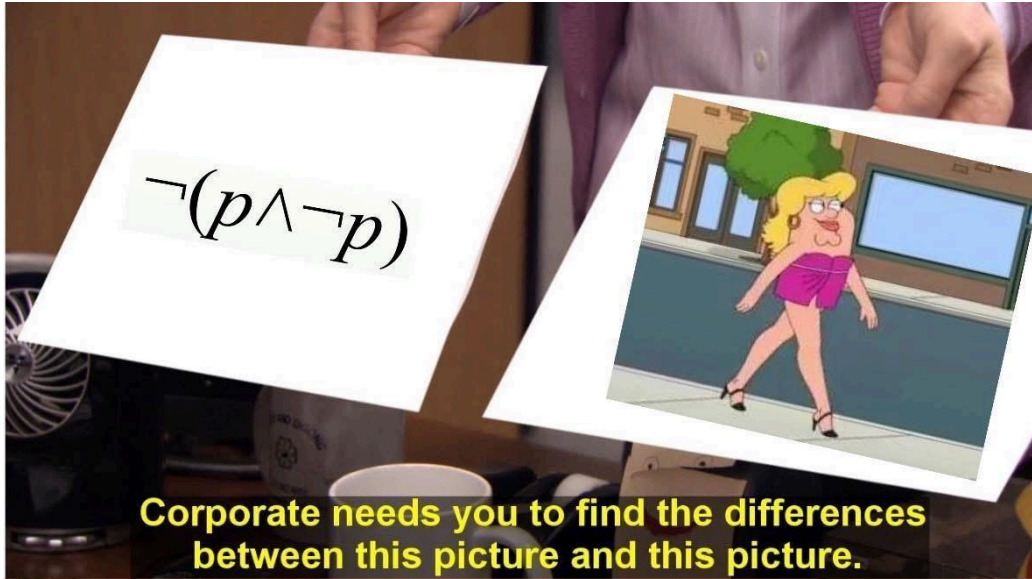
Traktowanie memów internetowych jako przykładów faktycznych memów odrzucili memetycy Richard Dawkins i Daniel Dennett. Dawkins w 2014 roku powiedział, że memy internetowe to „zawłaszczenie” oryginalnej idei, niespójne z jego koncepcją (Dawkins, 2014). Za to Dennett w roku 2017 określił memy internetowe jako ‘kulturowe śmieci’ (Dennett, 2017). Moim zdaniem obaj są w błędzie. Dawkins oburza się, że memy internetowe są projektowane przez internautów, a te ‘prawdziwe’ roznoszą się samodzielnie (Dawkins, 2014). Moim zdaniem nie ma tu żadnej sprzeczności, ponieważ skuteczny internetowy mem to ten, który zyska szerokie zasięgi i zostanie utrwalony, a tego aspektu już żaden twórca nie może do końca przewidzieć. Z Dennettem trudniej racjonalnie polemizować, bo nazwanie memów internetowych kulturowymi śmieciami to po prostu subiektywny sąd wartościujący. Jednak jestem przekonany, że Dennett, jako *cyfrowy imigrant* (Prensky, 2001), po prostu nie trafił na odpowiednią estetyczną niszę.

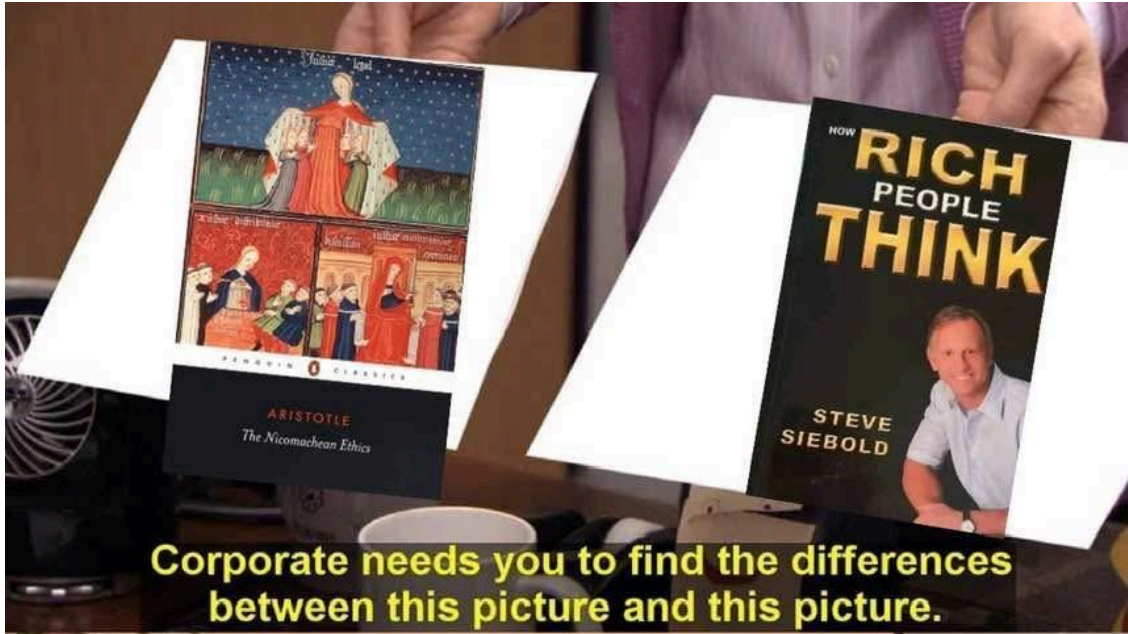
Rozdźwięk między memetyką i refleksją nad memami internetowymi pogłębiła Limor Shifman, tworząc własną definicję memu. Dla Shifman mem nie jest pojedynczą jednostką, ale grupą jednostek o wspólnych cechach. Ponadto, taka grupa jednostek musi posiadać kilka aspektów, które mogą podlegać imitacji. Badaczka szczególnie wyróżnia trzy takie aspekty: treść, formę i nastawienie (Shifman, 2014).

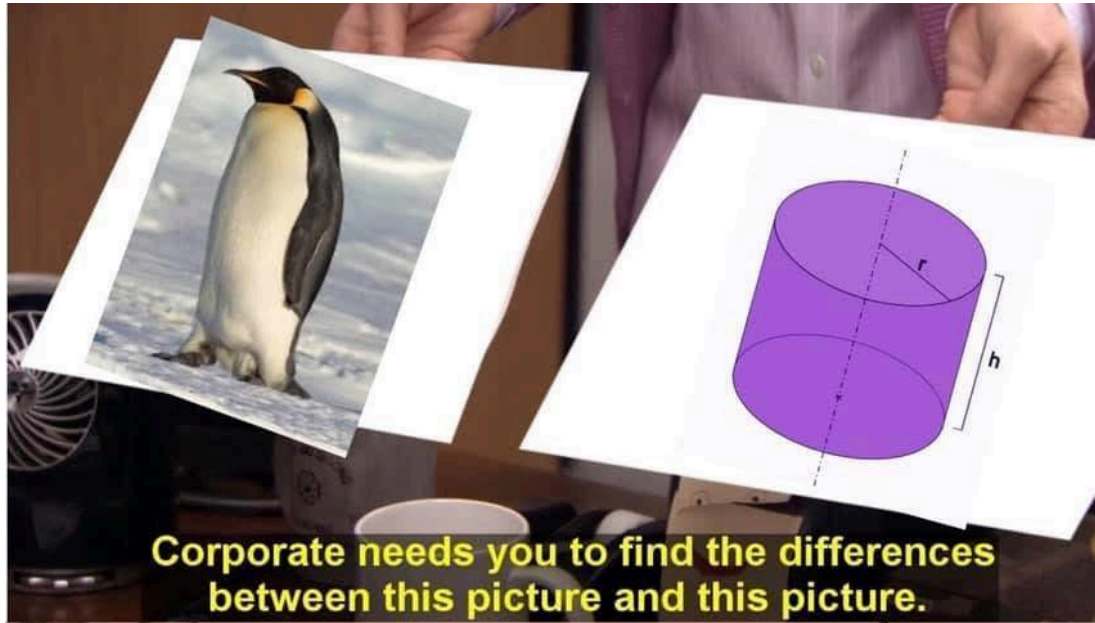
Shifman głównie zajmuje się analizą memów internetowych i jej koncepcja doskonale do nich pasuje. Ale nie tworzy definicji ‘memów internetowych’, tylko memów w ogóle, co prowadzi do pewnych komplikacji. Postrzeganie memu nie jako pojedynczego produktu kultury, ale całej ‘rodziny’ spokrewnionych produktów można potraktować jako rozwiązanie ogólnego problemu rekurencyjności w definiowaniu replikatorów, czyli

takich struktur informacyjnych, które potrafią tworzyć swoje kopie, które potrafią tworzyć swoje kopie, które potrafią tworzyć swoje kopie... a zatem pojedynczy mem jest z definicji powiązany ze wszystkimi swoimi kolejnymi kopiami.

Ale definicja Shifman nie odnosi się tylko do dziedziczenia w łańcuchu kolejnych pokoleń, ale także do memów podobnych, pokrewnych w relacji poziomej, różnych wariantów posiadających wspólnego przodka. W tym sensie dobrze pasuje to do zjawiska wzorców w memach internetowych (*templates*), czyli formuł lub schematów narracyjno-wizualnych, które stają się podstawą dla generowania całych rodzin pokrewnych memów. Shifman dodatkowo tworzy kategorię memów *egalitarnych*, czyli takich, w których żaden jeden wariant poza głównym szablonem nie jest uprzywilejowany (Shifman, 2014). Tego rodzaju szablony memetyczne pełnią rolę „kuzynostwa” kulturowego, w którym powiązania pomiędzy wariantami są równie istotne jak relacje dziedziczenia w łańcuchu kopii. Przykładem może być szablon *They're the same picture* oparty na scenie z serialu *The Office*, gdzie bohaterka ma porównać dwa obrazki i uznać je za identyczne (Gregory, 2021). Zastosowanie tego memu polega przeważnie na wklejeniu ewidentnie różnych obrazków, których postrzegane podobieństwo jest istotą dowcipu:







Corporate needs you to find the differences between this picture and this picture.



Physics textbooks: They're the same picture.



W koncepcji Shifman powyższe przykłady pokazują de facto jeden mem, znany pod nazwą *They're the same picture*. Mają tę samą formę, różną treść i przekazują różne postawy wobec różnych zagadnień. Dokładniej mówiąc, ich formy również są nieco odmienne. W dwóch pierwszych przykładach postać oceniająca obrazki nie została w żaden sposób określona, a zatem źródłem dowcipu jest samo porównanie bliskości treści obrazków, bez odniesienia do żadnego obserwatora. Drugi i trzeci przykład określają obserwatora wprost, a zatem przedmiotem żartu są właśnie ci obserwatorzy i ich błędne postrzeganie relacji wyżej wskazanych kategorii. Można zatem powiedzieć, że pierwszy i drugi tworzą 'rodzeństwo', trzeci i czwarty również tworzą 'rodzeństwo', a pozostałe relacje to o stopień dalsze 'kuzynostwo'.

Shifman jako przykład takiego jednego memu o wielu wersjach wskazuje *Harlem Shake* (Shifman, 2014). To mem szczególnie ciekawy dla muzykologii, ponieważ jego kluczowym elementem jest utwór muzyczny *Harlem Shake*. Jednak cały mem stanowi znacznie szersze zjawisko. Jest to specyficzny rodzaj tańca, który był wykonywany przez kolejne grupy, filmowany i udostępniany w sieci z tą samą muzyką. W lutym 2013 roku zjawisko to osiągnęło niespotykaną wcześniej skalę wiralnego rozprzestrzeniania się. Według Allocca (2013) do 11 lutego opublikowano około 12 000 filmów *Harlem Shake* o łącznej oglądalności przekraczającej 44 miliony, a tempo przyrostu wynosiło około 4000 nowych materiałów wideo dziennie, czyli jeden co 21 sekund. Już dzień później Stern (2013) informował, że liczba odsłon wzrosła jeszcze bardziej, co potwierdzało wyjątkową dynamikę wiralnego rozwoju. W krótkim czasie efekt ten przełożył się na przemysł muzyczny: Trust (2013) odnotował, że 20 lutego 2013 roku utwór *Harlem Shake* zadebiutował na pierwszym miejscu listy Billboard Hot 100 i utrzymał tę pozycję przez pięć kolejnych tygodni. Skalę zjawiska w dłuższej perspektywie podsumowuje Waxman (2014), wskazując, że do dziewiątej rocznicy YouTube'a w 2014 roku w serwisie znajdowało się około 1,5 miliona wersji filmowych opatrzonych etykietą *Harlem Shake*. Zjawisko miało także wymiar infrastrukturalny i prawny: Holpuch (2013) relacjonowała, że system identyfikacji treści *Content ID* przechwytywał tysiące wariantów memu, umożliwiając generowanie przychodów z reklam, a Soha i McDowell (2016) podkreślają, że przypadek *Harlem Shake* stanowił jedno z pierwszych studiów nad monetyzacją i remiksowaniem treści w ramach platformowej infrastruktury YouTube.

W mniejszym lub większym stopniu muzyczny charakter ma bardzo wiele memów, również w formie obrazkowej:



Powyższy mem przedstawia wykres amplitudowo-czasowy słynnej pętli perkusyjnej *Amen Break* z utworu *Amen, Brother* zespołu The Winstons (1969). Uchodzi ona za jeden z najczęściej smploowanych (zapożyczanych) fragmentów muzycznych w historii i stała się fundamentem wielu gatunków muzyki elektronicznej, w tym drum and bass (Schloss, 2014). Jeszcze innym bardzo złożonym zjawiskiem na pograniczu muzyki i internetowych memów jest estetyka *vaporwave*, ale nie podejmuję się opisu tego zjawiska w tej pracy [m.in.](#) dlatego, że zostało już bogato opracowane (Morissey, 2021; Nowak, 2018).

W kolejnym rozdziale zaprezentuję 5 wybranych przykładów takich memów internetowych, w których muzyka pełni szczególnie ważne funkcje. Nazywam je internetowymi memami muzycznymi i mają one na celu zaprezentować to, w jak różnych konfiguracjach może funkcjonować muzyka w internecie.

1.5.1. *Dancing Baby*



Pochodzenie:

Dancing Baby uchodzi za jeden z pierwszych memów internetowych (Börzsei, 2013). To bardzo wczesny wzorzec internetowej treści wiralnej. Dziś trudno wyobrazić sobie taką eksplozję zasięgów i widoczności w świecie bez algorytmów rekomendacyjnych, bez usieciowienia w mediach społecznościowych. A jednak wówczas było to możliwe. Tym samym *Dancing Baby* stanowi także zwrotny moment w ewolucji kultury cyfrowej (Shafer, 2024). Jego pojawienie się i gwałtowny wzrost popularności (Göke, 2024) pokazały rodzącą się potęgę kulturotwórczą internetu. Okazało się, że internauci mogli tworzyć współdzielone treści kulturowe o globalnym zasięgu poza kontrolą tradycyjnych

mediów. *Dancing Baby* w swoich pierwotnych iteracjach zyskał popularność w 1996 roku (Börzsei, 2013), w okresie, gdy możliwości korzystania z internetu były znacznie mniejsze niż w XXI wieku. Ponieważ łącza internetowe były bardzo powolne, sukces memu częściowo wynika ze skutecznej kompresji pliku animacji do lekkiego formatu GIF (Mirapaul, 1997). W tej formie mem rozprzestrzeniał się za pośrednictwem rozwijających się jeszcze kanałów komunikacji, takich jak poczta elektroniczna i fora internetowe (Zannettou, 2018).

Kulturowy przełom komunikacyjny stanowiło to, że ta krótka animacja przyciągnęła uwagę globalnej publiczności jeszcze bez wsparcia ze strony mediów głównego nurtu (Börzsei, 2013; Schafer, 2024). Choć później *Dancing Baby* pojawił się również w telewizji i prasie, jego wczesna popularność była zapowiedzią zmiany w dynamice propagacji treści kultury. Wzrastała rola zdecentralizowanych sieci, które przełamywały monopol mass mediów (Göke, 2024). Tak rodziła się kultura prosumencka, wypełniona treściami generowanymi przez użytkowników internetu (Jenkins, 2006, Lonnberg, 2020).

Treść:

Treść memu *Dancing Baby* zmieniała się w czasie. W pierwszej formie była to niema trójwymiarowa animacja niemowlęcia, powstała pierwotnie jako materiał demonstracyjny firmy Kinetix w 1996 roku. Studio zaprezentowało tak swój sposób przekładania ruchów na komputerowy model postaci (Börzsei, 2013). Początkowo animacja krążyła jako samodzielny plik w środowisku animatorów pod nazwą *sk_baby.max*. Popularność memu wzrosła dopiero wtedy, gdy do animacji została dołączona piosenka *Hooked on a Feeling* zespołu Blue Swede. Ta wersja memu funkcjonuje również pod nazwą *Oogachaka Baby*, gdzie *Oogachaka* to fonetyczny zapis charakterystycznej wokalizacji z początku utworu. Muzyka silnie wpłynęła na odbiór i charakter tego memu (Göke, 2024). Kolejnym etapem było pojawienie się animacji wraz

z muzyką w serialu *Ally McBeal* w 1998 roku, co wyniosło ją do statusu globalnej ikony popkultury (Schafer, 2024).

Dancing Baby, szczególnie w wariacie Oogachaka baby wywoływał ambiwalentne wrażenie. Zaczyna się od niepokojących, nienaturalnych ruchów i mimiki, a następnie tańczące dziecko nadaje charakteru komicznego. Ta mieszanka grozy i humoru zapewniła mu trwałe miejsce w historii memów internetowych (Börzsei, 2013). To wszystko doskonale odzwierciedla struktura towarzyszącej animacji piosenki *Hooked on a Feeling*. Według American Songwriter jej wstęp, oparty na rytmicznej wokalizie „ooga-chaka”, ma niemal plemienny charakter i działa jak zaskakujący sygnał otwierający (American Songwriter, 2022). Z kolei refren, według Stereogum, wnosi ‘chwytną melodyjność i euforyczną energię’, które stanowią kontrast z dziwnym początkiem. To przeciwstawienie i zaskoczenie wzmacniają komiczny wymiar memu (Breihan, 2019). Co ciekawe, jeszcze w 1997 do animacji *Dancing Baby* dołączane były różne utwory muzyczne, w tym *Digging in the Dirt* Petera Gabriela, ale te warianty się nie upowszechniły (Mirapaul, 1997). Skoro nieporównywalnie większą popularność zyskał wariant z jedną konkretną piosenką, możemy uznać, że ma ona decydujące lub co najmniej istotne znaczenie dla całości memu.

Potomstwo:

Obecność *Dancing Baby* w popularnym amerykańskim serialu *Ally McBeal* uchodzi za pierwszy przypadek przeniknięcia memów internetowych do mediów głównego nurtu, a zatem i do globalnej świadomości (Schafer, 2024). W XXI wieku to już standard, że treści powstające w przestrzeni sieciowej kształtują popkulturę na równych prawach z mediami tradycyjnymi (Göke, 2024). Dzięki tej roli *Dancing Baby* stał się już niemal klasycznym punktem odniesienia w badaniach nad kulturą i ewolucją memów internetowych (Göke, 2024).

1.5.2. Rickrolling

Rick Astley-Never Gonna Give You Up



Pochodzenie: Z powodu błędnie zastosowanej autokorekty językowej na stronie internetowej 4chan, słowo 'egg' zamieniało się automatycznie na 'duck'. Stąd wspomnienie dania 'egg roll' zamieniało się na absurdalną frazę 'duck roll', a wspomnienie zwyczaju wielkanocnego egg rolling (Gabbatt, 2015) zamieniało się na abstrakcyjne 'duck rolling'. W 2006 roku przez krótki czas zafunkcjonował mem nazywany *duckrolling* – wizualizacja, ekranizacja tego przejęzyczenia – kaczka na kółkach. Z czasem obrazek jadącej na kółkach kaczki został zastąpiony piosenką *Never gonna give you up* i tak powstał *Rickrolling* (Schonfeld, 2011).

Treść:

Rickroll (lub *Rickrolling*) to jeden z najslawniejszych memów w historii internetu (Baudry, 2022). Właściwie memu nie stanowi zamknięty materiał audiowizualny, ale również konkretne zachowanie. Mianowicie polega on na odtworzeniu, uruchomieniu utworu muzycznego *Never Gonna Give You Up* Ricka Astleya w zaskakującym, zwodniczym kontekście. To być może najważniejszy spośród wielu przykładów memów

internetowych, których charakter, tożsamość, działanie polega na zaskoczeniu oczekiwani.

Główny mechanizm zastosowania tego memu można zobrazować przykładowo:

Osoba A zapewnia osobę B, że przesyła link do dowolnej treści X.

Osoba B wierzy, że otrzymała link do treści X i otwiera ten link.

Osoba B zamiast treści X otworzyła i usłyszała utwór *Never Gonna Give You Up*.

Jest to w zasadzie internetowy *prank*, czyli dowcip, kawał, numer. Rickroll stanowi także jeden z najbardziej rozpoznawalnych przykładów trollingu (Scriven, 2024; Baudry, 2022). W 2025 roku hasło troll internetowy oznacza głównie osoby lub boty wygłaszające w internetowych dyskusjach poglądy uznawane za szkodliwe, prowokacyjne (Paakki, 2021), ale pierwotnie internetowy *trolling* oznaczał po prostu celowe próby wywołania odpowiedzi lub wyprowadzenia rozmówcy z równowagi w ramach żartu (Scriven, 2024). *Rickrolling* stanowi właśnie tę wczesną, niegroźną odmianę trollingu – rozmówca zostaje oszukany, bo nie otrzymał tego, czego oczekiwał. Usłyszał za to radosną piosenkę.

O popularności *Rickrollingu* świadczy chociażby fakt, że utwór na YouTube dobił już do miliarda odsłon. Najczęściej właśnie to do piosenki na tej platformie odnoszą się linki w tym żarcie (Britt, 2019). Co więcej, platforma YouTube w 2008 roku na prima aprilis sama przekierowywała wszystkie filmiki na *Never Gonna Give You Up*, tzn. sama wykonała *Rickrolling* na użytkownikach (Wortham, 2008. Według źródeł dziennikarskich w tym dniu na *Rickroll* nabrało się ponad 6 milionów osób (Van Buskirk, 2008).

Potomstwo:

Rickroll nie jest „samą piosenką”, lecz utrwalonym wzorcem zachowania. Ta internetowa praktyka oszukiwania, zaskakiwania bywa nazywana *bait-and-switch* (Shifman, 2011), co można przetłumaczyć jako ‘przynęta i zamiana’. Ważny jest tu walor społeczny. Audiowizualny klip stanowi pewien nośnik, ale mem definiuje praktyka jego użycia. Badania memów na YouTube sugerują, że memy o szerokich zasięgach charakteryzuje ich podatność na różne transformacje, remiksy (Shifman, 2012; Wiggins & Bowers, 2015). Mem zyskuje, kiedy internauci mogą go kreatywnie zmodyfikować. Ewoluującą historię tego memu, jego swoistą ‘zewnętrzną warstwę’, stanowią również komentarze na Youtube. Film ma 18 milionów kciuków w górę oraz 2.3 miliona komentarzy, z których wiele ma jeszcze swoje kciuki w górę - łącznie wiele milionów, trudnych do dokładnego przeliczenia. Wiele z tych komentarzy opisuje osobiste, pozytywne doświadczenia *Rickrollingu*, co prawdopodobnie wzmacnia jego rozprzestrzenianie się (Siersdorfer, 2010; Yang, 2022).

Jak wskazują Benoit Baudry i Martin Monperrus (2022), *Rickrolling* przeniknął także do literatury akademickiej. Ich systematyczne badanie wykazało, że do marca 2022 roku w bazie artykułów naukowych Google Scholar znajdowały się 33 dokumenty, które zawierały odnośnik do klipu *Never Gonna Give You Up*. 24 z nich stanowiły intencjonalne przypadki *Rickrollingu* - celowe umieszczenie w pracy zwodniczego linku. Najczęściej pojawiały się one w pracach dyplomowych (5 doktoratów, 11 prac magisterskich, 6 licencjackich), a sam fenomen nasilił się wyraźnie po 2017 roku (Baudry, 2017). Istotne znaczenie dla memu *Rickrolling* ma treść i charakter piosenki *Never Gonna Give You Up*. Utwór Ricka Astleya wydany w 1987 roku cechuje się prostą strukturą harmoniczną, zapadającą w pamięć melodią i powtarzalnym refrenem, a zatem spełnia podstawowe kryteria *earworma*, czyli ‘zaraźliwego’ fragmentu muzycznego (Taylor, 2014). Istotna była też pierwotna recepcja utworu. Jego status jako nieco zapomnianego starego przeboju umożliwił jego przewrotne wykorzystanie w żartach internetowych (Phillips, 2015).

1.5.3. *Keyboard Cat*



Pochodzenie: Nagranie, które później stało się memem znanym jako *Keyboard Cat*, zostało zrealizowane w 1984 roku przez Charliego Schmidta. Na filmie kotka Fatso ubrana w niebieską dziecięcą bluzkę zdaje się grać na elektronicznym instrumencie klawiszowym charakterystyczną melodię. Niestety główna gwiazda zmarła w 1987 roku i nie doczekała się swojej sławy (Haynes, 2018). W nagraniu został wykorzystany sampler Ensoniq Mirage. Jest on wyraźnie widoczny w kadrze, a poza tym to na nim Charlie Schmidt skomponował i nagrał muzyczną warstwę memu (Chaudhary, 2010).

Przez ponad dwie dekady nagranie spoczywało w archiwum na kasecie VHS. Charlie Schmidt opublikował swój film *charlie schmidt's 'cool cat'* na serwisie Youtube w 2007 roku, ale jego popularność rozpoczęła się dopiero w roku 2009. Wtedy to katalizatorem popularności stały się filmiki z hasłem 'play him off, keyboard cat' autorstwa Brada O'Farella, w których analogicznie do praktyki z wodewilu, po nieudanym występie następowało 'zgranie' (play off) kogoś ze sceny. W praktyce orkiestra rozpoczynała wykonanie głośnego utworu, który był sygnałem dla wykonawcy, żeby zakończyć występ. Ta praktyka dopełniała aktywną rolę publiczności w wodewilu. Najpierw to publiczność wyrażała dezaprobatę, która stawała się sygnałem dla orkiestry, że czas zagłuszyć niechcianego artystę i w upokarzającej praktyce przegonić go ze sceny (McLean, 1965; Butsch, 2000). Przy okazji warto zaznaczyć, że taka praktyka stanowiła ciekawy oddolny proces selekcji artystów i repertuaru. W swojej analogii do wodewilu O'Farrel dodawał *Keyboard Cata* na koniec znalezionych w internecie materiałów wideo, przedstawiających w zabawnym świetle różne ludzkie pomyłki i niepowodzenia, znane jako *fail videos* (Suddath, 2009; Coyle, 2009; Reeves, 2009). A zatem analogicznie *Keyboard Cat* funkcjonował w pewnej roli selekcjonera, krytyka, który decyduje, których ludzi należy 'przegonić' z internetu. Oczywiście co do zasady decydował o tym autor filmiku, ale w memach widoczny był kot. Tak samo jak ścieżka dźwiękowa była wykonana przez człowieka, ale materiał wideo sugeruje co innego.

Treść: Mem przedstawia kota ubranego w niebieską koszulkę, który pozornie gra na keyboardzie. Tę absurdalną scenę cechuje także specyficzna produkcja. Amatorski charakter wideo z nagrania kamerą VHS odpowiada także amatorskiej jakości dźwięku oraz amatorskiemu sposobowi wykonaniu prostej kompozycji muzycznej. *Keyboard cat* funkcjonuje głównie na YouTube i można wskazać konkretny adres, pod którym jego odsłon jest zdecydowanie największy – ale z nim nie jest jednak tak jak na przykład z memem *Rick Roll*, któremu odpowiada jeden konkretny adres na YouTube. Innymi słowy, *Keyboard Cat* wykazuje odmienną strukturę dystrybucji (Shifman, 2014). Mem ten zdobył sławę dzięki temu, że pojawiał się jako segment w innych materiałach wideo. Zatem jeden główny, najpopularniejszy materiał funkcjonuje jako pewien punkt

odniesienia, punkt powrotu, po tym, kiedy już *Keyboard Cata* zobaczyło i usłyszało się gdzie indziej.

Choć pełen materiał zawiera dwa krótkie muzyczne utwory, tylko pierwszy z nich zdobył rozpoznawalność i jest szeroko kopiowany. Sam ten fakt jest ciekawy dla analizy memetycznej. Pierwotny materiał zawierał wykonania dwóch oryginalnych utworów, jeden po drugim. Ale praktyka 'play him off, keyboard cat' wiąże się tylko z pierwszym z nich (*Keyboard Cat!*, 2010). Zatem memem de facto nie jest oryginalny utwór audiowizualny – tzn. nazwa *Keyboard Cat* nie odnosi się stricte do oryginalnego nagrania z 1984 roku. Ten mem ma bardziej złożoną, nieoczywistą strukturę. Zdobyl on największą popularność w skróconej formie jako fragment dodawany na koniec innych filmów. Ale poprzez to sam oryginalny film również zdobył dodatkowe odsłony. Pod wpływem sukcesu późniejszych iteracji oryginalny film zmienił nazwę – twórca pierwotnie nazwał swój film *cool cat*, ale po interwencji Brada O'Farella przyjęło się *keyboard cat* i tak też została przemianowana nazwa oryginalnego materiału. Z punktu widzenia teorii memetyki ta dynamika stanowi ewolucję memu pod wpływem presji selekcyjnych.

Podstawowa analiza muzycznej warstwy memu ujawnia szereg cech, które niewątpliwie przyczyniły się do jego wiralowego sukcesu. Mimo pozornej prostoty, kompozycja zawiera elementy sugerujące, że jej niekonwencjonalna struktura jest w istocie głęboko przemyślana. Główny motyw muzyczny memu to minimalistyczna kompozycja na syntezator samplujący Esoniq Mirage przy użyciu barw przypominających organy i perkusję. Jej analiza harmoniczna wskazuje na progresję akordów: C-dur, a-moll, G-dur, G-dur, po której następuje sekwencja C-dur, h-zmniejszony (Bdim), a-moll, C-dur. Użycie akordu zmniejszonego (Bdim) jest zabiegiem nietypowym dla prostych, popularnych melodii. Wprowadza on element dysonansowy i napięcie, które skutecznie przełamuje pogodny charakter utworu oraz dodaje mu nieoczekiwanej głębi.

Szczególną uwagę zwraca warstwa rytmiczna. Słychać, że nagranie zostało zrealizowane bez użycia metronomu. Skutkuje to naturalnymi, subtelnymi wahaniami

tempa. Poza tym istnieją również mikrorytmiczne niezgodności między partiami organów i perkusji. Ta mikrorytmiczna niedoskonałość paradoksalnie wzmacnia autentyczność przekazu. Podkreśla amatorski charakter wykonania, co jest spójne z wizerunkiem kota-muzyka. Mimo tych nieregularności, struktura harmoniczna utworu jest logiczna i zawiera klasyczne rozwiązania, jak następstwo dominantowo-toniczne G-dur – C -dur, co nadaje całości formy logiczne podstawy. Nieporadność wykonania i prosta harmonika większości przebiegu tworzą odczuwalny kontrast do niespodziewanego harmonicznego wtrącenia akordu zmniejszonego pod koniec. Można interpretować to jako metaforę radzenia sobie pomimo ograniczeń. Elementarna, banalna linia melodyczna podąża ruchem falowym za podstawowymi triadami w akordach. W ten sposób również melodia dodatkowo podkreśla pojawiający się jako ostatnia dominanta dysonansowy akord h-d-f, dzięki czemu zwiększa się jego wyrazistość. Przewidywalna, banalna struktura utworu i pojawienie się pod koniec elementu wyrazistego dysonansu najpierw buduje muzyczne oczekiwania, a następnie je zawodzi. Można to rozumieć w kategoriach opisywanego w drugiej części niniejszej dysertacji modelu uncertainty–surprise (Pearce, 2018). Taka struktura oraz jednoczesna popularność materiału zdają się potwierdzać ustalenia tradycyjnej muzykologii systematycznej. Rolę naruszeń przewidywalnych schematów jako kluczowego czynnika w estetycznym odbiorze muzyki podkreślali Leonard B. Meyer (Meyer, 1956) oraz David Huron (Huron, 2006).

W późniejszej części oryginalnego nagrania pojawia się drugi, kontrastujący motyw muzyczny. Charakteryzuje się on wolniejszym tempem i zmianą instrumentarium. Fundamentem jest tu pochod basowy typu *walking* wykonany na instrumencie cyfrowym o brzmieniu imitującym kontrabas, za to główna melodia wykorzystuje barwę przypominającą miauczenie kota. Zestaw perkusyjny pozostaje ten sam, co w części pierwszej, co zapewnia spójność całej kompozycji.

Potomstwo:

Keyboard Cat to jeden z najważniejszych internetowych memów, który znacznie przyłożył się do ustanowienia popularności i statusu kotów w kulturze memowej. Warto wspomnieć chociażby o istnieniu marki Cats on Synthesizers in Space, która bez wątplenia zawdzięcza swój sukces memowi *Keyboard Cat*. Cats on Synthesizers in Space sklep internetowy, strona na Facebooku o 250 tysiącach obserwatorów, a także oficyna wydawnicza na platformie Bandcamp.

Sam twórca *Keyboard Cata* stworzył bardzo wiele rozmaitych remiksów oraz kontynuacji, które publikuje na Youtube czy Facebooku. Warto wspomnieć chociażby o tym, że de facto *keyboard cat* to teraz pewna funkcja, a piastuje ją już trzecia kotka, które sam autor oryginału, Charlie Schmidt, nazywa *reinkarnacjami*. Kotka Bento wystąpiła na przykład w filmie *keyboard cat bongos*, który zdobył prawie dwa miliony odsłon na facebooku (Schmidt, 2016).

Keyboard cat pojawił się także w różnych mediach: W roku 2014 w programie Jimmy Kimmel Live! zaprezentowano krótki film stylizowany na trailer do nieistniejącego filmu *Ameowadeus* (Amiaudeusz), czyli parodię znanego filmu *Amadeus* o W.A. Mozarcie (Reed, 2014). W tym filmiku *Chomik na Fortepianie* (Hamster On a Piano) zagrany przez Christopha Waltza zazdrości talentu *Keyboard Catowi* (Kevin Spacey). W filmiku występują także Sir Ben Kingsley, Gary Oldman, Mandy Patinkin, Abbie Cornish oraz Jimmy Kimmel (Reed, 2014). Również w 2014 Keyboard Cat pojawił się w amerykańskim hollywoodzkim filmie o superbohaterach Teenage Mutant Ninja Turtles.

Powstała także kryptowaluta - \$KEYCAT, której faktycznym udziałowcem jest twórca oryginalnego filmiku. Istnieje też Non fungible Token NFT, tzn. można formalnie być właścicielem filmiku z Keyboard Catem. Ostatni kupiec zapłacił za niego 10 ETH, czyli na 31.05.2025 jest wart 96273,75 PLN (Coinmarketcap, 2025).

1.5.4. *Darude – Sandstorm*



@rawtrout007 5 lat temu

this doesn't sound like the song i was looking for



186 tys.



Odpowiedz

∨ 621 odpowiedzi



@yuckmane 4 lata temu

A moment of silence for those who have asked for a song and have been sent here



215 tys.



Odpowiedz

∨ 729 odpowiedzi

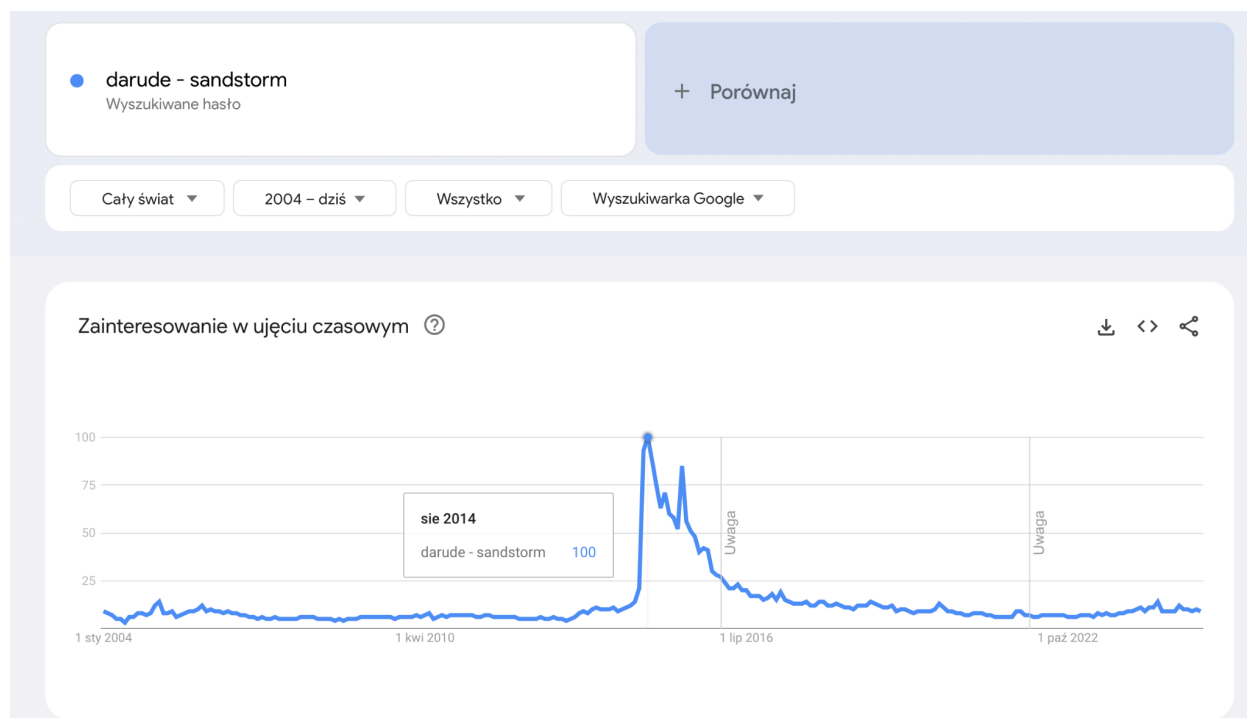
Pochodzenie: *Darude – Sandstorm* to kolejny bardzo ważny w kulturze internetowej mem muzyczny, który można zaklasyfikować jako formę trollingu, pod tym względem pewną pochodną *rickrollingu*. W tym wypadku także sam internetowy muzyczny mem esencjonalnie nie jest żadnym konkretnym utworem, filmikiem, ani inną zamkniętą w konkretnym pliku treścią. Memem jest zachowanie, zwyczaj, które jest naśladowane i osiągnęło status wiralny, stało się jasne i rozpoznawalne dla milionów internautów na całym świecie (Sloane, 2018). Zanim stał się internetowym memem, *Darude – Sandstorm* był już znanym utworem muzyki elektronicznej z gatunku trance, czego dowodzi chociażby jego 83 pozycja w amerykańskim notowaniu Billboard (Billboard, 2025). Utwór w tempie 136 uderzeń na minutę i tonacji e-moll zawiera wiele typowych cech dla gatunku trance, jak charakterystyczny miarowy rytm bębna basowego (uderzenia tzw. stopą, *kick drum*) na każdą ćwierćnutę, synkopowany rytm pary talerzy perkusyjnych (*hi-hat*) oraz stopniowo narastająca forma zawierająca charakterystyczne elementy kulminacji (*drop*) oraz wyciszenia (*breakdown*). Zawiera również

syntezatorowe brzmienia typowe dla muzyki trance z przełomu milleniów, w tym przypadku pochodzące m.in. z syntezatora Roland JP-8080.

Co ciekawe sama nazwa *Sandstorm*, która de facto stanowi co najmniej tak samo ważny element memu jak treść powiązanego utworu muzycznego, została zaczerpnięta i przetworzona. Producent Darude wyjawiał, że syntezator Roland JP8080 wyświetlał komunikat 'Sand Storm' na swoim wyświetlaczu po uruchomieniu i to stało się dla niego inspiracją (Mullen, 2018). Ponadto Sand Storm to nazwa pierwszego domyślnego fabrycznego brzmienia syntezatora Roland JP8080 i na tym właśnie brzmieniu producent oparł swój utwór (Mullen, 2018).



Utwór fińskiego producenta muzyki elektronicznej był przebojem w latach dwutysięcznych, aczkolwiek jego prawdziwy szczyt popularności, 'drugą młodość' należy powiązać z latami 2014-2016, kiedy to funkcjonował jako mem. Tutaj, podobnie jak w innych przykładach, istotą memu jest zwiedzenie oczekiwań w komunikacji, rodzaj żartu, czy trollingu. Pasuje tu także określenie *bait-and-switch* (Shifman, 2012).



Utwór w 2025 roku na serwisie YouTube przy głównym adresie audiowideo dobija do 300 milionów odsłon, a razem z remiksującymi go krótkimi klipami YouTube Shorts (łącznie około 120 milionów odsłon) przekracza 420 milionów odsłon, co świadczy o skali tego fenomenu.

Żeby określić początki internetowego zjawiska *Darude - Sandstorm* należy dokonać nieco internetowej 'archeologii' i wyszukać najwcześniejsze charakterystyczne zastosowania. Na YouTube wciąż istnieje film promocyjny z kanału GameStream, na którym utwór *Sandstorm* stanowi muzyczne tło do prezentacji rozgrywki w grze *Call of Duty 4: Modern Warfare* (GameStream, 2007). Film z 2007 roku przekroczył cztery miliony odsłon i może pokazywać wczesne skojarzenie *Sandstorm* z kulturą gier wideo

(Dictionary.com, 2018). W roku 2013 zaczął towarzyszyć transmisjom na żywo przedstawiającym rozgrywkę w gry wideo na portalu Twitch i prawdopodobnie właśnie w czacie na platformie Twitch narodził się charakterystyczny mem. Został on sugestywnie przedstawiony na poniższym internetowym memie obrazkowym typowym dla estetyki tej epoki:



Treść:

Przechodząc do meritum, internetowy mem *Darude - Sandstorm* polega na tym, że jeżeli w dowolnej sytuacji osoba A zapyta o tytuł słyszanej właśnie piosenki, osoba B odpowiada „Darude - Sandstorm”, niezależnie od tego, o jaką piosenkę chodziło. W roku 2014 na portalu Urban Dictionary hasło *Darude - Sandstorm* zostało opisane jako „Tytuł każdej istniejącej piosenki” (The name of every song in existence) (Urbandictionary, 2025).

Potomstwo:

Dziedzictwo memu *Darude - Sandstorm* jest rozpoznawalne dziś głównie poprzez zwiększenie popularności samej piosenki o tym tytule. W nowej fali jej popularności związanej z propagacją mema odbyły się też ciekawe wydarzenia, jak na przykład to, że serwis YouTube na prima aprilis w 2015 roku do każdego odtwarzacza filmu na serwisie dodał specjalny przycisk, który pytał użytkownika, czy jednak chodziło mu o utwór *Darude - Sandstorm* (*Did you mean Darude - Sandstorm?*) i po kliknięciu uruchamiał ten utwór (Gutelle, 2005). Omawiany przebój stał się również hymnem meczy e-sportowych, w tym też światowych finałów (Higgins, 2016). Utwór został również zagrany jako część ceremonii otwarcia olimpiady w 2024 roku i, mimo że nie można jednoznacznie powiedzieć, czy po prostu był to wybór klasycznego trance’owego energicznego utworu, czy może jednak utwór wybrano specjalnie ze względu na popularność memu, recepcja w internecie wskazuje na wysoką świadomość istnienia memu przez użytkowników internetu (Wong, 2024).

1. 5. 5. *MLG Airhorn*

[Myinstants](#) > [Sounds](#) > [Sound Effects](#) > MLG AIR HORN!!!!!!!!!!!!

MLG AIR HORN!!!!!!!!!!!!



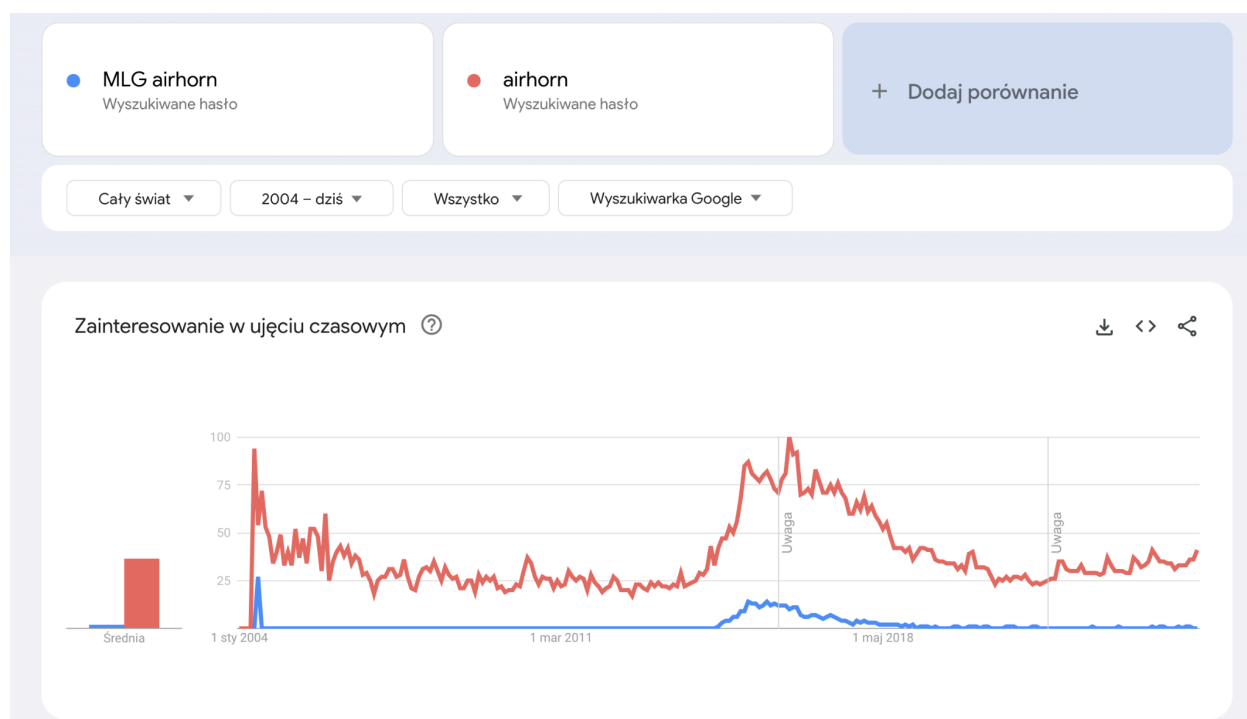
♥ **98,037 users** favorited this sound button

1,232,363 views

Pochodzenie: airhorn, czyli sygnalizator pneumatyczny lub przenośna trąbka pneumatyczna to puszka skompresowanego powietrza z zamontowaną piszczalką – urządzenie często używane przez amerykańskich kibiców sportowych, aby wywołać hałas (Mercer, 2014; NCAA, 2021). Stosowana była również na Jamajce już od lat 70-tych na imprezach z muzyką reggae i dub (Weiss, 2018). Wczesnym znanym dokumentem użycia w muzyce jamajskiej jest dźwięk niskiego airhornu w nagraniu Midnight Ravers & Ravers Version (Tuff Gong) Boba Marleya i the Wailers z 1972 (Weiss, 2018).

Treść: *MLG Airhorn* to kilkusekundowe nagranie dźwięku przenośnej trąbki pneumatycznej, które w kulturze internetowej niesie rozpoznawalne znaczenia w

określonych kontekstach, np oznacza ekscytację na meczach e-sportowych (Clownboss, 2012). To przykład tego, jak prosty może być muzyczny mem i ciekawy przypadek na pograniczu muzyki oraz dźwięku. Dźwięk airhorn stał się memem internetowym w roku 2010 (Clownboss, 2012). Nie zawsze jest kojarzony z nazwą *MLG Airhorn*, często jest to po prostu kilkusekundowy dźwięk tego instrumentu i nazywany jest po prostu airhorn, ale pełni taką samą funkcję. To łatwo rozpoznawalny sygnał. Denotował ekscytację, ale również, a może przede wszystkim – momenty humorystyczne. Sam dźwięk airhorn humorystycznie podkreślał najbardziej wyraziste fragmenty materiałów wideo.



Komentarz do wykresu: wyszukiwanie airhorn w internecie oczywiście oprócz efektu dźwiękowego odnosi się także do faktycznych trąbek pneumatycznych, które bywają w internecie sprzedawane. Konkretny dopisek 'MLG Airhorn' odnosi się za to tylko do memu internetowego.

Tytuł *MLG Airhorn* zdradza nam tożsamość dźwięku oraz jego kontekst – MLG to Major League Gaming, profesjonalna firma zajmująca się organizacją wysokiej rangi turniejów e-sportowych, głównie w grze Call of Duty (Caldwell, 2014; Bartel, 2016). Z czasem

używanie określenia 'MLG' przy opisywaniu własnych umiejętności zaczęło być kojarzone z pewnym elityzmem i zarozumiałością (Asddfds, 2016). Gracze umieszczali logo MLG przy filmikach pokazujących ich niezwykle dokonania i około roku 2015 przeważać zaczęły parodie takich popisów umiejętności (Wickman, 2015), rozpoczynając często pojawiającą się w periodyzacjach internetowych memów 'erę MLG', gdzie MLG kojarzone jest już przede wszystkim nie tyle z e-sportową organizacją, ile z absurdalną estetyką humoru, który pojawił się w tej niszy (Gwennaen, 2025). Dźwięk airhorn prawdopodobnie pierwotnie pojawiał się jako sygnalizator autentycznej sportowej ekscytacji, tak samo jak na trybunach stadionów. Następnie przeżył pewną ewolucję znaczeniową:

airhorn jako szczerzy sygnał ekscytacji przy e-sporcie

airhorn ironicznie użyty, aby wyśmiać ekscytację przy e-sporcie

airhorn jako szczerzy sygnał intensywnych momentów dowcipnych montażu wideo oraz streamów

airhorn jako podstawa 'airhorn remix', gdzie jego barwa sama w sobie stanowi pointę

Zastosowanie dźwięku airhorn do podkreślania momentów ekscytacji jest zgodne z praktyką kibiców sportowych, ale jednak przeniesienie tego dźwięku do internetu, a szczególnie do treści zupełnie niezwiązanych ze sportem jest pewną zmianą kontekstu. Charakterystyczne dla tego właśnie memu muzycznego jest jego podwójne zastosowanie w praktyce – nie tylko można go usłyszeć w filmikach, memach wideo czy ich kompilacjach, ale także w transmisjach na żywo, gdzie pełni ważną funkcję sygnalizacyjną. Na stronie myinstants służącej jako prosty sampler, czy też *soundboard* dla osób udostępniających na żywo treść swojego ekranu, czyli streamerów, dźwięk 'MLG AIR HORN!!!!!!' ma ponad 1 200 000 odsłon. Taki *soundboard* to zbiór pojedynczych dźwięków, sampli, w formie wirtualnej klawiatury, które dana osoba może mieć pod ręką do natychmiastowego wykorzystania (Soundalerts, 2023). W praktyce to ważne narzędzie dla osób prowadzących transmisje na żywo, czyli streamy, kiedy dzieje się coś szczególnego, co warto podkreślić. W takich momentach streamer

uruchamia dany dźwięk, który pojawia się w transmisji wraz z głosem prowadzącego i jednoznacznie narzuca odbiorcom wrażenie, że stało się coś wartego uwagi, zarazem podpowiadając pewną nastrojowość czy szczególną emocję. Takie *soundboardy* często zawierają właśnie bieżące w ramach danej społeczności memy dźwiękowe, co pomaga konsolidować wspólną narrację i poczucie społeczności.

Za dodatkowo wyróżnialną praktykę zastosowania dźwięku airhorn należy uznać airhorn remix - czyli pewnego rodzaju komiczny remix, cover, wykonanie określonej piosenki lub jej fragmentu używając charakterystycznego sampla airhorn. Wszystkie najbardziej popularne tego typu utwory pochodzą z lat 2014-2016 i mocno wiążą się z estetyką czy też erą memów MLG (Brandon, 2024; Overall-Estate1349, 2022). Wiele znanych airhorn remixów stworzył YouTuber o pseudonimie *Cyranek* czyli czołowa postać w środowisku memów MLG. Warto wymienić chociażby:

MLG Darude Sandstorm (Cyranek, 2015a) - 2.5 mln odsłon

Disney Intro (MLG Air Horn Remix) (Graving, 2015) - 6.9 mln odsłon

Let it go – MLG airhorn remix (Cyranek, 2015b) - 7.6 mln odsł

Undertale Megalovania - MLG Airhorn Remix (Cyranek, 2015c) - 7.2 mln odsłon

Beethoven – Moonlight sonata for air horns (Equalizer Magazine, 2013) - 671 000 odsłuchań

Gorillaz – Feel good inc. air horn remix (PhantomPhaze, 2014) - 250 000 odsłon

Potomstwo:

Airhorn sam w sobie przeżył pewną semantyczną ewolucję, polegającą na zmianach kontekstów użycia i sensów wraz z ewolucją kultury internetowej. Poza tym stanowi on bezprecedensowo popularny dźwięk reakcji '*reaction sound*', czyli krótki dźwięk, który ma sugerować odbiorcy jak ustosunkować się do danej treści. Zatem częściowo uTORował on drogę późniejszym dźwiękom reakcji, jak *Vine Boom* czy też *bruh*. Przede wszystkim jednak airhorn jako błyskawicznie rozpoznawalny, krótki i nasycony znaczeniowo sygnał przyczynił się do ustanowienia całej 'ery MLG' (Brandon, 2024; Overall-Estate1349, 2022), jednej z powszechnie wyróżnianych epok estetycznych w dotychczasowej historii memów internetowych.

Dźwięk airhorn i inne sygnały memów ery MLG (Brandon, 2024) znalazły się w materiale wideo z serii *You Can't Stump the Trump*. Spencer zwraca uwagę, że Donald Trump w 2015 roku udostępnił taki materiał na swoim oficjalnym koncie na platformie Twitter (Spencer, 2025). W tym materiale sygnały dźwiękowe typowe dla MLG zostały przekształcone w broń retoryczną: miały nadać Trumpowi wizerunek triumfującego drapieżnika i jednocześnie ośmieszać jego przeciwników (Spencer, 2025). Była to część szerszego zjawiska określanego przez badaczy i internautów jako „Great Meme War”, czyli wojna memów towarzysząca kampanii prezydenckiej w USA w latach 2015–2016, w której posługiwano się estetyką MLG (Spencer, 2025). W ten sposób airhorn stał się sygnałem agresywnej komunikacji politycznej.

2.1. Wprowadzenie memetyki

Pojęcie „mem” stworzył i spopularyzował Richard Dawkins w swojej książce *Samolubny Gen* (*The Selfish Gene*) w 1976 roku. Sama popularyzatorska książka dotyczyła przede wszystkim ewolucji genetycznej. Dawkins wprowadził pojęcie replikatora, które pięć lat później zdefiniował jako ‘cokolwiek we wszechświecie, co tworzy własne kopie’ (Dawkins, 1981). Badacz opisywał ewolucję biologiczną z perspektywy genu jako ‘samolubnego replikatora’, czyli jednostki informacyjnej, która przede wszystkim dąży do stworzenia własnych kopii. Zwracał uwagę przede wszystkim na proces kopiowania pewnej informacji, instrukcji. Ta koncepcja zaprowadziła go do tego, aby podać również inny przykład replikatora niż gen i był to właśnie mem. Warto podkreślić, że pojęcie memu zostało wprowadzone w bardzo znanej książce popularno-naukowej. Sam ten fakt bez wątplenia wpłynął na dalsze losy tego pojęcia. Mem to dziś hasło pojawiające się bardzo szeroko w kulturze popularnej (Shifman, 2014). Za to sfera, z której jest niemal wykluczone, to właśnie biologiczne nauki ewolucyjne. Również w biomuzykologii pojęcie memu wzbudza niechęć (Savage, 2025), co wiąże się z obiegową opinią, że memetyka to pseudonauka, lub „pozbawiona znaczenia metafora” (Aunger, 2000).

Co istotne, Dawkins w *Samolubnym Genie* pisał o memach bardzo ogólnie. Nie proponował powstania memetyki jako nauki, cała formuła zajmuje dosłownie 10 stron. W związku z ogromną popularnością książki nieostre pojęcie memu również zyskało na poplarności i wielu badaczy podjęło próby własnych kontynuacji i reformulacji memetyki. Do pseudonaukowego charakteru memetyki na pewno wielu przekonała skądinąd bardzo ciekawa książka *The Meme Machine* Susan Blackmore. Narracja badaczki sugerowała, że przy pomocy memetyki chce ona wyjaśnić ogrom zjawisk związanych z człowiekiem, co podniosło alarm wśród społeczności naukowej.

Od 1997 do 2005 działał *Journal of memetics*, którego treści określam jako pierwszą falę memetyki. Wyłaniający się z 38 opublikowanych w nim artykułów przede wszystkim

sugeruje istnienie memetyki jako autonomicznej dyscypliny naukowej. Ja za to w dalej w niniejszej pracy będę traktował memetykę nie jako niezależną naukę całkowicie oddzieloną od ewolucyjnych nauk o kulturze, ale jako program naukowy w rozumieniu Lakatosa (Lakatos, 1970), który można spójnie połączyć szerszym kontekstem teoretycznym. Przedstawiona koncepcja programu naukowego memetyki jest, jak sądzę, niezbędna do analizy ewolucji współczesnej kultury, i dająca się całkowicie zintegrować z naukami o ewolucji kultury.

Przedstawię wszystkie najważniejsze zarzuty odnośnie do memetyki, które pochopnie pogrzebały ją jako pseudonaukę i spróbuję podjąć z nimi przekonującą polemikę. Stawiam zatem kluczową dla niniejszej pracy tezę, że wzbogacone o memetykę nauki o ewolucji kultury stanowią najbardziej spójne podejście do stworzenia ewolucyjnej ontologii muzyki oraz do opisu jej dynamicznego rozprzestrzeniania się w oparciu o emergentną, ewolucyjnie rozwijaną sprawczość.

2.2. Istotne aspekty memetyki

Poza tym, że w niniejszej pracy opieram się na bardzo podstawowym modelu memetyki opisanym przez Richarda Dawkinsa, który rozwijam o różne opisane hipotezy pomocnicze zaczerpnięte z zewnętrznych dyscyplin, chcę wymienić także koncepcje dobrze zintegrowane w badaniach nad memetyką oraz ewolucją kulturową abstrahującą od pojęcia memu, które także zajmują w proponowanym przeze mnie modelu ważne miejsca. Na początku trzeba wspomnieć, że w badaniach ewolucji kulturowej funkcjonuje zapożyczona z genetyki koncepcja dryfu (Hahn, 2003), odnosząca się do zmian, które w pulach replikatorów zachodzą stochastycznie, jako pewnego rodzaju szum tła i nie są tym samym efektem jakiegokolwiek presji selekcyjnej. Oznacza to, że nie na każdym polu zmiennością kulturową zarządzają mechanizmy selekcji. Istnieje również zmienność określana jako losowa. Pokazuje to przykład zmienności imion nadawanych amerykańskim dzieciom (Hahn, 2003), gdzie nowe warianty stabilnie pojawiają się i znikają według tych samych regularności od stu lat. Bentley i współpracownicy (2006) pokazują także, że charakterystyczna dla dryfu dynamika towarzyszy również częstotliwości znikania konkretnych albumów muzycznych z listy Billboard top 100. Zjawisko to nie musi być jednak koniecznie dryfem.

Po pierwsze w tym wypadku losowość to tylko upraszczająca kategoria, która opisuje łączny, w praktyce graniczący z nieprzewidywalnością efekt działania niezliczonych pomniejszych presji selekcyjnych. Po drugie, to, czy w memetyce w ogóle istnieje dryf jest ontologicznie nieoczywiste. Każdy system poznawczy podlega ewolucji zgodnie z zasadą wolnej energii (Friston, 2006), a z założeń uniwersalnego darwinizmu wynika, że każdy mem musi wytwarzać choćby najmniejsze presje selekcyjne na swoje otoczenie – inaczej nie byłby memem. Zatem w memetyce, zamiast interpretować działanie dryfu jako pola, w którym nie zachodzi memetyczna selekcja, proponuję, tak jak wyżej, zastąpić pojęcie dryfu efektem równoważących się dynamicznych sił selekcji operujących w tle.

Kontynuując, to, że tempo znikania z listy albumów jest stabilne, nie może prowadzić do wniosku, że na pojawiające się na liście albumy nie działają presje selekcyjne. Kształt listy przebojów nie wynika z działania przypadku. Wręcz przeciwnie – ‘cykl życia’ piosenek na liście billboard można zamodelować przy użyciu bardzo zaawansowanych narzędzi statystycznych (Bradlow, 2001). Poza tym to tylko jeden z aspektów dynamiki popularności piosenek. Badania wykazują (Rosati, 2021), że cykl popularności piosenek dokładnie pokrywa się z modelem zarażeń epidemicznych – w uproszczeniu chodzi o szybki wzrost popularności, a następnie wolny, stopniowy jej spadek. Co ciekawe, ta sama, bardzo charakterystyczna dynamika rozprzestrzeniania się dotyczy również memów internetowych (Wang, 2011; Zhai, 2021). Wysokie podobieństwo rozprzestrzeniania się patogenów, popularnych piosenek oraz internetowych memów stanowi bardzo silne wskazanie, że posługująca się analogią wirusa memetyka muzyki to wartościowy poznawczo model badawczy.

Omówienia wymagają także dwie koncepcje z memetyki Daniela Dennetta. Pierwsza z nich to kompetencja bez zrozumienia (*competence without comprehension*) (Dennett, 1995). Dennett wielokrotnie podkreśla ‘ślepe’ działanie procesu ewolucji, które jednak poprzez konsekwentne eliminowanie błędów prowadzi do odnajdywania optymalnych, skutecznych rozwiązań. W ten sposób w naturze możemy odnaleźć niezliczone przykłady zachowań, które wyglądają na odgórnie, inteligentnie zaprojektowane, a w rzeczywistości wynikają z powolnych, bezmyślnych oddolnych procesów selekcji (Dennett, 1995). Ten sam sposób nabywania kompetencji funkcjonuje w sferze kultury. Dennett mówi o przedziwnym odwróceniu rozumowania (*strange inversion of reasoning*), które zauważa zarówno u Darwina, jak i Turinga (Dennett, 2017). W rzeczywistości ‘odwrócone’ rozumowanie obu badaczy to optymalny sposób opisu przedmiotów ich badań. Zatem Dennett zauważa pewne podobieństwo w logice zachowania organizmów żywych oraz funkcjonowania komputerów. Organizm nie musi rozumieć, co robi, żeby robić to skutecznie, a komputer nie musi znać zasad arytmetyki, żeby ją skutecznie uprawiać (Dennett, 2017).

Rozwinięcie tej koncepcji w memetyce muzyki może mieć donośne skutki dla dydaktyki muzycznej. Klasyczna dydaktyka nastawiona jest na propagowanie muzycznej wiedzy, która ma stanowić podstawę dla muzycznych kompetencji. Za to Dennettowskie przedziwne odwrócenie rozumowania stanowi solidną podstawę do stwierdzenia, że kompetencja gry na instrumencie nie wynika z racjonalnej wiedzy obecnej w umyśle i pamięci muzyka wykonawcy. Pozwala to przypuszczać, że nauka gry poprzez ustalenie prymatu w bezpośrednim doświadczeniu i intuicji może być skuteczniejsza niż metoda opierania wykonawczej kompetencji na intelektualnym rozumieniu.

Poza tym ta koncepcja ma też istotne znaczenie dla opisu tego, skąd pochodzi i na czym polega sprawczość memów, która objawia się na przykład w ich sposobach na prowokowanie własnej replikacji. Piosenka nie musi intencjonalnie, celowo, dynamicznie zabiegać o replikację na wzór tego, jak zwierzęta zabiegają o pożywienie. Ale wśród rozlicznej i różnorodnej populacji piosenek znajdzie się taka, która akurat ową replikację sprowokuje skutecznie i to ona zostanie więcej razy usłyszana i to ona z większym prawdopodobieństwem 'wyda potomstwo', to znaczy, kolejne powstałe piosenki będą pod pewnymi względami właśnie ją przypominać. Patrząc zatem na całą populację piosenek jako spójną grupę. Proces częściowo ślepej generacji rozlicznych wariantów i eliminacji tych nieskutecznych, czyli proces ewolucyjny, doprowadza do zachowania takich memów, które wykazują się skutecznością. Według Michaela Levina w takiej perspektywie możemy mówić o sprawczości, agencjalności samego procesu ewolucyjnego (Levin, 2025).

Ostatnia koncepcja to Dennettowska przestrzeń darwinowska (*darwinian space*) (Dennett, 2017). Idea ta określa, że procesy darwinowskie oraz wszystkie procesy ewolucyjne można zlokalizować jako punkt na trójwymiarowym spektrum. Według Dennetta zerojedynkowe ustalenie, czy dany proces jest ewolucyjny, czy też nie jest, to przesadne uproszczenie. Ma to ważne znaczenie dla procesu ewolucji memetycznej, ponieważ w wielu przypadkach powstaje pytanie – czy na pewno można powiedzieć, że mem ewoluował, skoro wygląda na to, że został świadomie zaprojektowany? Ważnym dylematem jest kwestia kreatywności, twórczości: czy możemy mówić o ewolucji

muzyki, jeśli tym procesem rządzi celowe projektowanie utworów przez kompozytorów? Ma tutaj znaczenie także logika intencjonalności wynikająca z koncepcji kompetencji bez zrozumienia. Moja propozycja skrótowej odpowiedzi na ten kontrowersyjny dylemat brzmi: konceptualizując proces ewolucji memetycznej, możemy pominąć kwestię pochodzenia danego memu. Źródło jego pochodzenia jest drugorzędne – ważne jest to, czy jego cechy replikują się w populacji, czy też nie. Innymi słowy, czy te cechy można dostrzec w populacji, czy są w niej obecne. Jeżeli są, to proces zmienności częstości pojawiania się takich cech w populacji jest de facto procesem ewolucji darwinowskiej. Taki proces nie musi świadczyć o tym, że kompozytor nie miał twórczego udziału lub wolnej woli przy tworzeniu utworu. Jeżeli na zasadzie statystycznej analizy cech populacji utworów muzycznych możemy zauważyć ewolucyjne prawidłowości, wtedy można wyciągnąć wniosek, że prawdopodobnie w umysłach kompozytorów funkcjonują pewne wspólne populacje memów. Ale nie ma dowodów czy nawet wskazań na to, że całe funkcjonowanie i wola człowieka jest zależna tylko i wyłącznie od memów.

2.3. Odpowiedzi na zarzuty

Koncepcję memów zaproponował biolog Richard Dawkins (1976). Oryginalnie stanowiła pewną teoretyczną abstrakcję. Co istotne, w artykułach z zakresu prawdopodobnie najprężniej rozwijającej się gałęzi memetyki, czyli algorytmiki memetycznej, badacze nazywają teorię Dawkinsa 'filozoficzną' (Neri, 2012, Tang, 2019). Zapewne decyduje o tym to, że dla obliczeń stosujących memetyczne algorytmy wszystko, co nie jest ściśle obliczalne staje się filozoficzne. W każdym razie, jak widać, dziś memetykę można odnaleźć nie tylko w literaturze teoretycznej. Znajduje także zastosowania m.in. w badaniach z pogranicza biologii, kognitywistyki i nauk o informacji (np. Fitch, 2022; Duch, 2021). Odnoszą się do niej również badacze, którzy rozwijają modele poznawcze i biologiczne, tacy jak Karl Friston czy Michael Levin (Levin, 2025). Mimo tej szerokiej obecności w naukach przyrodniczych i kognitywnych, memetyka spotyka się z różnorodną krytyką. Nasila się ona zwłaszcza wśród antropologów kulturowych i kulturoznawców, którzy podkreślają ograniczenia redukcjonistycznych ujęć kultury. Sceptyczne narracje istnieją też w samej biologii: przykładem jest Eva Jablonka, która podważa zasadność traktowania memów jako odrębnych replikatorów kulturowych (Jablonka, 2005).

Poniżej zatem staram się przedyskutować wszystkie moim zdaniem najważniejsze zarzuty. Spróbuję ukazać zarówno podobieństwa, jak i różnice między memetyką a nauką o ewolucji kulturowej, kiedy konceptualizowane są jako odrębne nauki. Przy omówieniu zarzutów i odpowiedzi przedstawię dlaczego uważam, że memetyka może oraz powinna być włączona do nowej syntezy ewolucyjnych nauk o kulturze. Choć prace badacze z przełomu wieków z kręgów *Journal of Memetics*, które dziś nazwałbym pierwszą falą memetyki, były w naukowym konflikcie z pracami badawczymi z zakresu głównego nurtu badań nad ewolucją kulturową, a ich dyskusje sugerowały, że należy wybrać jedną lub drugą tradycję, dziś można wykazać, że poprawną drogą rozwoju nie jest alternatywa, ale synteza.

Przede wszystkim rozważę najważniejsze punkty krytyczne ze zbiorowej publikacji *The Status Of memetics as a Science* (Aunger, 2000). Jest to zbiór publikacji badaczy z kręgów biologii ewolucyjnej, kognitywistyki, psychologii i ewolucyjnych nauk o kulturze, gdzie splatają się pozytywne oraz negatywne wizje memetyki. Możemy tu też znaleźć szereg metodologicznych wątpliwości. Odnoszę się do wszystkich, które uznają za najbardziej adekwatne i wymagające wyjaśnienia. Niektóre należy kategorycznie odrzucić, inne doprecyzować, a jeszcze z innymi trzeba się zgodzić. Jeżeli dane zarzuty negują bardzo istotny obszar pojęciowy, zdarza się, że szeroko zakrojona z konieczności odpowiedź od razu stanowi odparcie szeregu innych sceptycznych propozycji.

Najpoważniejszy zarzut Roberta Aungera – sugestia, że memy nie istnieją:

(...) many researchers blithely discuss features of memes, ignoring the fact that their existence has yet to be proven. (Aunger, 2000)

(...) wielu badaczy beztrąsko dyskutuje o memach, ignorując fakt, że nikt nie wykazał jeszcze, że istnieją.

Takie podejście przenosi na memetykę ciężar dowodzenia, że można w naturze spotkać mem, zarazem sugerując, że jest to trudne zadanie. Memetycy najprościej mogą ten problem rozwiązać, wskazując na dowolny produkt dzisiejszej kultury, ponieważ według podstawowej wykładni memetycznego paradygmatu kultura ma memetyczny charakter, wykazuje memetyczne „zachowanie”, co nie musi być wprost odzwierciedlone w ostatecznej materialnej formie danej informacji kulturowej. Ale tego typu odpowiedź może być niewystarczająca dla twardego sceptyka.

Czołowy memetyk Daniel Dennett odpowiada na sceptycyzm wobec istnienia memów, cytując poglądy ewolucyjnego językoznawcy Jackendoffa (Dennett, 2017). Jackendoff definiuje słowa jako „na wpół niezależne struktury informacyjne” (semi-autonomous informational structures) (Jackendoff, 2002). Dennett wnioskuje zatem, że jeśli słowa istnieją, to memy także. Spójność perspektywy Jackendoffa z memetyką, a także szerzej rozumianą memetykę w językoznawstwie opisują również inni badacze (Planer,

2024; Jan, 2016). Sceptyczni wobec memetyki ewolucjoniści nadal mogą odpowiedzieć, że to 'zaledwie teoria' i przypuszczenia.

Zatem, w kolejnym kroku odwołuję się do realnych przykładów empirycznych, które stanowią silny argument wspierający memetykę. Konsekwencją istnienia memetycznej selekcji musi być to, że kolejne pokolenia memów posiadają cechy aktywnie napędzające ich replikację. Klasyczny model ewolucji kulturowej bez istnienia kulturowego replikatora po prostu nie prowadzi do takiej predykcji. A zatem należy przedstawić przykłady takich produktów kultury, które konkurują ze sobą o replikację. Takie, z których struktury wynika, że nie tylko wpasowują się w ludzką niszę poznawczą, ale manipulują człowiekiem, żeby sprowokować własną replikację.

Jako przykład takiej wzajemnej konkurencji o ludzką uwagę oraz istnienia szeregu wyłonionych oddolnie, niezaprojektowanych inteligentnie cech produktu kultury, przy pomocy których skuteczniej napędza on swoją replikację możemy podać treści z serwisu TikTok. Jest to szczególnie ciekawe dla niniejszej pracy, ponieważ TikTok jest zarówno platformą muzyczną jak i społecznościową w warstwie wideo oraz interakcji. A zatem przytoczone poniżej cechy i strategie mają kluczowy wpływ na jedno z największych mediów dystrybucji muzyki (w szczególności trendów muzycznych) na świecie. A przytoczone kryteria dotyczą wyłącznie tego, co muzykologia nazwałaby treściami pozamuzycznymi.

Memetyka bywa kojarzona z ilościowymi opisami ewolucji kulturowej, co w czasach przełomu wieków, czyli najgorętszych debat odnośnie memetyki, było znacznie trudniejsze do przeprowadzenia niż obecnie. W badaniach TikToka w tym celu zastosowano uczenie maszynowe. Tego typu analiza (Cervi, 2021) miała za zadanie odnaleźć w badanych treściach z platformy TikTok oraz związanych z nimi metadanych wskaźniki predykcyjne dla osiągnięcia sukcesu replikacyjnego. Badanie empiryczne wskazało, że jednym z najsilniejszych predyktorów sukcesu jest popularność samego twórcy, ale istotną rolę odgrywają również konkretne cechy samego materiału

audiowizualnego.

Twórcy doszli do wniosku, że cztery konkretne cechy treści szczególnie wzmagają zasięgi. Po pierwsze są to materiały, w których kamera pokazuje zbliżenia lub plany średnie. Kolejną pożądaną cechą są napisy słowne na ekranie. Duże znaczenie mają ponadto popularne w danym momencie hasztagi. Co więcej, filmy, w których twórca zwraca się bezpośrednio do widza (co nazywane jest perspektywą drugiej osoby), również wykazują wyższy potencjał wiralowy (Cervi, 2021). Zapewne odrębna analiza obecnych w tych treściach klipów muzycznych również dostarczyłaby ciekawych wniosków, ale badacze nie uwzględnili tego w swojej metodyce.

Oczywiście niezależnych badań dotyczących ‘zaraźliwej’ muzyki jest wiele (np. Findeisen, 2017; Seabrook, 2015). Przede wszystkim jednak w tym kontekście należy przytoczyć literaturę na temat zjawiska earworms (Jakubowski, 2016), gdzie badacze statystycznie ustalili, jakie cechy muzyczne są skutecznym wskaźnikiem tego, że utwór muzyczny czy melodia ‘utknie nam w głowie’, że będziemy w pętli mimowolnie ją sobie wyobrażać po usłyszeniu. Pośród innych kryteriów zauważono m.in., że utwory takie zazwyczaj łączą ogólnie znane i proste kontury melodyczne z kilkoma nietypowymi interwałami, co sprawia, że melodia jest zarówno prosta do przetworzenia, jak i wystarczająco ciekawa, by przykuć uwagę. Warto podkreślić, że kryterium wyróżniania się na tle reszty można już rozumieć jako typowe kryterium selekcji memetycznej. Jan wspomina zresztą o aspekcie wyróżnialności w swojej koncepcji memetyki muzyki (Jan, 2007). Adekwatność do memetyki wynika stąd, że w tym przypadku to różnica pomiędzy produktami kultury w ramach jednej populacji staje się kryterium selekcyjnym, a nie tylko relacja produktu do człowieka. Tym niemniej memosceptyczni ewolucyjni kulturoznawcy mogą i ten mechanizm sprowadzić do tego, że o wszystkim decydują wyłącznie ludzkie preferencje i uprzedzenia.

W tak daleko idącym przypadku należy głębiej sformułować problem: w jaki sposób możemy jeszcze dobitniej wykazać, że memy istnieją, skoro ani ich spójne konceptualizacje, ani przykłady empiryczne nie są wystarczające? Najbardziej

przekonujące byłoby wykazanie ich istnienia w kategoriach formalnych, matematycznych, fizycznych. Oczywiście w niniejszej pracy nie podejmuję się zadania stworzenia matematycznego dowodu istnienia memów. Sformułuję za to argumenty, które konsekwentnie prowadzą mnie w tym kierunku. Formalna koncepcja memu opiera się na dwóch filarach: koncepcji samoreplikującej się informacji oraz algorytmie uniwersalnego darwinizmu. Uniwersalny darwinizm (UD), po raz pierwszy wprost nazwany przez Dawkinsa w 1983 (Dawkins, 1983) został koncepcyjnie zaproponowany przez tego samego autora już w 1976. Za Dawkinsem (1983), Dennettem (1995) i Blackmore (1999) UD możemy sformułować następująco:

Dobór naturalny to owoc działania uniwersalnego algorytmu, który może się realizować w dowolnym medium. Wariacja, selekcja i dziedziczenie w dowolnym substracie prowadzą do doboru naturalnego.

Susan Blackmore opisała to w ten sposób:

Argument Darwina wymaga trzech głównych cech: wariacji, selekcji oraz retencji (lub dziedziczenia). Właściwie, najpierw musi pojawić się wariacja, tak, żeby nie wszystkie stworzenia były identyczne. Po drugie, musi istnieć środowisko, w którym nie wszystkie stworzenia mogą przetrwać – niektóre radzą sobie lepiej niż inne. Po trzecie, musi istnieć proces, przy pomocy którego potomstwo dziedziczy cechy rodziców. Jeśli wszystkie te trzy warunki są spełnione, cechy użyteczne do przetrwania w tym środowisku muszą wzrastać.

Darwin's argument requires three main features: variation, selection and retention (or heredity). That is, first there must be variation so that not all creatures are identical. Second, there must be an environment in which not all the creatures can survive and some varieties do better than others. Third, there must be some process by which offspring inherit characteristics from their parents. If all these three are in place then any characteristics that are positively useful for survival in that environment must tend to increase (Blackmore 1999).

Blackmore jednak pisze o stworzeniach, co może wciąż sugerować, że algorytm ten dotyczy organizmów żywych. Za to w pismach Dawkinsa algorytm darwinowski opisywany jest w odniesieniu do replikatora, a sam replikator to „cokolwiek we

wszechświecie, co tworzy własne kopie” (1982). Zatem replikatorem jest także muzyka — zresztą Dawkins w swojej pierwotnej koncepcji (Dawkins, 1976) wymienia melodie dosłownie jako pierwszy przykład tego, czym w ogóle są memy. Uniwersalny Darwinizm ma bardzo donośne konsekwencje dla ewolucji kultury, a w tym ewolucji muzyki, ponieważ pozwala traktować ją nie jako metaforę, lecz jako rzeczywisty proces o strukturze algorytmu. Esencją doboru naturalnego jest bowiem mechanizm algorytmiczny. Skuteczność takiej perspektywy potwierdzają badania empiryczne (Aguera y Arcas, 2024). Nie jest to jednak tylko twierdzenie empiryczne: UD ma również wymiar formalny, ponieważ opiera się na regule, że jeśli istnieją jednostki różniące się cechami, które podlegają dziedziczeniu i wpływają na szanse przetrwania, to w populacji musi zachodzić proces selekcji. W tym sensie jest to twierdzenie tautologiczne, a zarazem, jak sugerują dane empiryczne, opisuje ono rzeczywiste zjawiska. UD stanowi więc uogólnioną formalizację oryginalnych spostrzeżeń Darwina, który zauważał działanie doboru także poza światem dzikiej przyrody – chociażby w sferze języka i słów.

Skoro memy to ewoluujące replikatory kulturowe, a sama ewolucja kulturowa może być dobrze opisana na sposób formalny, pozostaje jeszcze wykazać, że informacje kulturowe mogą się replikować – tworzyć własne kopie. Tutaj podparcie mogą stanowić te same badania, które wymieniłem wcześniej. Spontaniczne powstawanie replikującej się informacji zostało wykazane w matematycznych eksperymentach w środowisku komputerowym (Aguera y Arcas, 2024), co już mocno zbliża nas do formalnego dowodu ich istnienia. Należy jednak podkreślić, że są to dowody formalne, które dopiero w zestawieniu z empirycznymi badaniami zyskują walor opisu rzeczywistości. Takie abstrakcyjne dowodzenia na strukturach informacyjnych można także przenieść do świata materialnego dzięki zasadzie Landauera, która nierozzerwalnie wiąże informację z termodynamiką – skasowanie pojedynczego bitu informacji musi się wiązać z wytworzeniem ciepła (Landauer, 1961). Różnica między 0 a 1 w komputerowym kodzie sprawi, że zostanie wykonany inny program, a różnica między D# a E w muzycznym wykonaniu znanego dzieła wywoła natychmiastową reakcję w umysłach i zachowaniu publiczności. Informacja posiada niepodważalną fizyczną moc sprawczą.

Istnieje kilka poziomów replikacji. Operowanie w próżni przykładem memów jako samo-replikujących się informacji sprowadza się do jałowego filozoficznego sporu o lokalizację sprawczości – albo to sprawczy mem manipuluje człowiekiem, albo to człowiek manipuluje pasywnym produktem kultury. Rozważając ten problem w izolacji, łatwo można ulec prostemu antropocentryzmowi, w którym tylko człowiek, lub ewentualnie inne rozumne zwierzę, może wywierać sprawczy wpływ na świat zewnętrzny poprzez wolicjonalne działanie. Żeby wyjść z tego impasu, należy pokazać, że taka alternatywa (aktywny/pasywny człowiek) to ogromne uproszczenie, które nie oddaje rzeczywistości. Poniżej demonstruję w tym celu całe spektrum procesów replikacji różnych struktur informacyjnych. Te kontrprzykłady zostały uszeregowane w kolejności malejącej zależności od człowieka:

Łańcuszek pocztowy: wiadomość o treści pokroju „roześlij mnie do 10 osób, lub czeka cię 10 lat nieszczęścia”. Produkt kultury wprost zawiera instrukcję dla człowieka, aby go zreplikował. Oczywiście sceptycy opiszą ten przykład jako inteligentny plan i sprawczość człowieka, który eksploatuje psychologię innego człowieka.

Earworm: nachalny fragment muzyczny nieposiadający wprost wyrażonej instrukcji, aby ludzki odbiorca dokonał replikacji, a jednak poinstruowany człowiek tę replikację memu przeprowadza. Instrukcja jest zatem zawarta w samej treści fragmentu muzycznego – musi być on tak zaprojektowany, że wpasowuje się w mechanizmy poznania i zachowania człowieka jak klucz w zamek i swoją strukturą wywołuje replikację. Jego pochodzenie może być różne, ale nie ma to znaczenia dla jego sprawczości. Przykładowo w reklamach taki earworm jest świadomie zaprojektowany, skomponowany. Za to w kulturach oralnych, o czym więcej piszę w 3.2.1., earworm może także powstać bez żadnego projektanta, tylko w wyniku oddolnych niewolicjonalnych mutacji materiału muzycznego wynikających z charakterystyki ludzkiej niszy poznawczej. Empiryczny dowód takiej oddolnej ewolucji materiału muzycznego bez projektanta pokazuje również eksperyment DarwinTunes (Maccallum, 2012). Ale również i w tym przykładzie sceptyk odnajdzie tylko sprawczość człowieka: w

końcu system ewolucji muzyki DarwinTunes został zaprojektowany przez człowieka i to ludzie go obsługują. Za to earwormy w kulturze oralnej może i nie mieszczą się w sferze świadomych zachowań człowieka, ale to dalej człowiek kształtuje je i replikuje.

DNA: Mechanizmy replikacji DNA są niezależne od świadomego działania czy kontroli człowieka, choć wciąż odbywają się dzięki działaniu naszego ciała. Przez całe życie nasze organizmy replikują DNA, tworząc nowe komórki, żebyśmy mogli żyć (Alberts, 2002). Elementarna sprawczość leży tutaj po stronie każdej pojedynczej żywej komórki (Levin, 2022).

Wirusy: W powyższych przykładach (negując istnienie memów) sprawczość leżała po stronie żyjącego, biologicznego agenta. Sytuacja znacznie bardziej się komplikuje w przypadku wirusów. Ponieważ nadal nie istnieje naukowy konsensus odnośnie do tego, czym jest życie, nie można definitywnie powiedzieć, że wirusy są martwe, choć uważa tak wielu specjalistów (Moreira, 2009; Raoult, 2009). Koonin i Starokadomskyy twierdzą, że paradygmat replikatora wyjaśnia zachowanie wirusów: replikują się pomimo tego, że są martwe (Koonin, 2016). Zatem grypa, COVID czy HIV wykorzystują ludzkie ciała do tego, aby wytworzyć własne kopie, które następnie infekują inne organizmy i przekazują innym komórkom instrukcje, aby dokonywały replikacji wirusa – cykl trwa, replikacja się odbywa, sprawczość leży po stronie wirusa – pomimo tego, że cały czas jest on martwy. Ma to donośne znaczenie dla debaty o sprawczości memów. Dennett opisuje to w taki sposób:

Memy ewoluują tak jak zwierzęta, rośliny czy wirusy. Wirusy nie są żywe, a jednak skutecznie ewoluują poprzez dobór naturalny. Zatem nie trzeba żyć, żeby ewoluować na drodze doboru naturalnego, co jest dobrą informacją, bo słowa nie są żywe. A jednak ewoluują na drodze doboru naturalnego. Memy nie są żywe. A jednak podlegają doborowi naturalnemu. Mam swoje określenie na to, czym są wirusy: to kwas nukleinowy z charakterkiem. Oznacza to, że jakaś cecha ich kształtu daje im siłę i kompetencje do tego, żeby sprowokować własną replikację, kiedy dostaną się do wnętrza komórki.

Memes evolve just as animals and plants and viruses do. Viruses aren't alive, but they sure do evolve by natural selection. So you don't have to be alive to evolve by natural selection, which is a pretty good thing, because words aren't alive. But they evolved by natural selection. Memes are not alive. They are subject

to natural selection. I have a phrase for what viruses are. They're nucleic acid with attitude. That means something about their shape gives them the power, the competence to provoke their own replication when they get inside a cell. (Dennett, 2017)

Wirus komputerowy: wirusy komputerowe jeszcze do niedawna były projektowane wyłącznie przez człowieka. Określenie 'projektowane' wprowadza tu ważne rozróżnienie, ponieważ jeżeli wirus komputerowy tworzy własne kopie, to tylko ten pierwszy, oryginalny egzemplarz możemy uznać za 'stworzony' przez człowieka. Kolejne kopie są już tworzone przez własne poprzedniczki. A zatem człowiek mógł projektować wirusy, ale replikacji dokonywały same. Obecnie rola człowieka jeszcze znacznie bardziej umniejszona. Instytut Alana Turinga w 2024 wydał bogaty raport na temat Generatywnej AI oraz cyberbezpieczeństwa (Mercer, 2024). Wynika z niego, że wirusy projektowane przez AI stanowią już powszechne zagrożenie. W części 3 niniejszej pracy przedstawiam argumenty za tym, że różne technologie AI de facto stanowią procesy ewolucyjne. Zatem tutaj sprawczość człowieka sprowadza się do tego, że uruchamia na komputerze proces ewolucji, system ewoluuje tysiące konkurujących ze sobą i mutujących architektur, po czym informuje człowieka o odnalezieniu tego najbardziej skutecznego rozwiązania (Mercer, 2024). Tak samo zatem, jak John Cage w utworze 4'33" deleguje stworzenie treści utworu przypadkowi oraz publiczności, tak samo człowiek używający generatywnej AI do stworzenia wirusa deleguje stworzenie treści wirusa procesowi ewolucyjnemu. Wraz z rozwojem agencjalnego AI twórcza i sprawcza rola człowieka odchodzi i będzie odchodzić na coraz dalszy plan.

W naturze znajdziemy jeszcze więcej przykładów samoreplikujących się struktur, takich jak priony (Bieschke, 2004), zbiory autokatalityczne (Liu, 2020; Hordjik, 2010), czy też kryształy (Zervacic, 2014). W tej perspektywie, mając empiryczne weryfikacje replikatorów spontanicznie powstających w materii nieożywionej, materii ożywionej, przestrzeni informacyjnej, a także earwormów, sceptycyzm co do istnienia replikatorów w ogóle, a także muzycznych replikatorów wydaje się nieuprawniony.

Aunger za Janionem przedstawia drugi bardzo poważny zarzut: 'memy nie czynią żadnych predykcji i są nefalsyfikowalne' (Aunger, 2000, s.2).

Po pierwsze, w poetyckim akcie dobrej woli odrzucamy antropomorfizację dokonaną przez autora i rozwijamy jego skrót myślowy. Chodzi o to, że jego zdaniem memetyka nie czyni żadnych predykcji, a zatem tezy memetyki są nefalsyfikowalne, bo niesprawdzalne empirycznie.

Przede wszystkim cytowany zarzut pochodzi z 1999 roku, czyli czasu, kiedy autorzy artykułów w *Journal of Memetics* (JoM) (np. Boyd, 2001; Gatherer, 1997; Calvin, 1997) usiłowali stworzyć spójny zbiór pojęć, który miałby stać u podstaw memetyki. Nieustannie odnosili się do zarzutów filozoficznych, więc modelom predykcyjnym praktycznie nie poświęcano uwagi. Co ciekawe, jeden z głównych redaktorów oraz wydawca *Journal of Memetics* pisząc o tym, czemu pismo upadło, zdaje się w ogóle nie zauważać tego problemu filozofowania (Edmonds, 2005). Autor pisze za to, że jego zdaniem JoM upadł, ponieważ nie udało się dzięki analogii gen-mem zyskać żadnej dodatkowej mocy eksplanacyjnej w memetyce (Edmonds, 2005). Po pierwsze to nieprawda - dzięki słabej analogii gen-mem można wyciągnąć użyteczny poznawczo wniosek, że kultura zbudowana jest z replikatorów. Po drugie to zdanie reprezentuje szerszy pogląd obecny w artykułach z JoM – że memetyka musi być oparta na analogiach do genu. Autorzy nieustannie próbowali stworzyć czy też dostrzec mocne analogie, co wywołało mętlik pojęciowy i było drogą donikąd. Ale to nie jest jedyny możliwy kierunek badawczy dla memetyki, co rozwinę poniżej.

Zatem zarzut o brak predykcyjności w memetyce odnosi się do jej filozofującej, wczesnej gałęzi - która doprowadziła do upadku jedyne recenzowanego pisma memetycznego i zarazem rozpowszechnienia wieści o tym, że memetyka jest pseudonauką. Nic zatem dziwnego, że badacze, którzy podejmują de facto problemy memetyczne od tamtej pory do dziś, nie odwołują się do terminologii memetycznej. Nie zdążyła ona zdobyć znacznej widoczności (Edmonds, 2005) a poza tym prowadzi donikąd. Zatem predykcyjną naukę memetyczną musimy rozpoznawać pod innymi

nazwami, innymi hasłami, ale zachowując najważniejsze pytania badawcze – w jaki sposób przebiega ilościowo mierzalna ewolucja skwantyfikowanych produktów kultury? Jak produkty kultury zabiegają o własny byt? Na takie i pokrewne pytania odnajdziemy mnóstwo empirycznych odpowiedzi.

Co szerzej omówię w części trzeciej, obecnie sukces marki czy aplikacji jest w dużym stopniu zależny od tego, jak informacja na jego temat rozchodzi się w internecie – czy trafia do zainteresowanych klientów (Kaplan, 2011). To wiąże się z działaniem systemów rekomendacyjnych, które określam silnikami memetycznej replikacji. U podstaw mierzenia, przewidywania i napędzania ruchu w internecie leży nauka o dyfuzji informacji w sieciach społecznościowych. Dwa modele, na których oparte są tysiące innych prac, zostały sformułowane już w 2003. Kempe, Kleinberg i Tardos stworzyli model Niezależnych Kaskad (Independent Cascade) oraz Model Progu Liniowego (Linear Threshold Model) (Kempe, 2003). W pierwszym z nich każda jednostka, która już przyjęła daną informację, ma jednorazową szansę przekazania jej swoim sąsiadom w sieci z określonym prawdopodobieństwem. W drugim przyjmuje się, że każdy węzeł ulega wpływowi wielu sąsiadów i „aktywizuje się” dopiero wtedy, gdy łączna siła tych wpływów przekroczy indywidualny próg. Te sposoby formalizacji i mierzenia przepływu informacji w internecie są zarazem predykcyjnymi modelami ewolucji memetycznej. Przepływ informacji w internecie oznacza jej replikację.

Dziś modelowanie przepływu informacji i skuteczne predykcje zachowania użytkowników stoją u podstaw muzycznych systemów TikToka (Zhou, 2024) i Spotify (Verma, 2025). To wszystko sprowadza się do przewidywania replikacji informacji i metainformacji związanych z nagraniami utworów muzycznych. Predykcje dotyczące jasno skwantyfikowanych, podzielnych, policzalnych, replikujących się jednostek informacji to czysta empiryczna predykcyjna memetyka. Warto pójść jeszcze o krok dalej w określaniu naukowych nisz, które tak naprawdę są memetyczne, pomimo tego, że nie stosują tej nazwy. Ostatecznym i najważniejszym przykładem będzie pole algorytmów ewolucyjnych. Na sam początek należy wyraźnie podkreślić, że istnieje aktywna, płodna i przydatna tradycja naukowa badająca i wytwarzająca tzw. algorytmy

memetyczne. Recenzowane czasopismo naukowe Memetic computing w 2024 mogło pochwalić się pięcioletnim impact factorem na poziomie 2.6 (Ting, 2024). Artykuł określający cele i zakres przedmiotowy pisma (Springer, 2025) uwidoczni tendencję w informatyce i AI, żeby po prostu empirycznie pracować z memami i memetyką, zamiast zatrzymywać się na problemach filozoficznych. To doskonały przykład memetyki empirycznej. Spróbuję pokazać, że jest ich jeszcze więcej, choć kryją się pod innymi nazwami.

Wyszczególniona w informatyce kategoria algorytmów memetycznych (MA) to podzbiór szerszej kategorii algorytmów ewolucyjnych (EA). W istocie większość tych algorytmów należy do kategorii algorytmów genetycznych (GA). Wszystkie algorytmy ewolucyjne z założenia mają symulować procesy ewolucyjne na komputerze. Z perspektywy uniwersalnego darwinizmu, a zatem także memetyki, już tutaj należy wprowadzić terminologiczną poprawkę. Jeśli definiujemy ewolucję jako algorytm, to na komputerze można przeprowadzić jej emulację, a nie symulację. Symulacja to powierzchowne naśladowanie pewnych przejawów procesu, a emulacja to de facto przeniesienie i przeprowadzenie danego procesu w nowym środowisku.

W ramach dodatkowego wyjaśnienia mogę zaproponować porównanie oparte na przykładzie: aplikacja iBeer - Beer Simulator to symulator kufła piwa na smartfonie. Na ekranie wyświetla się pewien poziom napoju, który zmniejsza się wraz z przechyleniem telefonu. Za to aplikacja internetowa Metronome Online nie jest symulacją metronomu, tylko jego emulacją. Piwo składa się z cząsteczek chemicznych, które są zaledwie reprezentowane na telefonie. Metronom za to jest systemem nadającym stałe tempo przy pomocy komunikatów dźwiękowych. Metronom na telefonie nie jest symulowanym metronomem, tylko prawdziwym metronomem. Podobnie kalkulator, który fundamentalnie jest systemem przetwarzania informacji. Kalkulator na telefonie to nie symulowany kalkulator, ale prawdziwy. Tak samo ewolucja jest systemem przetwarzania informacji, który nie staje się mniej prawdziwy, jeśli zachodzi na telefonie czy w komputerze. Ewolucja, która zachodzi w komputerze to prawdziwa ewolucja - a jej przedmiotem są informacje kulturowe.

Informatyczne pojęcie algorytmów genetycznych odnosi się do takich procesów ewolucji informacji, w których ta informacja nie jest obdarzona dziedzicznością cech nabytych (Krasnogor, 2005). Zachodzą tylko losowe mutacje jednostek informacji i dziedziczenie tych mutacji. Za to algorytmy memetyczne stanowią takie algorytmy genetyczne, których funkcjonowanie jest poszerzone o mechanizm dziedziczenia lamarckowskiego (Ong, 2004) – czyli właśnie cech nabytych. Taki podział należy z punktu widzenia nauk ewolucyjnych określić jako bardzo okrojone inżynierskie uproszczenie, wymodelowanie w celu uzyskania algorytmicznej jasności i skuteczności. Ale, co dla nas jest najistotniejsze, zarówno algorytmy genetyczne (GA) jak i algorytmy memetyczne (MA) operują na informacjach kulturowych, czyli w tym przypadku danych wprowadzonych do systemu przez człowieka i w znakomitej większości dotyczących ludzkiej działalności. Zatem wszystkie algorytmy ewolucyjne działające w sferze technologii cyfrowych w świetle uniwersalnego darwinizmu są algorytmami memetycznej ewolucji.

Na takiej właśnie memetycznej, emulowanej w technologiach cyfrowych ewolucji opierają się w dużej mierze najważniejsze osiągnięcia w dziedzinie AI. W 2016 AlphaGo firmy DeepMind pokonał jednego z najlepszych graczy na świecie w grze w Go, Lee Sedola 4:1 (BBC, 2016). Go uchodziło za najbardziej skomplikowaną grę planszową na świecie i tego typu przełom nadszedł znacznie wcześniej, niż oczekiwali eksperci. Stało się to dzięki jego unikatowej architekturze – AlphaGo najpierw analizował 30 milionów ludzkich ruchów w Go, a następnie rozgrywał mecze przeciwko samemu sobie, ucząc się od siebie przez wzmocnienie (Silver, 2016). Podobieństwo ewolucji do uczenia się przez wzmocnienie jest tak uderzające, że w niektórych przypadkach można je ze sobą zrównać (Frank, 2014; Whiteson, 2006). W obu procesach co do zasady odbywa się ślepe poszukiwanie optymalnych rozwiązań poprzez konkurencyjne uczenie się na próbach i błędach. Całkowitym ucieleśnieniem tej strategii była młodsza siostra AlphaGo czyli AlphaGo Zero. Zero miała zerową wiedzę na temat tego, jak grać w Go i jakie są ludzkie strategie. Po prostu od zera grała przeciwko samej sobie, uczyła się od siebie i próbowała wygrać. Po trzech dniach takiego treningu wygrała ze swoją wyposażoną w ekspercką wiedzę starszą siostrą AlphaGo sto na sto gier, tzn

zakończyły mecz z wynikiem 100:0 (Silver, 2017). Gra w Go ma 2500 lat, a sztuczne sieci neuronowe uczące się przez wzmocnienie - czyli poddane cyfrowej memetycznej ewolucji - przerosły umiejętnościami człowieka w trzy dni. Sukces AlphaGo umożliwił także stworzenie odmiennego, ale pokrewnego systemu AlphaFold, do przewidywania trójwymiarowej struktury białek. W uznaniu tego spektakularnego sukcesu predykcyjnego szefowie projektu w 2024 otrzymali nagrodę Nobla w dziedzinie chemii (Nobel Prize, 2024). Wskazane powyżej argumenty przekonują, że memy istnieją i że memetyka ma moc predykcyjną.

We wspomianej na początku rozdziału pracy zbiorowej (Aunger, 2000) znajdziemy również wiele uwag krytycznych dotyczących memetyki, które w porównaniu do tych wyżej omawianych możemy potraktować jako kosmetyczne i odeprzeć je w sposób skrócony. Pierwsza grupa takich zarzutów dotyczy ontologii i definicji memu. Kilka z nich można zebrać w grupę pod ogólnym tytułem zarzutów o redukcjonizm w sensie psychologicznym. Rosaria Conte (Conte, 2000) argumentuje, że memetyka błędnie postrzega ludzi jako wektory transmisji memów, a nie aktywnych agentów (s. 87). Henry Plotkin (Plotkin, 2000) uważa, że memetyka jest naiwna, kiedy ignoruje złożone procesy psychologiczne (s. 70). Dodatkowo Kevin Laland i John Odling-Smee (Laland, 2000) piszą o tym, że memetyka, opisując zaraźliwe memy, zaniedbuje miarę podatności gospodarza, a skupia jedynie na zaraźliwości memu (s. 134). Wszystkie te zarzuty oceniam jako konstruktywne, słuszne i w pełni się z nimi zgadzam. Na szczęście implementacja złożoności, sprawczości oraz podatności człowieka wspaniale wzbogacają memetykę. Zatem nie są to argumenty przeciwko memetyce, a raczej świetne pomysły na jej rozwój.

Adam Kuper (s. 185) oraz David Hull (s. 55) uważają, że analogia gen-mem jest nadużywana. Z tym zarzutem także całkowicie się zgadzam, szczególnie w odniesieniu do filozoficznej memetyki rodem z JoM. W memetyce obliczeniowej tego typu zarzuty już nie sprawiają problemów. Antropolog Maurice Bloch ubolewa, że memetyka ignoruje dokonania antropologii w dziedzinie badań nad kulturą (s. 190). Tu pozostaje wyrazić

nadzieję, że w przyszłych modelach uda się zaimplementować jak najbogatszą wiedzę z dziedziny antropologii.

Przytoczyłem i odpowiedziałem na szereg zarzutów wobec memetyki. Oczywiście literatura krytykująca memetykę jest szersza, a zarzutów można znaleźć więcej. Tym niemniej przytoczyłem te moim zdaniem najważniejsze i starałem się wykazać, że nawet te najcięższe zarzuty wcale nie dyskredytują memetyki. Co więcej, problematyka memetyczna, chociaż określana przy pomocy innych nazw, z powodzeniem opracowywana jest w rozmaitych naukach predykcyjnych.

2.4. Hipotezy pomocnicze

Jednym z najpoważniejszych zarzutów wobec memetyki jest jej postrzegana izolacja i trudności w redukowaniu jej twierdzeń do ugruntowanych ram naukowych, takich jak biologia ewolucyjna, psychologia czy nauki kognitywne (Aunger, 2000). Częściowo przyznają temu rację w odniesieniu do memetyki prezentowanej na łamach *Journal of Memetics*. W dużej mierze dlatego, że autorzy koncentrowali się na abstrakcyjnych i często nierozstrzygalnych sporach filozoficznych (Boyd, 2001), zamiast na systematycznym budowaniu pomostów z naukami empirycznymi, co utrzymywało izolację dyscypliny. Ponadto, wczesne wysiłki ustanowienia memetyki jako całkowicie autonomicznej dziedziny sprzyjały zamykaniu się na integrację z innymi obszarami wiedzy. Problemem nie była sama idea niezależności (biologia czy lingwistyka też funkcjonują jako samodzielne nauki), lecz fakt, że w przypadku memetyki często brakowało wysiłku w kierunku redukcji lub konstruktywnego dialogu interdyscyplinarnego.

Dziś sytuacja wygląda zupełnie inaczej i nowoczesna memetyka może współpracować z różnorodnymi naukami. W tym rozdziale przedstawię zbiór różnych teorii, hipotez, paradygmatów, programów badawczych spoza memetyki, które moim zdaniem są z nią głęboko spójne, zwiększają jej moce eksplanacyjne i wspólnie są użyteczne poznawczo. Niektóre z nich w praktyce rozwijam w trzeciej części tej pracy. Najszerzej rozwijam te, które uznałem za najbardziej obiecujące. Uważam, że integracja z innymi dyscyplinami nie tylko wzmacnia użyteczność poznawczą memetyki, ale także przyczynia się do budowania spójnego, naturalistycznego obrazu procesów kulturowych.

Najbardziej fundamentalnym punktem odniesienia jest teoria informacji Claude'a Shannona (1948). Proponowany przez Shannona i Weaver'a schemat komunikacji, który tak przysłużył się do rozwoju zarówno telekomunikacji jak i informatyki, stanowi także podstawowy model mechanizmu memetycznej mutacji. To także bardzo ważne w

memetyce muzyki, szczególnie dotyczącej kultur oralnych, gdzie transmisja informacji przebiega w bezpośredniej dźwiękowej komunikacji. Shannon w pracy Matematyczna Teoria Komunikacji (1948) poszukiwał takich elementów komunikacji, w których może pojawić się zaburzenie, nieścisłość, szum. Podzielił on akt komunikacji na kilka odrębnych kategorii: źródło informacji, wiadomość nadana, przekaźnik, sygnał wysłany, źródło szumu, sygnał odebrany, odbiorca, wiadomość odebrana, adresat.

W sytuacji zachodnie muzyki artystycznej można ten schemat opisać na następującym przykładzie: **źródłem informacji** jest kompozytor, który zaczyna z pewną dźwiękową intencją. Następnie przekłada ją na papier nutowy, z wszelkimi jego ograniczeniami. Nie tylko samo medium nutowe jest niedoskonałe, ale również sama zdolność kompozytora do przenoszenia abstrakcyjnego wyobrażenia na konkretny system notacji. Taka zapisana partytura stanowi **wiadomość nadaną**. Następnie uruchamia się **przekaźnik** – wykonawca muzyczny. Wykonawca z konieczności nadaje dziełu, choćby szczątkowo, autorską interpretację (Cook, 2001), i tak powstaje **sygnał wysłany**. Kiedy muzyk wykonuje dzieło, nieuchronnie mierzy się ze złożonym **źródłem szumu**: szumem informacyjnym oraz szumem akustycznym. Składają się na niego wszelkie okoliczności towarzyszące wykonaniu, takie jak (potencjalnie niedoskonały) program koncertu, oświetlenie, temperatura, zapachy towarzyszące wykonaniu, ale też, przede wszystkim faktyczny szum akustyczny. Poza wykonawcą na sali może pracować klimatyzator, projektor, szumiący wzmacniacz. Również i nagłośnienie zawsze ma jakąś charakterystykę, nie jest obiektywne i neutralne. Poza tym dźwięki wydaje publiczność: krząta się, chrząka, kaszle. Te oraz inne szумы wchodzą w interakcję z sygnałem wysłanym i tworzą **sygnał odebrany**. To, że pewna fala akustyczna dotarła do słuchacza, nie oznacza jeszcze, że zakończył się odbiór muzyki. Wręcz przeciwnie, teraz dopiero się on rozpoczyna – samo ucho oraz umysł odbiorcy musi aktywnie i kreatywnie przetworzyć bodziec dźwiękowy, który dopiero w umyśle stanie się muzyką. Ucho każdego odbiorcy jest nieco inne – mamy m.in. inne zakresy słyszenia (Jacewicz, 2023). Po akcie percepcji zaczyna się kształtowanie znaczenia [m.in.](#) w oparciu o dotychczasowe doświadczenia, skojarzenia, preferencje (Juslin, 2008). Poza tym trudno, aby umysł słuchacza był w pełni skoncentrowany na doświadczaniu muzyki –

doświadczenie muzyczne jest kształtowane przez takie czynniki jak bieżący nastrój czy osobowość odbiorcy (Vuoskoski, 2011). Te wszystkie czynniki razem kształtują **wiadomość odebraną**, która trafia do świadomego umysłu **adresata** (to uproszczony opis - muzykę odbieramy także poprzez układ przedsionkowy oraz mechanoreceptory w skórze, np. ciała Paciniego, które reagują na wibracje niskich częstotliwości (Phillips-Silver, 2008; Trainor, 2009; Merchel, 2020)). Taka przykładowa droga utworu muzycznego z umysłu kompozytora do świadomego umysłu odbiorcy przekonująco pokazuje, że idealnie niezakłócony akt komunikacji muzycznej praktycznie nie zachodzi. Za to każda zmiana potencjalnie może się stać dziedziczną memetycznie mutacją komunikowanej muzycznej informacji.

Model komunikacji to jedno z dwóch narzędzi autorstwa Shannona, które istotnie wpłynęły na naukę opisującą zagadnienia pokrewne memetyce. Shannonowska koncepcja informacji fundamentalnie odnosi się do przewidywania i zaskoczenia, pomijając przy tym znaczenie komunikatu. Im mniej prawdopodobne zdarzenie, tym większą ilość informacji wnosi jego pojawienie się. Informacja jest miarą redukcji niepewności (Shannon, 1948). Ta koncepcja już w latach 50. XX wieku pomogła ukształtować pierwszą teorię percepcji muzyki opartą na oczekiwaniach – a jej autorem był Leonard B. Meyer (Meyer, 1956). Znacznie odświeżoną, wzbogaconą psychologicznie wersję tej hipotezy przedstawił w 2006 David Huron (Huron, 2006). Jeszcze bardziej poszerzony model empirycznego badania muzycznych oczekiwań stworzył Marcus Pearce (Pearce, 2005, 2010, 2018; Cheung, 2019) i w jego przypadku znacząca rola przewidywań i oczekiwań w percepcji muzyki została nawet wykazana w empirycznych badaniach z użyciem rezonansu magnetycznego fMRI.

W tych muzykologicznych pracach dotyczących oczekiwań nie ma jednak wzmianek o memach czy ewolucji kulturowej muzyki, ale można zauważyć pewne istotne logiczne koneksje. Meyer (1956) z rozważań dotyczących muzycznych oczekiwań i niespodzianek wychodzi do rozważań nad bardziej uogólnionym muzycznym stylem. Meyer zauważa, że style muzyczne podlegają nieustannej zmienności (s. 64). Ponadto pisze, że style posiadają swoje normy oraz odstępstwa, które tworzą auto-degradujący

system prawdopodobieństw (s.72). Poza tym, co być może najważniejsze, pisze o tym, że styl nie wynika tylko z inwencji kompozytorów, ale także grup kompozytorów, oraz norm i odstępstw, które dziedziczą po poprzednikach i z chęcią modyfikują (s.69).

Czyli u Meyera obok sprawczego kompozytora twórczą rolę w kształtowaniu stylu muzycznego pełni także grupa o rozproszonej sprawczości, a także współdziedziczony zbiór muzycznych elementów, powiązanych wzajemnie logiką prawdopodobieństw występowania, nieustannie podlegający zmienności. Taki pogląd na rozproszoną sprawczość jest całkowicie spójny z memetyką muzyki, a owe relacje prawdopodobieństw to doskonały sposób na ukazanie sprawczej siły zastanej puli memetycznej.

Huron i Pearce piszą z perspektywy psychologii muzyki, a zatem sprawczość dostrzegają wyłącznie po stronie kompozytora, a w ich pismach, choć mogłoby, nie pojawia się hasło ewolucji kulturowej muzyki. Za to w artykule z 2018 Pearce opisuje to, że różne kultury wytwarzają różne współdzielone probabilistyczne modele muzycznych oczekiwań. Pisze również o tym, że jest to związane z uczeniem statystycznym i stanowi obliczeniowy model enkulturacji (Pearce, 2018). Autor explicite tego nie wskazuje, ale tę koncepcję idealnie dopełnia mechanizm ewolucji kulturowej. Przecież nabywane przez enkulturację wspólne chwilowo dla danej grupy modele oczekiwań nie mogą być dziedziczone genetycznie. Są przekazywane przez imitację, ciągle podlegając grupowej negocjacji, zmienności – czyli są dziedziczone memetycznie.

Kolejnym, odrębnym narzędziem opartym na teorii Shannona, które może znaleźć zastosowanie w memetyce muzyki, jest zasada wolnej energii (*Free energy principle, FEP*) Karla Fristona (2010). Zasada wolnej energii to ogólna teoria wynosząca shannonowskie informacyjne oczekiwania i zaskoczenia na znacznie wyższy poziom abstrakcji niż tylko komunikacja. FEP zakłada, że optymalną strategią przetrwania systemów poznawczych jest minimalizacja wolnej energii, czyli unikanie zaskoczenia. To znaczy, że pożądanym stanem jest pełna zgodność między przewidywaniami a napływającymi danymi sensorycznymi (Friston, 2010; Friston, 2006). System

poznawczy to tutaj nie tylko żywy organizm, ale także grupa ludzi, instytucja lub sztuczny autonomiczny system taki jak AI (Clark, 2013; Constant, 2022, Veissière, 2023). W kulturze muzycznej system poznawczy stanowi na przykład orkiestra symfoniczna wraz z partyturami i dyrygentem. Całość działa jako system kolektywnego przetwarzania informacji, wnioskowania, w którym pozornie niezależne jednostki wspólnie minimalizują niepewność poprzez łączony generatywny model rzeczywistości. Jeszcze bardziej radykalny przypadek to systemy poznawcze oparte wyłącznie na produktach kultury, takie jak algorytm rekomendacyjny Spotify, który sam w złożony sposób przetwarza informacje i kształtuje oczekiwania kolejnych użytkowników (Veissière, 2019).

W tym miejscu należy dodać, że proces minimalizacji wolnej energii można ująć jako cykl formułowania różnych hipotez na temat otoczenia, testowania ich skuteczności oraz zachowywania i rozwijania tych, które się sprawdziły. Taki proces jest matematycznie tożsamy z procesem ewolucyjnym w rozumieniu uniwersalnego darwinizmu (Campbell, 2016). Czyli – układ poznawczy, który minimalizuje wolną energię, opiera swoje funkcjonowanie na nieustannie ewoluującym modelu świata zewnętrznego. A zatem na podstawie FEP można skonstruować przekonujący model memetycznej ewolucji kultury, ponieważ w języku matematycznych formalizmów FEP ilościowo opisuje ewolucję konkretnych układów czy też przedmiotów kultury.

Poza tym, pozornie całkowicie spójna z FEP jest koncepcja informacjożerów (infovores) (Biedermann, 2006). Autorzy twierdzą, że człowiek posiada wrodzoną predyspozycję do poszukiwania nowych informacji i że ich postrzeganie pobudza układ nagrody. Mechanizm ten opiera się na działaniu receptorów μ -opiodowych w obszarach asocjacyjnych mózgu: nowe i bogato interpretowalne bodźce powodują uwalnianie endomorfina, co wywołuje odczucie przyjemności (Biedermann, 2006). Badacze argumentują, że wyszukiwanie nowych informacji było ewolucyjnie skuteczną strategią i dlatego zostało fizjologicznie utrwalone - za dostrzeganie nowych informacji otrzymujemy chemiczną nagrodę w postaci pobudzenia wydzielania opiodów. Wykazują to szczególnie na podstawie bodźców wizualnych, ale po analizie literatury

(Goldstein, 1980) sugerują, że podobny mechanizm powinien działać w domenie dźwiękowej, ponieważ muzyczny efekt ciarek nie pojawił się u badanych po podaniu środka blokującego receptory opioidowe (Goldstein, 1980). To, że subiektywnie przyjemna muzyka pobudza u słuchaczy uwalnianie opioidów, potwierdzają także niedawne badania empiryczne (Putkinen, 2025). Jednocześnie w przyszłości w kwestii percepcji muzyki owocne może być porównanie założeń informacjożerności i pearcowskiej koncepcji muzycznego zaskoczenia. Z badań Pearce'a wynika, że pewną optymalną percepcyjnie strukturą jest charakterystyczny punkt równowagi pomiędzy przewidywalnością i niespodzianką (Pearce, 2010). Za to według koncepcji informacjożerności to informacyjna nowość, czyli niespodzianka, jest przyczyną uwalniania opioidów. Prawdopodobnie są to dwa odrębne mechanizmy, których praca wzajemnie się nie zakłóca. Muzyczna (percepcyjna) nowość pobudza ośrodek nagrody, a równowaga bodźców nowych i przewidywalnych pomaga przede wszystkim utrzymać uwagę. Badania fMRI wskazują, że równowaga bodźców nowych i przewidywalnych angażuje różne struktury mózgowie: ciało migdałowate, hipokamp, korę słuchową oraz jądro półleżące, które odpowiadają odpowiednio za emocjonalne i poznawcze aspekty nagradzającego doświadczenia muzyki (Cheung, 2019).

Koncepcja FEP mówi, że wyszukiwanie i nieustanne korygowanie stanu wiedzy na temat środowiska jest strategią optymalną. Koncepcja człowieka jako informacjożercy dodaje, że skłonność poszukiwania nowych informacji jest ewolucyjnie utrwalona w naszych ciałach, w tym mózгах. Stąd wyprowadzam wniosek, że poszukiwanie nowinek jest dla człowieka naturalne i intuicyjne. Można to powiązać z problematyczną konsumpcją memów, TikToków, shortsów, czy mediów społecznościowych. W popularyzatorskich artykułach często widnieje slogan, że nadmierna konsumpcja treści na smartfonie oznacza uzależnienie od dopaminy (Kuc, 2023; Multi.life, 2025, Terapia nad wartą, 2025). W rzeczywistości sytuacja jest bardziej skomplikowana: dopamina nie jest prostym hormonem przyjemności, lecz elementem działania złożonego układu nagrody (Schultz, 2016). Subiektywna przyjemność z konsumpcji treści w mediach społecznościowych nie jest wyłącznie efektem dopaminy, wynika także z działania układu opioidowego, podobnie jak w przypadku doznań muzycznych (Tereshchenko,

2023). Dopamina odpowiada przede wszystkim za proces wzmacniania i utrwalania nawyku sięgania po kolejną porcję treści (Schultz, 2016), natomiast opioidy mogą być związane z poczuciem nagrody i redukcją napięcia (Tereshchenko, 2023). A zatem najprawdopodobniej przesadna konsumpcja mediów społecznościowych wynika ze wspólnego działania dwóch systemów: dopaminowego, który utrwała cykl nawykowego sprawdzania, oraz opioidowego, który nagradza nowość i interpretowalność bodźców (De, 2025). Te właśnie dopaminowe i opioidowe układy są eksploatowane przez zastosowanie inteligentnych algorytmów rekomendacyjnych, które dzięki swojej ewolucyjnej naturze - i memetycznej sprawczości - skutecznie pobudzają ludzkie skłonności do poszukiwania informacji (Lin, 2021).

W memetyce Daniela Dennetta człowiek nie składa się tylko z genów. Jest złożonym bytem hybrydycznym, którego tożsamość i zdolności poznawcze kształtują zarówno geny jak i memy (Dennett, 1995). Doskonałą ramą do opisu takiego układu jest koncepcja poznania 4E (framework 4E cognition), również szeroko rozwijany na gruncie muzyki. 4E oznacza, że poznanie jest ucieleśnione (embodied), usytuowane (embedded), rozszerzone (extended) i enaktywnie realizowane (enactive) (Newen, 2018). Kompozytor oraz tablet wyświetlający nuty (Magnusson, 2019) czy też pianista i fortepian (Leman, 2007), wspólnie tworzą złożony system poznawczy. Ponadto, sama obecność w studiu nagrań zmienia myślenie muzyczne (Sawyer, 2006) i jest ono dodatkowo dynamicznie kształtowane w akcie wspólnego działania takim jak jam session (Krueger, 2014). Łączenie 4E cognition z memetyką stanowi pewne wyzwanie, ponieważ może zacierać się granica pomiędzy tym, które memy dokonują replikacji, a które memy oraz geny stanowią dla nich środowisko. Tym niemniej, w niniejszej pracy tego typu dylematy proponuję zawsze rozwiązywać doraźnie, to znaczy przy konkretnym badaniu empirycznym podkreślać, gdzie zaczynają i kończą się analizowane memy i mempleksy. Na przykład, w badaniu ewolucji fortepianu można skupić się na samym materialnym instrumencie, a nie na całym układzie obejmującym instrument oraz człowieka. Wtedy człowiek stanowi środowisko.

Zarówno 4E jak i FEP doskonale dopełnia teoria konstruowania niszy ekologicznej sformułowana przez Johna Odling-Smee i współpracowników (Odling-Smee, 1988, 1996, 2003). Teoria ta zakłada, że ewolucja nie zachodzi w próżni – ewoluujący organizm wpływa na swoje środowisko, kształtuje je, i ono ewoluuje wraz z nim, kształtując sprzężenie zwrotne (Laland, 2016). Jako przykład takiej koewolucji człowieka i jego niszy autorzy wskazują niebywale szybką ewolucję genetyczną w kierunku tolerancji laktozy u człowieka, co wynikało z kulturowego wynalazku hodowli bydła (Gerbault, 2011). Zatem ewolucja kulturowa niszy wytworzyła szczególną presję selekcyjną na ewolucję genetyczną. Autorzy mówią również o innych aspektach takiego koewolucyjnego procesu:

Zatem konstruowanie niszy przez człowieka, częściowo zależne od memów przekazywanych społecznie, nie tylko częściowo kształtuje środowisko selekcyjne ludzkich genów, ale także środowisko selekcyjne memów. Ludzka kultura materialna w postaci narzędzi, artefaktów czy domów może dosłownie być przekazywana z pokolenia na pokolenie - to jeden z aspektów ekologicznego dziedziczenia w naszym gatunku.

Thus, human niche construction, partly dependent on socially transmitted memes, not only partly shapes the selective environment of human genes, but also the selection environment of memes. Human material culture, in the form of tools, artifacts, and homes, may literally be transmitted from one generation to the next, as one aspect of the ecological inheritance of our species. – Laland, 2000

Ten, zdaje się, zbyt oczywisty aspekt może faktycznie umknąć uwadze badaczy ludzkiej ewolucji – dziedziczenie obiektów i dziedzictwa materialnego to także element procesu genetyczno-memetycznej koewolucji. Kiedy na przykład rodzic przekazuje dziecku swój wysoko ceniony instrument, aktywnie i głęboko wpływa na rozwój przyszłej kariery muzycznej dziecka. Autorzy piszą również o tym, że w ewolucji memów jedne memy stanowią środowisko dla innych. Wyjaśniają to w ten sposób, że porównanie memu do wirusa powinna dopełniać także koncepcja odporności (Laland, 2000). Sugerują, że sukces replikacyjny memu zależy od jego zaraźliwości, ale także od podatności gospodarza oraz warunków środowiska społecznego (Laland, 2000).

O sprawczej sile informacji pisał także fizyk David Deutsch, który to w swoim modelu wprost integruje pojęcie memu. Informacja nie jest dla niego ani formalną abstrakcją, ani Shannonowską redukcją niepewności, ale realnie oddziałującym bytem, który przekształca świat zgodnie z prawami natury: to konstruktorowa teoria informacji (Deutsch, 2015). Przede wszystkim Deutsch pisze o wiedzy, którą definiuje jako informacje trafnie odzwierciedlające, modelujące jakiś odrębny system (Deutsch, 2011). Ujmuje to także w inny sposób:

Ludzkie mózgi i cząsteczki DNA mają wiele funkcji, ale między innymi są to wszechstronne nośniki informacji: co do zasady mogą przechowywać każdy rodzaj informacji. Co więcej, zarówno mózgi jak i DNA wyewoluowały, aby przechowywać konkretny typ informacji, i oba te typy łączy pewna własność o kosmicznej istotności. Mianowicie, kiedy zostaną one fizycznie ucieleśnione w odpowiednim środowisku, zwykle sprawiają, że tak pozostaje. To wysoce nieprawdopodobne, aby takie informacje – które nazywam wiedzą – powstały w inny sposób niż dzięki procesom korekcji błędów obecnym w ewolucji lub myśleniu. (...) oba te rodzaje wiedzy oparte są na wystarczająco podobnej logice, aby teoria ewolucji w znaczącym stopniu odnosiła się do wiedzy ludzkiej.

Human brains and DNA molecules each have many functions, but among other things they are general-purpose information-storage media: they are in principle capable of storing any kind of information. Moreover, the two types of information that they respectively evolved to store have a property of cosmic significance in common: once they are physically embodied in a suitable environment, they tend to cause themselves to remain so. Such information – which I call knowledge – is very unlikely to come into existence other than through the error-correcting processes of evolution or thought. (...) the two kinds of knowledge share enough of their underlying logic for the theory of evolution to be highly relevant to human knowledge. - Deutsch, 2011

Deutsch tworzy zatem jeszcze inną definicję – jego sprawcza ‘wiedza’ to również takie informacje, które raz ucieleśnione, oddziałują na otoczenie w taki sposób, że ten stan się utrzymuje. Dostrzega też on w tej dynamice ważną rolę ewolucji. Innymi słowy, w systemie Deutscha, wiedza prowokuje własną replikację, czyli jest replikatorem. Nie jest to nadinterpretacja – Deutsch wprost odnosi się do uniwersalnego darwinizmu (Deutsch, 2011) i rozwija własne aspekty memetyki.

U Deutscha skuteczny replikator zawiera jakiegoś rodzaju prawdziwe informacje, czy też adekwatnie odnoszące się do jakichś zewnętrznych elementów środowiska i ta właśnie synergia prowadzi do tworzenia kopii. Według badacza wiedza – czyli także replikator – składa się z takich informacji, które zawierają w sobie funkcjonalny model czegoś zewnętrznego (Deutsch, 2005). Idąc dalej, fizyk tworzy kategorie memów racjonalnych i memów nieracjonalnych (Deutsch, 2011), z których to tym racjonalnym przypisuje pewną wyższość, ponieważ ich treść trafniej modeluje rzeczywistość (Deutsch, 2011). Racjonalne memy są skuteczne, użyteczne, podlegają lamarckowskiej ewolucji na drodze krytyki, za to nieracjonalne opierają się krytyce, wiążą się z przesądami i sztucznie podtrzymywanymi konwencjami (Deutsch, 2011). To rozróżnienie świetnie się sprawdza do opisu ewolucji wiedzy naukowej, ale wytwarza nietrywialne problemy przy próbie opisu świata muzyki. Spróbuję jednak je rozwiązać. Współpracowniczka Deutscha, współautorka koncepcji teorii konstruktorów, Chiara Marletto, w swoim artykule wprost podaje przykład zbioru nut muzycznych (Marletto, 2015). Wnioskuje więc, że racjonalne memy muzyczne mogą zatem tworzyć teorię muzyki opartą na metodzie naukowej, za to nieracjonalne memy muzyczne składają się na muzyczną teorię opartą na fantazyjnej spekulacji. Ale sama brzmiąca muzyka, rozumiana w uproszczeniu jako pewien rodzaj bodźca zmysłowego, nie posiada żadnego oczywistego kryterium racjonalności. Wnioskuje za to, że w memetyce Deutscha brzmiąca muzyka wtedy stanowi mem racjonalny, a zarazem skuteczny replikator, kiedy zawiera przydatną wiedzę na temat swojego środowiska. To znaczy, jej struktura w odpowiedni sposób odzwierciedla, dopasowuje się, a zatem i modeluje, istotne do sprowokowania replikacji elementy środowiska zewnętrznego, jak na przykład ewolucyjnie ugruntowane preferencje estetyczne (Bowling, 2015).

W kwestii fizyczności informacji dodatkowym narzędziem eksplanacyjnym jest zasada Landauera (1961), która głosi, że usunięcie bitu informacji musi wiązać się z wytworzeniem ciepła. Zatem informacja podlega zasadom termodynamiki. Interesującym polem rozwoju dla memetyki może być także biologia poznawcza (Lyon, 2025). Zgodnie z biologią poznawczą wszelkie struktury biologiczne są nośnikami pewnej powstałej na drodze ewolucji wiedzy na temat tego, jak optymalnie zbudować

organizm żywy (Kovač, 2000). Ponadto, w tym nurcie rozwinęła się także koncepcja poznania bazowego (basal cognition), według której właściwości poznawcze takie jak uczenie się, pamięć czy podejmowanie decyzji nie są zarezerwowane wyłącznie dla zwierząt z mózgiem, lecz występują również u organizmów jednokomórkowych, roślin i tkanek (Levin, 2022). Jeżeli w tak prostych strukturach istnieją formy uczenia się i pamięci, a jednocześnie zgodnie z FEP minimalizują one wolną energię (Friston, 2010), to możliwe, że już w takiej mikro-skali zachodzą procesy ewolucji nabytych informacji analogiczne do ewolucji memetycznej i są one w tej uproszczonej formie znacznie ewolucyjnie starsze niż mózgi.

Ponadto, ewolucją informacji jeszcze przed początkiem życia zajmuje się teoria złożzeń (assembly theory) (Sharma, 2022). Proponuje ona, że każdy posiadający kopie obiekt we wszechświecie powyżej pewnego określonego poziomu złożoności musiał powstać na drodze procesów selekcji. Wynika to stąd, że osiągnięcie wysokiego poziomu złożoności na drodze czysto losowych procesów jest w praktyce skrajnie nieprawdopodobne. W świecie o odpowiednio ograniczonym czasie i zasobach nawet nieskończona liczba małych stukających w maszyny do pisania nie odtworzyłaby kompletu dzieł Szekspira (Marshall, 2022). Analiza w duchu teorii złożzeń prowadzi do rozbioru każdego obiektu na pomniejszych, bardziej pierwotne elementy, żeby wykazać, jaki realny proces był niezbędny do powstania danego obiektu (Walker, 2024). Tego typu analiza może być bardzo przydatna do opisu ewolucyjnej drogi złożonych memów, w tym memów internetowych. Współczesne memy internetowe często funkcjonują jako złożone struktury znaczeniowe, powstające na drodze rekombinacji licznych odniesień kulturowych, stylistycznych i afektywnych. Ich sukces komunikacyjny zależy nie tylko od pojedynczego aktu transmisji, lecz od synergii warstwowego montażu i zapośredniczenia.

Acerbi zauważa (2020), że cyfrowe środowisko medialne sprzyja właśnie takim formom kultury: złożonym, powielanym z wysoką wiernością, umożliwiającym przekształcania. Połączenie teorii złożzeń z memetyką potencjalnie pozwoli zatem zbudować formalny model nie tylko dla analizy rozprzestrzeniania się memów, ale także dla ich wewnętrznej

struktury, głęboko sprzężonej z historią ich ewolucji (Walker, 2024). Ewolucja sprawczych memów nieuchronnie doprowadza do ich rosnącej autonomii. Szczególnie wyraziste tego przykłady możemy obserwować w środowisku cyfrowym. Badacze i inżynierowie pracujący w dwóch powiązanych choć rozłącznych dziedzinach: sztucznej inteligencji (AI) oraz sztucznym życiu (AL), analizują oraz konstruują układy informacyjne dążące do niezależności, celowości, intencjonalności i podmiotowości (Totschnig, 2020; Williams, 2010; Popa, 2019, Ferrag, 2025). Połączenie z memetyką na gruncie formalnym ułatwione jest przez to, że w rozwoju AI są stosowane algorytmy memetyczne (Ai, 2021).

Współczesny rozwój AI, szczególnie w zakresie modeli językowych, ale także systemów rekomendacji, częściowo opiera się na dziedzinie NLP, czyli przetwarzania języka naturalnego. Jego fundamentalna hipoteza i jej dalekie praktyczne konsekwencje mają donośne znaczenie dla memetyki muzyki, a być może nawet przysłużą się do rozwoju badań nad semantyką muzyki w ogóle. Owa fundamentalna hipoteza NLP to hipoteza dystrybucyjna Johna Firtha (1957). W najprostszej formie głosi ona, że znaczenie słowa można poznać na podstawie kontekstu jego użycia (Firth, 1957). Ta nieintuicyjna hipoteza okazała się wystarczająco prawdziwa, aby z powodzeniem służyć jako podstawa zaawansowanych systemów AI (Turney, 2010). Dzięki metodom algorytmicznym komputer analizuje przepastne korpusy tekstów i ekstrahuje informacje na temat częstości współwystępowania słów. Tworzy na tej podstawie wektorowe przestrzenie znaczenia, gdzie każde słowo stanowi punkt w określonej lokalizacji (Turney, 2010). Geometryczna bliskość dwóch słów w takiej przestrzeni wektorowej określa ich podobieństwo semantyczne. To mapowanie jest tak dokładne, że w semantycznej przestrzeni wektorowej można skutecznie przeprowadzać nawet operacje arytmetyczne, takie jak Londyn + Niemcy - Anglia = Berlin (Fares, 2017).

Badanie empiryczne zastosowania metody wektoryzacji słów do treści muzycznych dało pewne pozytywne rezultaty przewidywania muzycznych struktur (Chuan, 2018). Szersze zastosowanie ma jednak pewne uogólnienie tej techniki: algorytmy item2vec operują w przestrzeni semantycznej nie słów, a obiektów (Barkan, 2016), a nawet

utworów muzycznych (Du, 2022). Te same zasady kalkulacji semantycznych, które umożliwiają kształtowanie poprawnych zdań w języku naturalnym, umożliwiają także tworzenie interesujących muzycznie playlist. To sugeruje, że obok semantyki dystrybucyjnej języka może istnieć również pole badań semantyki dystrybucyjnej muzyki. W niniejszej pracy nie podejmuję się dalszej analizy tego złożonego tematu, ale zaznaczam, że taka praktyczna, zawarta już w systemach rekomendacji wiedza znajduje popularne praktyczne zastosowania, które mają donośne znaczenie dla memetyki muzyki. Taki opis utworów jako wektorów semantycznych może pomóc ilościowo opisać zasady kształtowania playlist jako mempleksów muzycznych.

Do programu badawczego memetyki doskonale pasują jeszcze trzy teorie, które szczegółowo przedstawię później, kiedy owe teorie będą stosowane w praktyce. Chodzi o model systemu wiabilnego (Viable System Model) (Beer, 1972), teorię aktora-sieci (Actor Network Theory) (Latour, 2005) oraz architekturę wieloskalowych kompetencji (Multiscale Competency Architecture) (Levin, 2022).

2.5. 20 tez memetyki muzyki Stevena Jana

Steven Jan to pionier i prawdopodobnie najważniejsza postać na polu globalnej memetyki muzyki. Dzięki publikacji w *Journal of Memetics* zyskał miejsce również w historii szeroko rozumianej memetyki. W końcu muzyka, czy też melodie, to pierwszy wymieniony przez Richarda Dawkinsa (1976) przykład memu, zatem memetyczna analiza muzyki ma swoje donośne znaczenie także poza muzykologią, jako przykład tego, w jaki sposób można uprawiać memetykę w ogóle. Tym niemniej zgodnie z określeniem ukutym w tej pracy można przyporządkować Jana do grupy pierwszej fali memetyki, która miała swoje poważne problemy metodologiczne. Przede wszystkim badacze nowo zafascynowani memetyką z pasją walczyli o uznanie jej jako samodzielnej nauki. Tego typu myślenie można odnaleźć również w publikacjach Jana. To efekt pierwotnych kontrowersji i emocji wzbudzanych przez memetykę, która albo miała być w całości odrzucona, albo w całości zaakceptowana.

Jan w pracy *The memetics of music: a neo-darwinian view of musical structure and culture* (2007) podaje w wątpliwość twierdzenie, że memetyka jest młodą nauką, skoro *Samolubny Gen*, czyli de facto publikacja oryginalnego manifestu memetycznego już wtedy, w 2007, miała 30 lat. To doskonałe miejsce na podkreślenie pewnego rozdroża – 30 lat dla niszowej, nowoczesnej subdyscypliny, jak np. biologia kognitywna, to już odpowiedni czas na rozwinięcie podstawowych tez, metod, ukształtowanie elementarnego programu badawczego czy nawet niekiedy nowego paradygmatu. Z drugiej strony 30 lat dla poważnej, niezależnej nauki, jak muzykologia czy matematyka to bardzo krótki czas. Wydaje się niemożliwe i absurdalne, żeby zakładać, że w 30 lat może powstać poważna i dojrzała tradycja naukowa.

Według modelu historii nauki Thomasa Kuhna (1964) rozwój każdej nauki odbywa się w cyklach, na które składają się dwa stadia: stadium rewolucyjne oraz stadium nauki

normalnej. Stadium rewolucyjne możemy powiązać z odważnymi eksperymentami, nowymi tezami, negocjacją nowych znaczeń pojęć. W wyniku tego procesu wyłania się consensus odnośnie do podstawowego paradygmatu i może rozpocząć się nauka normalna – czyli masowa produkcja sieci hipotez i testów empirycznych, która wyprowadza z ustalonej bazy dalsze konsekwencje logiczne i wiąże ów nowy program badawczy z zastaną wiedzą.

Moim zdaniem memetyka, a w szczególności memetyka muzyki może być obecnie znacznie bardziej użyteczna w praktyce, jeśli zrozumiemy ją albo jako dojrzały niewielki element szerszego programu badawczego nauk ewolucyjnych badających kulturę, albo potraktujemy ją jako spekulatywną i rodzącą się dopiero poważną, niezależną naukę w stadium rewolucyjnym. Możemy zatem albo tworzyć konkretne empiryczne i teoretyczne badania mocno zapośredniczone w dotychczasowej tradycji ewolucyjnych nauk o kulturze zaledwie wzbogacając je o memetykę, w szczególności memetykę muzyki, albo możemy oddać się daleko idącym teoretycznym spekulacjom, wiedząc, że w zasadzie muszą one póki co pozostać w sferze niedookreślonej spekulatywnej filozofii.

Z książki Jana przebija zupełnie inny tok myślenia. Badacz sam w pojedynkę podejmuje się tu i teraz stworzenia znikąd dojrzałego paradygmatu memetyki muzyki, wytwarza złożoną sieć pojęć i od razu proponuje szereg eksperymentów do ich testowania. Od 2007 minęło jednak 18 lat i nadal nie ustabilizowała się normalna praktyka memetyki muzycznej, zatem jego dzieło stworzenia ex nihilo nie przyniosło jeszcze oczekiwanych skutków. Ten szalenie ambitny projekt wydaje się przerastać możliwości oddziaływania na środowisko naukowe i upowszechniania się z przyczyn paradoksalnie wskazywanych przez samą memetykę, co jednym zdaniem podsumowuje znane z nauki o systemach prawo Galla: złożony system, który działa, musiał wyewoluować z prostego systemu, który działał (Gall, 1986). Wielką tradycję memetyki, w szczególności muzycznej, musimy budować iteracyjnie, krok po kroku, umożliwiając jej - memetyczną przecież - ewolucję.

W tym właśnie duchu i z przekonaniem o donośnej wadze książki Stevena Jana chciałbym przytoczyć jej główne tezy i odnieść się do nich krytycznie. Janowskie tezy w kontekście niniejszej pracy stanowią bardzo ważną podstawę. Tym niemniej krytyczne komentarze wykażą, że mankamentem jest poziom złożoności tego częściowo spekulatywnego, a zarazem sztywnego systemu, którego nie można funkcjonalnie rozwinąć dalej bez dokonywania pewnych istotnych modyfikacji. Ta sekcja ma na celu wykazanie mojego sceptycyzmu wobec obecnego kształtu naukowej dyscypliny memetyki muzyki i zarazem nadziei i perspektyw na przyszłość. Zanim przejdziemy do krytycznego spojrzenia na 20 najważniejszych tez książki wskazanych przez samego Stevena Jana, pragnę wyszczególnić problemy związane z nomenklaturą centralną dla memetyki muzyki tego autora. Będzie to także prezentacja pojęć, których badacz używa w swoich 20 tezach. Mówimy przede wszystkim o dwóch triadach pojęć, które poniżej wymieniam i w razie potrzeby polonizuję:

Meme, Memotype, Femotype – Mem, Memotyp, Femotyp

Mnemon, Mnemotype, Phnmnemotype – Mnemon, Mnemotyp, Fmnmemotyp

Obie te triady mają stanowić bezpośrednie analogie do klasycznych pojęć stosownych w genetyce:

Gene, genotype, phenotype – Gen, Genotyp, Fenotyp

Dzięki tej grupie pojęć Jan stara się naraz rozwiązać trzy problemy. Po pierwsze, zgodnie z duchem pierwszej fali memetyki usiłuje stworzyć bezpośrednie analogie między genetyką i memetyką. W genetyce genotyp to pełny zbiór genów kodujący informacje o danym organizmie; jest to zatem przepis, instrukcja, jaki organizm „wyprodukować” (Fortuna, 2017). To, co ostatecznie powstaje, w wyniku interakcji ze środowiskiem zewnętrznym, to fenotyp. W przypadku bliźniąt jednojajowych ten sam genotyp wytwarza dwa nieco różne fenotypy (Fraga, 2005).

Podobną konceptualizację Jan przeprowadza na gruncie memetyki. Memotyp to u niego kompletny opis memu zakodowany w konstelacji neuronów (Jan, 2007, s. 30) czyli zapisany w ludzkiej pamięci. A zatem jeden memotyp może produkować różne femotypy. Prowadzi to do pewnych komplikacji ontologicznych, które dokładniej opisuję w dalszej analizie głównych tez Jana. Badacz wyróżnia także trio mnemon, mnemotyp, fmnemotyp, które mają odnosić się w analogiczny sposób do treści ludzkiej pamięci.

Drugim problemem, który taka typologia ma rozwiązać, jest właśnie oddzielenie memów rozumianych jako treść pamięci przekazywana na drodze imitacji od treści pamięci nieprzekazywanej na drodze imitacji. To mało intuicyjne rozróżnienie podkreśla bliskość memu (myśl imitowana) i mnemonu (myśl nieimitowana), za cenę tworzenia chaotycznych kategorii pojęciowych: triada mnemon-mnemotyp-fmnemotyp jest bezpośrednią analogią triady gen-genotyp-fenotyp pomimo że mnemon z definicji (s. 30) w żaden sposób nie podlega ewolucji.

Trzeci problem to rozróżnienie memów jako bytów istniejących wewnątrz umysłów (intencja kompozytorska, melodia w głowie) od memów jako materialnych produktów kultury, takich jak partytura, płyta CD, patefon, playlista na Spotify. Jan podchodzi do tego tematu radykalnie, wielokrotnie podkreślając, że mem to wyłącznie mentalny byt, treść pamięci, przepis, z którego wytwarzane są 'femotypiczne' artefakty zewnętrzne, które już samego memu nie stanowią.

To zawołane podejście prowadzi dalej do wielu skomplikowanych i nieodwracalnych sprzeczności, które zostaną omówione poniżej. W proponowanym w niniejszej pracy memetycznym paradygmacie zamiast tego złożonego problematycznego systemu używam klasycznych dla memetyki pojęć i-meme (internal meme) oraz e-meme (external meme) jako memu wewnętrznego i zewnętrznego (McNamara, 2009), niemniej również i taki podział traktuje jako uproszczony, płynny, umowny. Zgodnie z zasadami 4E cognition (Schiavio, 2018) artefakty zewnętrzne takie jak partytury czy odtwarzacze mp3 możemy traktować jako przedłużenie umysłu (extended mind), przez co granica między memem wewnętrznym i zewnętrznym zaciera się. Tym niemniej

zarówno w ludzkiej pamięci, w partyturze, na taśmie magnetofonowej i w komputerze muzyka - memy muzyczne - funkcjonują w pewien podwójny sposób, podobny do dualności elementarnego rozróżnienia na i-meme oraz e-meme: jeden to struktura wynikająca z samego standardu zapisu a drugi to forma brzmiąca, odpowiadająca muzycznej treści umysłu.

Poniżej znajduje się 20 głównych tez memetyki muzyki Stevena Jana opisanych w książce „Memetics of music: A Neo-Darwinian view of musical structure and culture” (2007), które zostały przytoczone w oryginalnych wersjach, następnie w polskim przekładzie. Warto, aby obie wersje widniały obok siebie, z uwagi na specyfikę niektórych terminów oraz zwięzłość tych sformułowań. Każda teza została opatrzona komentarzem wyjaśniającym lub krytycznym i w wielu miejscach znajdują się odnośniki do tych części książki, w których Jan wypowiada się szerzej na określone tematy.

1. Most fundamentally, memes exist somatically, as constellations of neural interconnections. Their information content is reflected extrasomatically, as patterns in various physical media. (s. 227)

Na najbardziej podstawowym poziomie, memy istnieją cieleśnie, jako konstelacje połączeń pomiędzy neuronami. Zawarte w nich informacje są odzwierciedlane poza ciałem, jako komunikaty nadawane w różnych środkach przekazu.

W powyższej pierwszej tezie autor odwołuje się do niedookreślonego, problematycznego obszaru w memetyce, mianowicie problemu ‘podstawowego’ sposobu istnienia memów. Problem ontologii memu w niniejszej pracy został szerzej omówiony we wcześniejszych rozdziałach. Jan wskazuje w tabeli (s.30), że w ten sam określony powyżej sposób możemy zdefiniować nie tylko mem, ale także mnemon, czyli ‘jednostkę pamięci przechowywanej w mózgu’. Mnemonem można więc z materialnej perspektywy nazwać pojedynczą iterację Janowskiego memotypicznego memu, który jeszcze się nie zreplikował. Za to szerzej rozumiany memotypiczny mem sam w sobie istnieje częściowo wirtualnie, ponieważ jest to ewolucyjny ciąg, który obejmuje wszystkie kolejne iteracje informacji, która jest postrzegana, zniekształcana, oraz przekazywana dalej przez każdego kolejnego odbiorcę. Wąsko rozumiany pojedynczy

memotypiczny mem w jednym egzemplarzu nie istnieje, ponieważ mnemon staje się memem wtedy, kiedy dokona replikacji. Zatem, żeby wskazać poszczególny memotypiczny mem, należy wskazać ten mem oraz przynajmniej jedną jego kopię. Kolejny poziom komplikacji wprowadza to, że nie mamy dookreślenia, czy ta jedna dodatkowa kopia musi być memem, czy może być mnemonem. Ta definicja, najprawdopodobniej niezamierzenie, może prowadzić do nieskończonej pętli rekurencyjnej – ‘mem to taka jednostka informacji reprezentowana w mózgu, która poprzez kopiowanie wytworzyła mem’. Janowski podział na mem i mnemon rodzi wiele pytań.

Ponadto, Janowską koncepcję memów jako ‘konstelacji połączeń neuronalnych’ możemy dziś uznać za istotną historycznie, lecz zarazem wymagającą korekty i doprecyzowania. U Jana mem i mnemon w mózgu posiadają tę samą statyczną formę połączeń w sieci. Inaczej jednak, w sposób znacznie bardziej dynamiczny, memy w mózgu definiuje neuronaukowiec i kognitywista Włodzisław Duch. Duch (2021) określa mem jako „quasi-stabilny stan atraktorowy sieci pamięci asocjacyjnej” (quasi-stable associative memory network attractor state). Zatem według jego koncepcji mem to raczej dynamiczny proces niż statyczna materia. Koncepcja Jana to filozoficzna dywagacja, a z kolei Duch proponuje modele testowalne empirycznie.

2. The geno/m(n)emotype-pheno/ph(mn)emotype parallelism relates the ontology of the gene to that of the meme and affords evidence for the close affinities between replicator forms hypothesized by universal Darwinism. (s. 227)

Paralela między genotypem i m(n)emotypem oraz fenotypem i f(mn)emotypem zbliża ontologię genów i memów. Dostarcza ona dowodów na bliskie pokrewieństwo z hipotetycznymi replikatorami z teorii uniwersalnego Darwinizmu.

Sam Janowski podział na memotyp i femotyp jest faktycznie analogią podziału na genotyp i fenotyp. Za to pojęcie mnemonu, który dzieli się na mnemotyp i fmnemotyp nie ma odzwierciedlenia w terminologii genetycznej. Co więcej, samo istnienie paraleli nie jest żadnego rodzaju dowodem na prawdziwość ‘hipotezy memetycznej’.

Badacz wpada w pułapkę tworzenia niepotrzebnie dokładnej analogii pomiędzy ewolucją genetyczną oraz memetyczną, co prowadzi go do budowania skomplikowanego systemu pojęć – a niektóre z nich odnoszą się do niejasnych filozoficznych koncepcji. Tworzenie złożonego systemu niejasnych pojęć może prowadzić do nawarstwiania problemów językowych oraz trudności ze zrozumiałością tej koncepcji dla nowych czytelników. Uniwersalny Darwinizm definiuje ewolucję przez dobór naturalny jako uniwersalny algorytm (Dawkins, 1983), który może zostać ucieleśniony w różnych mediach. To naturalne, że w wyniku tego ucieleśnienia moc sprawczą wpływającą na sam materialny proces mają cechy danego medium. Zatem ten dobór naturalny w biologii oraz w kulturze może (i prawdopodobnie musi) z powodu materialnego powiązania z mediami replikacji realizować się nieco odmiennie.

Na przykład, w genetyce dziedziczność zachodzi tylko w jednym kierunku. Rodzic nie może odziedziczyć po dziecku koloru oczu. Za to w memetyce nie ma tego ograniczenia. Muzyka grunge w momencie powstania częściowo opierała się na muzyce punk, dziedziczyła jej cechy (Stafford, 2018). Ale muzyka punk nie przestała się rozwijać – podlegać memetycznej ewolucji. W późniejszych latach do muzyki punk przenikały wpływy z muzyki grunge (Ozzi, 2022). Takie dwukierunkowe dziedziczenie jest opisywane również w ewolucyjnym językoznawstwie (Brown, 2008). Reasumując, szukanie bezpośrednich paraleli do ewolucji memetycznej w procesach ewolucji genetycznej jest błędnym podejściem prowadzącym do fałszywych wniosków.

3. A musical meme is a replicated pattern in some syntactic/digital element of music - principally pitch and, to a lesser extent, rhythm - transmitted between individuals by imitation as part of a neo-Darwinian process of cultural transmission and evolution. (s. 228)

Mem muzyczny jest zreplikowanym wzorcem w określonym syntaktycznym/skwantyzowanym parametrze muzyki - przede wszystkim w wysokości dźwięku, rzadziej w rytmie - który jest przenoszony pomiędzy ludźmi poprzez imitację. Jest to element procesu ewolucji i transmisji kultury wynikający z założeń neo-Darwinizmu.

Z tezą trzecią, podobnie odważną jak poprzednie, nie można się zgodzić. Ten bardzo ogólny wniosek został prawdopodobnie oparty na analizach muzyczno-memetycznych Jana, które zostały przeprowadzone przy uwzględnieniu pewnych bardzo konkretnych, ograniczających założeń. Jan mówi wprost, że tekst nutowy to jego zdaniem najlepszy materiał do analizy muzycznej i ten właśnie materiał wybiera. Nie analizuje on nagrań ani innych 'femotypicznych produktów memów muzycznych'. Być może, jeśli memy przede wszystkim przenoszą się z człowieka na człowieka, warto byłoby jako przedmiot analizy porównawczej przyjąć produkty psychologiczne związane z recepcją oraz imitacją danego memu? Analiza zapisu nutowego tak, jakby był on podstawowym medium istnienia memów memotypicznych jest z całą pewnością niewystarczająca - chociażby dlatego, że obejmuje ona tylko i wyłącznie te parametry, które w zapisie nutowym zostały skwantyfikowane i znajdują się na pierwszym planie - jak wysokość chromatyczna, metrum i rytm. Nie ma żadnych dowodów na to, że barwa, dynamika, czy intonacja to parametry dźwięku, które nie podlegają memetycznej imitacji.

Co więcej, analizy Jana dotyczą tylko i wyłącznie europejskiej muzyki artystycznej okresu klasyczno-romantycznego. Zatem wniosek o prymacie skwantyfikowanych struktur harmoniczno-melodycznych oraz rytmicznych w procesie imitacji odnosi się głównie do muzyki tego miejsca i okresu. Skupienie na takich jasno skwantyfikowanych wartościach zbliża muzyczną memetykę Jana do paradygmatu memetyki prezentowanego na łamach *Journal of Memetics* (w którym zresztą Jan publikował) – ta pierwsza fala memetyków zwracała szczególną uwagę na dyskretną strukturę memów, co miało umożliwiać kopiowanie ich z wysoką wiernością.

4. Some pitches and rhythms in a piece, principally articulated and unharmonized monads or dyads, may be non-memetic, being generated spontaneously/ analogically, not replicated memetically/homologically. (s. 228)

Niektóre z wysokości oraz rytmów w utworze, przede wszystkim pojedyncze, nieharmonizowane dźwięki oraz dwudźwięki mogą być nie-memetyczne, kiedy są tworzone spontanicznie lub poprzez analogię, a nie replikowane memetycznie i homologicznie.

Teza czwarta pozornie brzmi jak pewne lekkie załagodzenie pierwotnej ortodoksji pierwszej fali memetyki – skoro memy istnieją, to pochopnie wyciągano wnioski, że całkowicie zdominowały one umysły ludzi i wszystkie nasze zachowania oprócz podstawowych instynktów są po prostu replikacją memów. Odnosi się to także do krytyki poglądu, jakoby kreatywność oraz wolicjonalność kompozytorów (i ludzi w ogóle) była całkowicie podległa memom. Jan pozwala sobie na odejście od tej ortodoksji – zezwala na dostrzeżenie niewielkiej klasy dźwięków nie-memetycznych i ściśle określa ich dozwoloną strukturę. Ta refleksja wydaje się daleko idącą spekulacją, która jako precyzyjny wyjątek ma de facto potwierdzić regułę.

W części II i III niniejszej dysertacji zauważam bliską relację między imitowaniem i kreatywnością, ale jednak rozgraniczam oba te pojęcia. Ludzki umysł zasiedlony memami nadal może wykazywać autentyczną, spontaniczną kreatywność. Co więcej, imitacja bez absolutnie żadnej kreatywności to przypadek graniczny, który obserwujemy przede wszystkim w maszynowej reprodukcji fizycznych przejawów memów takich jak płyty CD czy pliki mp3. Człowiek z natury działa kreatywnie, nawet jeśli kreatywnie kopiuje zastany wzorzec (Okada, 2016). Inną perspektywę nadaje tutaj zewnętrzna klasyfikacja ludzkiej twórczości, także muzycznej. Osoba absolutnie nieznająca repertuaru muzycznego ludzkości, nauczona podstawowych chwytów na gitarę, bardzo szybko skomponuje piosenkę, którą zewnętrzny obserwator uzna za plagiat, niekreatywną kopię, kopiowanie gotowców, powielanie wzorców. Jednak z perspektywy tej osoby będzie to owoc autentycznej kreatywności, wynik sprzężonego działania indywidualnych psychologicznych, poznawczych, fizjologicznych uwarunkowań postrzegania, przetwarzania i wytwarzania muzyki.

To nie przypadek, że taki ‘samorodek’ można jednak uznać za memetycznie spokrewniony z szerszym repertuarem. Sama gitara, system strojenia, budowa akordów i sposób kształtowania formy przez akordowe następstwa to kulturowe wynalazki, memy muzyczne wyższego rzędu abstrakcji niż sam repertuar. A zatem ten instrument sprzężony z konkretnymi kompetencjami można już uznać za maszynę do replikacji

memów, w której ludzka autentyczna kreatywność stanowi rodzaj napędu czy też ograniczonej warunkami brzegowymi podstawy.

Idąc dalej tym tropem, jeśli zakładamy, że memotypiczne memy to abstrakcyjne zbiory, które obejmują całą ewoluującą kulturę, nie możemy w ogóle mówić o nie-memetycznych dźwiękach w kulturze. Za to Jan sugeruje, że w ramach jednego dzieła mogą występować dźwięki memetyczne oraz nie-memetyczne (s. 79). Daje zatem do zrozumienia, że proces komponowania utworu jest tu rozumiany jako proces dobierania odpowiednich memów dźwiękowych oraz przeplatania ich 'oryginalnymi' łącznikami. Ale tradycję dodawania takich 'oryginalnych' łączników również trzeba uznać za pewien zwyczaj, element kultury muzycznej, który podlega imitacji i ewolucji. Kiedy Jan zastanawia się nad tym, czy istnieją jakieś nie-memetyczne dźwięki (s. 79), dochodzi również do wniosku, że niektóre z nich mogą składać się ze zbyt małej ilości informacji, aby zostać odpowiednio postrzeżone, lub mogą być niemożliwe do zimitowania. Nie są znane jednak żadne ścisłe zasady grupowania dźwięków do pojedynczych memów, a poza tym memy mogą swobodnie przeplatać się ze sobą, za każdym postrzeżeniem dobierane indywidualnie przez postrzegającego. Co do niemożliwości imitacji - niedookreślona imitacja jest procesem tak szerokim, że potencjalnie obejmuje każdą możliwą zmianę parametrów dźwięku. Podsumowując, analiza kultury muzycznej jako złożonego zjawiska, obejmującego szerszą praktykę, a nie tylko zbiór treści dzieł muzycznych, wskazuje, że nie można mówić o żadnych nie-memetycznych dźwiękach wewnątrz szeroko pojmowanej ludzkiej kultury.

5. The work is merely a phemotypic manifestation of a temporary alliance between memes. For all the celebration of the work in western discourse on music since c. 1750, the selfish meme is the fundamental unit of all musical cultures. (s. 228)

Dzieło muzyczne jest zaledwie femotypiczną manifestacją chwilowego układu (sojuszu) memów. We wszystkich dziełach i wszystkich kulturach muzycznych omawianych w zachodnim dyskursie muzycznym od około 1750 roku podstawową jednostką jest samolubny mem.

Istotą tego punktu jest określenie memu jako podstawowej materii każdej nauki o kulturze, również kulturze muzycznej. Model Jana zakłada, że jesteśmy w stanie stworzyć całą historię muzyki (w pierwszej kolejności tej klasyczno-romantycznej) z perspektywy pojedynczych, femotypicznych memów muzycznych; jeśli chodzi o historię muzyki klasyczno-romantycznej, takim podstawowym femotypicznym memem, który Jan analizuje, są figury melorytmiczne. Komputerowa analiza ma pomóc w przebadaniu klasyczno-romantycznego repertuaru pod kątem takich melorytmicznych figur, które w nieco zmienionej, zimitowanej formie pojawiają się w kolejnych dziełach i w ten sposób napisać historię muzyki z perspektywy pojedynczych, ewoluujących figur. Samolubny mem to pojęcie oparte na samolubnym genie Dawkinsa, które ma podkreślić, że pojedynczy replikator (gen lub mem) zachowuje się tak, jakby dbał o własny interes. Tym samym każdy fragment muzyczny będzie mniej lub bardziej 'dążył' do replikacji.

Jan pisze, że dzieło muzyczne to 'zaledwie' chwilowy sojusz memów (s. 228). Takim językiem ('zaledwie') podkreśla, że to, co najważniejsze, odbywa się gdzie indziej. Z reszty pracy można łatwo wyciągnąć wniosek, że to, co jego zdaniem najważniejsze, to muzyczne motywy, tematy, frazy, akordy – części utworu muzycznego wydzielane przez klasyczną analizę muzyczną (Riemann, 1929). W ten sposób Jan zdradza swoją nieracjonalną preferencję i proveniencję związane z analizą muzyczną. Dziś znacznie bardziej użyteczną perspektywę daje nam zapożyczone z biologii rozwojowej Michaela Levina określenie multiscale competency architecture – wieloskalowa architektura kompetencji (Levin 2023). Memy mogą tworzyć złożoną sieć, współpracującą na wielu poziomach abstrakcji i organizacji, z których każdy może posiadać pewną swobodę i dowolność, a także pewne kompetencje, zarazem nie posiadając żadnego poziomu uprzywilejowanego, najważniejszego, centralnie zarządzającego resztą struktury. Więcej o strukturach różnych rzędów Steven Jan pisze w swoim punkcie 20.

Jan zapewne opiera swój wywód także na pewnym specyficznym rozumieniu myśli Daniela Dennetta (Dennett, 1995), którego w swojej pracy często cytuje i parafrazuje. Zgodnie ze stwierdzeniem wprost zawartym w innej części książki (s. 181), Jan deklaruje wiarę w oddziaływanie dzieło-styl w odwrotnym kierunku niż ten sugerowany

przez analizę Schenkerowską. Według myśli Schenkera to szeroki, głęboki archetyp generuje pomniejsze, niewiele różne od siebie dzieła. Jan, za Dennettem, mówi o bottom-up design – projektowaniu oddolnym. W tej perspektywie sprawczość zawarta jest na samym dole struktury – pojedyncze, najmniejsze memy muzyczne składają się w większe i większe całości, aż ostatecznie wytwarzają się ogólne kompleksy stylów muzycznych. Problem w sformułowaniu Jana odzwierciedla przyjęty przez niego – moim zdaniem nieracjonalnie – jednoznaczny model najmniejszego możliwego muzycznego memu i przypisanie jemu elementarnej i zarazem kluczowej sprawczości w systemie muzycznym.

6. Musical memes, like those propagated in other substrates, manifest the attributes of longevity, fecundity and copying-fidelity. These attributes interact in complex ways but, invariably, some memes enjoy great longevity as a result of the preservation of their phenotypic forms, some are more fecund than their alleles and all manifest a higher degree of copying-infidelity than their genetic counterparts. (s. 228)

Memy muzyczne, podobnie jak memy innych typów, charakteryzują się trwałością, płodnością i kopiowalnością. Te cechy są ze sobą powiązane w skomplikowany sposób. Niektóre memy cieszą się dużą trwałością dzięki temu, że ludzie dbają o ich fenotypiczne inkarnacje, inne są bardziej płodne niż ich allele, a wszystkie z nich przejawiają niższą precyzję w kopiowalności niż w przypadku genów.

Trwałość fenotypicznych form memów muzycznych powiązana jest tutaj między innymi z takimi mediami jak zapis nutowy, kaseata audio lub formaty zapisu cyfrowego. Te pośrednie modusy istnienia memów muzycznych pozwalają na ich długowieczność oraz pozwalają przełamywać bariery geograficzne. Obecnie, dzięki swobodnemu przekazywaniu plików dźwiękowych przez internet, ewolucja danych memów dźwiękowych może przebiegać równolegle w różnych miejscach na świecie. Koncepcja płodności memów jest związana z tym, jak bardzo 'prowokują' one kolejnych nosicieli do imitowania ich, tworzenia kolejnych kopii czy iteracji. W rozdziale 4.2 Jan przypomina znany fenomen earworms jako melodii, które prowokują 'nosiciela', aby je

nucii. Jest to przykład bardzo płodnych memów muzycznych - takich, które mają tendencje do częstej, wręcz epidemicznej replikacji.

Jan posługuje się jednak terminem allele, który to jest bezpośrednią paralełą do terminu genetycznego. Według koncepcji Jana allele w memetyce muzycznej powstają przy memach bardziej skomplikowanych strukturalnie. Dwa memy melodyczne mogą więc rywalizować niczym allele danego genu o miejsce w temacie danej formy sonatowej. Podobna rywalizacja może się odbywać na poziomie formalnym. Jan zauważa istnienie także 'strukturalnych' memów muzycznych. Na tej płaszczyźnie można mówić na przykład o allelicznej rywalizacji dwu- i trzyczęściowej formy koncertu. Kopiowalność oznacza zaś to, czy dany mem w ogóle może zostać zreplikowany i jaki stopień utraty informacji może to spowodować. Twierdzenie, że kopiowalność memów jest mniejsza niż kopiowalność genów, znaczy, że kolejna iteracja danego fenotypu może znacznie różnić się od poprzedniej. Ewolucja może postępować szybciej, a zawartość informacyjna może zmieniać się w większym stopniu, niż ma to miejsce w ewolucji genetycznej. Jan postuluje zdecydowane stwierdzenie, że wszystkie memy przejawiają mniejszą kopiowalność niż geny. Z taką tezą można zgodzić się tylko przy pewnych precyzyjnie określonych założeniach.

Po pierwsze te założenia muszą dotyczyć sposobu mierzenia kopiowalności w genetyce (Kunkel, 2004). Jeżeli precyzyjnie kopiowalność informacji genetycznej miałyby oznaczać to, jaki procent informacji genetycznej zostaje bez mutacji przekazany od rodzica do dziecka, ta ilość informacji genetycznej u człowieka nie przekracza 50% - ponieważ dziecko łączy geny obu rodziców (Roach, 2010). Mówilibyśmy zatem o niskiej kopiowalności, z pewnością niższej niż przy kopiowaniu muzycznego albumu na płycie CD, gdzie możemy realnie uzyskać kopiowalność na poziomie 100%.

Jednak Jan konceptualizuje to w inny sposób. Nie bierze pod uwagę tego, że replikacja biologiczna człowieka wiąże się z utratą połowy informacji genetycznej. Skupia się tylko na tym, że to właśnie wybrane 50% jest kopiowane z niemal idealną skutecznością. A zatem skupia się tylko na części procesu replikacji DNA. Z drugiej strony, odnośnie do

memów, postuluje istnienie bardzo złożonego cyklu replikacji i bierze pod uwagę jego całość: mem w umyśle jako pewien schemat, który potem dopiero zostaje zapośredniczony w jakiejś femotypicznej strukturze, czy femotypicznym memie, i dopiero potem w umyśle kolejnego 'nosiela' wytwarza następną iterację prawdziwego memu. Cykl może być również jeszcze dłuższy: kompozytor tworzy dzieło w głowie (powstaje mem), następnie zapisuje je w partyturze (powstaje mem femotypiczny), wykonawca je wykonuje (powstaje kolejny mem femotypiczny) i dopiero w umyśle słuchacza powstaje kolejny prawdziwy mem. Memy femotypiczne w skrócie możemy nazwać tutaj nieprawdziwymi, ponieważ Jan nie uwzględnia ich w swoim modelu kopiowalności. W jego modelu femotypiczna forma memu to tylko forma, a nie właściwy mem. Dlatego na przykład w mojej analizie nie mogę omówić ewolucji czy też mutacji następującej pomiędzy partyturą dzieła a jego wykonaniem, co, jak wiadomo chociażby ze słynnych wykonań Bacha przez Glenna Goulda, bywa bardzo istotną siłą w procesie ewolucji muzyki (Nussbaum, 2007). Jan w swojej mierze wierności kopiowania memu muzycznego pomija ten etap. Trudno nie ulec wrażeniu, że jest to subiektywna i nieracjonalna selektywność wynikająca wprost z niekonsekwencji pojęciowej stojącej u podstaw tego modelu.

7. Particulate and digital, musical memes normally consist of at least three components, two elements normally possessing insufficient information content to impart to a pattern sufficient perceptual and cognitive salience to function as a meme. There is no fixed upper limit to the number of components, but this value and the copying-fidelity are inversely proportional.

Naturalne oraz skwantyzowane muzyczne memy zwykle składają się przynajmniej z trzech elementów. Dwie części składowe zwykle zawierają zbyt mało informacji, aby stworzyć wzór, który zostanie dostrzeżony i zafunkcjonuje jako mem. Nie istnieje żaden górny limit liczby elementów, ale ta liczba jest odwrotnie proporcjonalna do kopiowalności.

W powyższym punkcie starałem się wykazać, że Jan do swojej miary wierności kopiowania memów unika traktowania 'memów femotypicznych' jako memów prawdziwych. A jednak już w tym punkcie Jan mówi o dwóch lub trzech elementach wyszczególnionych z partytury zgodnie z miarą cognitive salience. Jeśli nie miałyby tak

być, to co autor może mieć na myśli, mówiąc o dwóch lub trzech komponentach memu obecnego w umyśle? Jan konceptualizuje memy obecne w umyśle jako połączenia między neuronami, czy zatem może mieć na myśli istnienie dokładnie dwóch lub trzech synaps? Czy miałyby one łączyć konkretną liczbę neuronów, a może posiadać jeszcze inne parametry? Z analizy treści książki wynika jednak coś innego. Jan ma na myśli dwa lub trzy elementy dostrzeżone w partyturze. A zatem na podstawie partytury definiuje istotne cechy memu muzycznego. Pomiędzy punktem 6 i 7 istnieje zatem całkowita sprzeczność w definicji ontologii memu. Nie można tego interpretować inaczej niż tak, że Jan żongluje tymi terminami tak, żeby dopasować je do omawianej lokalnie koncepcji.

Dalszą analizę prowadzimy zatem odnośnie muzycznych memów femotypicznych - czy też po prostu branych de facto przez Jana najczęściej na tapet partytur klasyczno-romantycznej artystycznej muzyki Zachodniej. Wyrazista, wyróżnialna wobec tła struktura muzyczna to główny temat 7 punktu Jana. Ten punkt odnosi się przede wszystkim do częściowo nieprzetłumaczalnego pojęcia 'cognitive salience'. Termin salience ma swoje zastosowanie między innymi w psychologii, socjologii i językoznawstwie (Liu, 2025; Stryker, 1968; Giora, 1997). Salience jest tą cechą danej struktury informacyjnej, która sprawia, że jest ona dostrzegalna na tle swojego kontekstu i otoczenia. Steven Jan w swoich opisach cognitive salience (choćby w 4.1) odwołuje się do modelu I-R (implication-realization), stworzonego w dużej mierze przez Leonarda B. Meyera (Meyer, 1956) oraz Eugene'a Narmoura (Narmour, 1984). Jest to model analizy muzycznej opartej na oczekiwaniach słuchacza - rozwijająca się struktura muzyczna powoduje, że słuchacz spodziewa się tego, jakie zmiany mogą w niej nastąpić. Zaskoczenie słuchacza, 'rozczarowanie' oczekiwań miałyby zwiększać współczynnik cognitive salience.

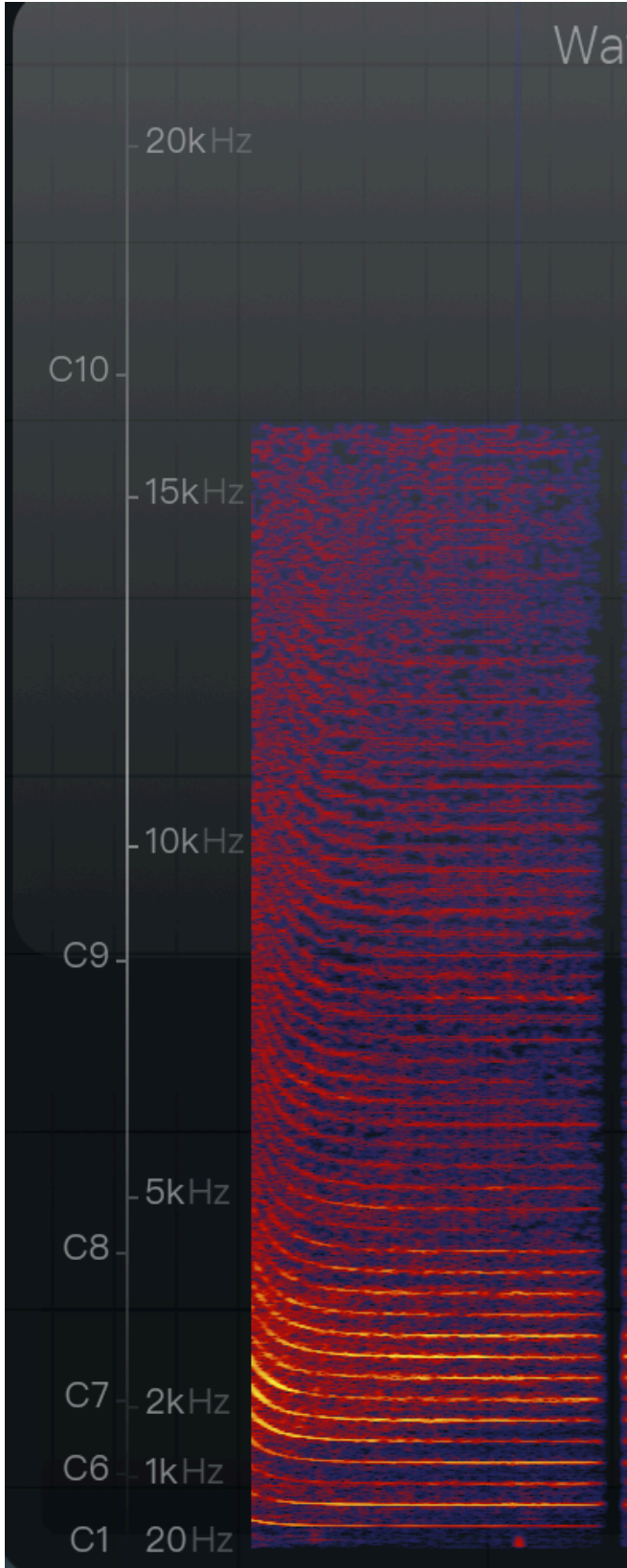
To jednak, że Jan pisze o 'trzech elementach', które muszą znaleźć się w danej strukturze muzycznej, jest jednak wysoce spekulatywne. Elementy musiałyby być albo 'zimnymi' bitami informacji, co wymagałoby ustalenia znacznie bardziej precyzyjnej definicji, aby nie odnieść wrażenia, że autor ma na myśli bity jako elementarne części

w sensie digital physics, albo muszą być psychologicznie zauważonymi informacjami, które równie trudno jednoznacznie podzielić i sklasyfikować.

Określenie liczby bitów informacji może mieć sens, jeśli mówimy o zapisie nutowym, ponieważ łatwo w nim jednoznacznie wykazać zmianę określonego parametru. W całej książce Jan ma tendencję do analizy właśnie zapisów nutowych, ale trudno na ich podstawie wyciągać uprawnione wnioski dla każdego medium fonotypicznych memów muzycznych, skoro zapis nutowy to tylko jedna, bardzo skompresowana forma symbolicznego zapisu brzmiącej muzyki. Zatem de facto siódmy punkt Jana należałoby rozumieć jako propozycję tego, żeby za najmniejszy możliwy fonotypiczny mem muzyczny odnaleziony w muzycznej partyturze traktować trzy nuty. Badacz opiera tę subiektywną wartość na mierze potencjału implikacji-realizacji. Zgodnie z tym tokiem myślenia, żeby powstała wyrazista struktura muzyczna, musi najpierw zaistnieć jeden i drugi dźwięk, które budują pewne oczekiwanie odnośnie do trzeciego dźwięku, a wtedy ten trzeci dźwięk może to oczekiwanie wypełnić lub zawieść. Jeżeli konceptualizujemy muzykę wyłącznie jako nuty o atrybutach wysokości i następstw czasowych, ten argument może brzmieć przekonująco.

Tym niemniej, ten naciągany, spekulatywny argument po prostu nie zgadza się z empirią wykraczającą poza spekulacje z dziedziny analizy muzycznej. W części I niniejszej pracy cały podrozdział poświęciłem analizie memu muzycznego „MLG Airhorn”, który w najprostszej formie składa się dokładnie z jednego dźwięku – przynajmniej w formie jednego dźwięku można by go zapisać na pięciolinii.

Inaczej wygląda to na spektrogramie:



Jak przeliczyć na kategorie nutowe ilość informacji zawartych w spektrum dźwięku? Czy bogata barwa może być tożsama z akordem? A może dwudźwiękiem? Czy skoro dźwięk airhorna ewidentnie wykazuje schodzące glissando, można mówić tu o dwóch klasach wysokości, czy jednej? A może o każdej dostrzegalnej w nagraniu klasie wysokości pomiędzy wartością maksymalną i minimalną? Te wszystkie pytania pozostają bez jednoznacznych odpowiedzi i zwracają uwagę na to, że ten punkt jest sformułowany w sposób bardzo problematyczny. Co więcej, autor z dużą dozą ostrożności pisze o tym, że memy 'zwykle' posiadają 'przynajmniej trzy elementy'. Słowo 'zwykle' może nas prowadzić albo do konkluzji, że badacz po prostu sam nie jest pewien, czy jego teza jest trafna – albo, traktując je jako zastosowane precyzyjnie, możemy je odczytać jako informację o prawidłowości statystycznej. Mianowicie: 'memy muzyczne przeważnie posiadają co najmniej trzy elementy'. Nadal – owe 'elementy' to zbyt rozmyta kategoria, aby wyciągnąć te przytoczone przez Jana konkretne wnioski na temat liczby elementów w memie, nawet w sferze zapisu nutowego. Czy oznaczenie artykulacji jest elementem? Ile elementów zawiera tryl?

Trzeba także dodać, że w omawianym właśnie punkcie 7 Jan mierzy się z problemem zdefiniowania najmniejszego możliwego memu muzycznego, ponieważ mówi o przynajmniej (at least) trzech elementach. W tej pracy, we wcześniejszym rozdziale, omawiam ten problem i usiłuję przekonująco wykazać, że nie da się go rozwiązać w sensie ogólnym, bez przyjęcia lokalnych definicji i ograniczeń. Jan usiłuje w zasadzie rozwiązać ten problem dla ograniczonych warunków, gdzie ontologia muzyki sprowadza się do zapisu w partyturze, ale nie informuje o tym w swoich wnioskach; jego teza ma charakter ogólny. Nie możemy jej nawet uznać za fałszywą, a raczej za głęboko problematyczną i nacechowaną niespójnością pojęciową i logiczną z resztą tez.

8. The segmentation of a musical string into components capable of functioning as memes is determined partly by innate, bottom-up attributes of human perception and cognition, and partly by learned, top-down factors, including the presence of a coindex. (s. 228)

Segmentacja muzycznych całości na części, które mogą funkcjonować jako odrębne memy, jest zdeterminowana częściowo przez wrodzone cechy ludzkiego poznania i percepcji (oddolnie), a z drugiej

strony przez czynniki wyuczone, takie jak coindex (odgórnie).

W tezie 8. Jan wprowadza pojęcie koindeksu (coindex). Jest ono zapożyczone z językoznawstwa, konkretnie gramatyki generatywnej Noama Chomsky'ego (Chomsky, 1981). Koindeks to to wspólna etykieta mająca podkreślić powiązanie dwóch elementów. Replikowanie memów musi faktycznie być zależne od zdolności percepcyjnych człowieka. Również grupowanie ich w większe całości prawdopodobnie przebiega zgodnie ze sposobem postulowanym przez autora. Z tym punktem, który ma połączyć memetykę muzyki z Dennettowskimi kategoriami top-down i bottom-up (Dennett, 1995, s. 135; Dennett, 2017, s. 152), niniejsza dysertacja jest zgodna.

Na marginesie można dodać, że w podanej narracji istnieje jednak pewna nieścisłość - autor mówi o ograniczeniach percepcji słuchowej, jednocześnie posługując się przykładami nutowymi (3.3). Problematyczny przypadek dla tak sformułowanej tezy stanowi na przykład muzyka dodekafoniczna. Nietrudno wyobrazić sobie, że muzyka zapisana w postaci nutowej może ewoluować w sposób zupełnie niezależny od ewolucji 'brzmiących' dźwięków. Zapis nutowy to przede wszystkim formalna koncepcja, która jest faktycznie powiązana z dźwiękami, jednak do postrzegania oraz przetwarzania jej nie jest w ogóle potrzebny słuch. Jeśli rozumiemy zapis nutowy jako postrzegany diachronicznie zapis abstrakcyjnych informacji, widzimy też, że ograniczenia związane z czasem percepcji są tu równie nieistotne. Właściwie rozwój artystycznej muzyki zachodnioeuropejskiej jest w dużym stopniu zależny od memetycznych struktur wyższego rzędu, takich jak ewoluujący zespół formalnych zasad 'poprawności' muzycznej oraz warunki społeczne, ekonomiczne polityczne.

9. Coindexation describes the mapping of two discrete musical patterns against each other. The more elements such units have in common, the greater the probability that both are instatations of a given meme and therefore that they are related to each other in a nexus of replication (and vice versa).

Koindeksacja to opis wzajemnych zależności między dwoma dyskretnymi fragmentami muzycznymi. Im więcej zawierają one wspólnych elementów, tym większe prawdopodobieństwo, że są różnymi iteracjami jednego mema i są ze sobą spokrewnione w drodze replikacji.

W punkcie 9 Jan rozwija swoją koncepcję memetycznomuzycznej koindeksacji. W jego typologii (3.3.3), koindeksacja memów muzycznych dzieli się na kilka podtypów. Przede wszystkim służy ona temu, aby odszukać memy, które są sobie bliskie na drodze imitacyjnego pokrewieństwa. Jeśli dwa femotypiczne memy muzyczne są do siebie bardzo podobne pod względem struktury, możemy założyć, że mają one wspólnego przodka, lub że jeden jest ewolucyjnym następcą drugiego.

Co więcej, zgodnie z jego systemem, w którym realnie mem istnieje tylko jako schemat pamięciowy w umyśle, a partytury to tylko artefakty, efekt uboczny istnienia tych mentalnych wzorców, Jan sugeruje, że znaczne podobieństwo w partyturze może prowadzić nas do tego samego wzorca mentalnego – czyli pełny koindeks oznaczałby pokrewieństwo całkowite, tożsame z tym obserwowanym w genetyce na poziomie bliźniąt jednojajowych. Choć Jan nie odnosi się do teorii złożań (assembly theory) (Sharma, 2023), jego teza o wspólnym pochodzeniu ewolucyjnym dwóch obiektów jest z nią spójna. To o tyle ważne, że teoria złożań proponuje szersze materialne i ewolucyjne wyjaśnienie, dlaczego tak się dzieje.

Teoria złożań (assembly) zauważa, że złożoność (complexity), a szczególnie złożoność dostrzeżona w większej liczbie tożsamych egzemplarzy dowolnego obiektu nie jest dziełem przypadku. W zgodzie z teoretycznymi wskazaniem Jana, zwolennicy teorii złożań interpretują zaobserwowanie większej liczby bardzo podobnych złożonych struktur jako informację o ich wspólnym pochodzeniu – to znaczy uczestnictwa w tym samym procesie selekcji. Precyzyjnie, teoria złożań mówi, że prawdopodobieństwo spontanicznego powstania obiektu spada wykładniczo wraz ze wzrostem jego złożoności, mierzonej za pomocą miary nazwanej indeksem złożań (assembly index) (Walker, 2024).

W konsekwencji, zaobserwowanie obiektu o wysokim indeksie złożań w wielu kopiach jest traktowane jako jednoznaczna ‘sygnatura selekcji’ (signature of selection) (Sharma, 2023) – dowód na istnienie nieprzypadkowego, napędzanego informacją mechanizmu,

który doprowadził do wygenerowania tego obiektu. W tym ujęciu, Janowska koindeksacja, identyfikująca 'wiele wspólnych elementów' w różnych instancjach memu, jest koncepcyjnym odpowiednikiem ustalania wysokiego indeksu złożenia. Jego wniosek, że wskazuje to na uczestnictwo w tym samym procesie replikacji, jest tożsamy z fundamentalnym postulatem teorii złożzeń, która interpretuje takie odkrycie jako dowód na działanie procesu selekcyjnego, który umożliwił powstanie i powielenie mało prawdopodobnej struktury. Obie teorie, wychodząc z różnych dziedzin, dochodzą do tego samego wniosku: obfitość złożoności nie jest dziełem przypadku, lecz wynikiem selekcji. Ponadto, Jan, wyjaśniając swoją koncepcję koindeksu, wprowadza ciekawą metaforę. Zaznacza, że koindeksacja może pod względem semiotycznym działać podobnie jak hiperłącza w internecie - prowadzić słuchacza bezpośrednio do innego dzieła, w którym znajduje się mem, oznaczony tym samym koindeksem.

10. The greater the number of components in a meme, the greater the probability that a given coindex is directly derived from it - that it is not an element of the common memetic currency or a dialect (and vice versa). (s.228)

Im bardziej złożony jest dany mem, tym większe prawdopodobieństwo, że określająca go koindeksacja wynika bezpośrednio z jego struktury i nie jest elementem lokalnego memetycznego dialektu.

Jedną z koncepcji Jana jest to, że memy muzyczne funkcjonują wewnątrz skomplikowanej hierarchii, która wiąże się między innymi ze strukturą stylów muzycznych. Jeśli dany styl jest wyjątkowo rygorystyczny, wtedy dzieła, które są w nim wyrażone, w dużym stopniu będą zawierały jego piętno w swojej strukturze muzycznej. Możemy w tej sytuacji odnaleźć wiele koindeksacji i nie będą one wynikały z ewolucji femotypicznych memów najmniejszego szczebla, ale ze wspólnej przynależności wszystkich komórek brzmieniowych do jednego, odgórnego stylu.

Jeśli zaś konkretny femotypiczny mem charakteryzuje się dużą złożonością, możemy założyć, że odnaleziony podobny femotypiczny mem jest raczej selekcyjnie spokrewniony z nim, niż że ich podobieństwa wynikają z bezpośredniego zaczerpnięcia

z generatywnej puli podstawowych elementów języka muzycznego danej puli memetycznej (jak kadencje, zwroty harmoniczne, ozdobniki itd.). A zatem punkt 10 ma proponować narzędzie do odróżniania złożonych, wyrazistych, spokrewnionych selekcyjnie memów od generatywnego tła ogólnego stylu i konwencji.

Zachodzi tu jednak problem metodologiczny. Teza 10, choć ma być kluczowa dla odróżnienia bezpośredniej replikacji od zwykłej konwencji stylistycznej, wpada w pułapkę zastawioną przez wcześniej sformułowaną teorię Jana. Skoro memy, aby przetrwać proces replikacji, muszą być relatywnie proste i odporne na błędy kopiowania (zgodnie z Tezą 7.), to skąd wziąć złożone, unikatowe wzorce, które jako jedyne miałyby stanowić wiarygodną 'sygnaturę selekcji' świadczącą o bezpośrednim zapożyczeniu? Innymi słowy – jak jednoznacznie odróżnić jeden złożony mem od grupy prostych memów? Teoretycznie możemy powołać się na pojęcie mempleksu i poszukiwać wyrazistości poznawczej (cognitive salience). Taka odpowiedź jest teoretycznie zgrabna, ale w praktyce analitycznej staje się subiektywnym aktem wiary. To analityk, a nie obiektywna miara, musi arbitralnie zdecydować, czy dana sekwencja jest jednym, spójnym mempleksem (którego złożoność jest dowodem na bezpośrednią replikację), czy zaledwie przypadkowym zbiorem prostszych memów funkcjonujących w ramach wspólnej konwencji. W rezultacie teza, która obiecuje analityczną precyzję w tropieniu pochodzenia idei muzycznych, pozostawia kluczowy moment identyfikacji w nierozstrzygalnej, szarej strefie interpretacji.

11. Memes exist at several structural-hierarchic levels within a piece, from the immediate musical surface, where they have a clear perceptual-cognitive reality, to the deepest and most remote level, where they have a distributed, virtual existence. A given meme may therefore exist at several hierarchic levels within the same piece - the scale-degree succession 3-2-1, for instance, may simultaneously function as a foreground element, a middleground structure and as an Urlinie - or at different levels in different pieces. (s. 229)

Memmy istnieją na kilku poziomach strukturalno-hierarchicznych dzieła. Na najbliższej muzycznej powierzchni są bezpośrednio słyszalne oraz rozpoznawalne, a na najgłębszych strukturalnych poziomach

działa ich istnienie jest rozdystrybuowane oraz wirtualne. Dany mem może więc istnieć na kilku hierarchicznych poziomach wewnątrz jednego utworu - na przykład pochod tercja-sekunda-pryma może jednocześnie funkcjonować jako element warstwy powierzchniowej, środkowej oraz jako praosnowa - lub na innych poziomach w innych utworach.

Jednym z punktów memetyki muzyki Stevena Jana jest jej powiązanie z analizą Schenkerowską (Schenker, 1979). Wynika to przede wszystkim z trzech powodów. Po pierwsze, przedmiotem analizy memetycznej Jana są tylko i wyłącznie zapisy nutowe. Musi więc odnieść się do takiej metody analizy, która pasuje do tego medium. Po drugie, Jan odwołuje się do ewolucyjnego, organicznego charakteru analizy Schenkerowskiej (1.3) jako odpowiedniego narzędzia dla memetyki. Po trzecie, ten rodzaj analizy dobrze odpowiada stylowi muzycznemu, którym zajmuje się Jan, czyli artystycznej muzyce okresu klasyczno-romantycznego.

Teza 11, postulująca istnienie memów na wielu poziomach strukturalno-hierarchicznych, jest centralnym punktem teorii Jana, w którym łączy on memetykę z analityczną tradycją Heinricha Schenkera. To powiązanie wynika z metodologicznych wyborów Jana: jego koncentracja na zapisie nutowym muzyki klasyczno-romantycznej oraz skojarzeniach ewolucyjnych w naturalny sposób prowadzi go do Schenkerowskiej, hierarchicznej wizji dzieła muzycznego. Jednakże teza ta rodzi fundamentalne napięcie wewnątrz samego systemu Jana, stając w pozornej sprzeczności z logiką tezy 10. Wcześniej, aby odróżnić bezpośrednią replikację od konwencji stylistycznej, kluczowa była złożoność memu, rozumiana jako liczba jego komponentów. Wymagało to postrzegania memu jako konkretnej, namacalnej struktury. Tymczasem w tezie 11, mem taki jak pochod 3-2-1 staje się abstrakcyjnym, skalowalnym wzorcem, który może manifestować się zarówno jako kilkudźwiękowy motyw na powierzchni utworu, jak i Urlinie rozciągająca się na całą formę. Następuje tu ontologiczne przesunięcie: mem przestaje być konkretnym 'femotypem', a staje się wirtualnym, relacyjnym 'memotypem', co podważa spójność kryteriów analitycznych stosowanych w poprzednich tezach.

Tę wielopoziomową organizację można jednak zinterpretować w znacznie bardziej dynamiczny sposób, wykorzystując zapożyczoną z biologii rozwojowej koncepcję

skalowalnej architektury kompetencji (multiscale competency architecture, MCA) Michaela Levina (Levin, 2023). Levin argumentuje, że systemy biologiczne składają się z zagnieżdżonych w sobie, agencjalnych/sprawczych (agential) modułów (komórek, tkanek, organów), z których każdy posiada pewne kompetencje do rozwiązywania problemów w swojej skali. Podobnie, memy muzyczne można postrzegać jako elementy wieloskalowej architektury kompetencji kulturowej. Wzór 3-2-1 nie jest jedynie pasywną strukturą, ale funkcjonalnym modułem, który może być wdrażany do rozwiązywania różnych "problemów" kompozycyjnych na różnych poziomach: na poziomie mikro (powierzchni): służy jako lokalny gest melodyczny; na poziomie mezo (planu środkowego): realizuje średnioskalowe połączenie harmoniczne; na poziomie makro (planu dalszego): zapewnia globalne domknięcie tonalne całej formy.

W tym ujęciu hierarchia Jana staje się statyczną mapą dynamicznego procesu, w którym te same kompetentne moduły kulturowe są elastycznie wykorzystywane na różnych poziomach organizacji dzieła. Tak bezpośrednie odniesienie do biologii rozwojowej może wydawać się nadużyciem, zatem wymaga nieco szerszego doprecyzowania. Michael Levin argumentuje, że ewolucja biologiczna nie operuje na pasywnym materiale, lecz na 'materiałach agencjalnych' (agential materials) – komórkach i tkankach, które same w sobie posiadają kompetencje do rozwiązywania problemów (np. regeneracji czy adaptacji do nowych warunków) (Levin, 2023). Przenosząc tę logikę na grunt memetyki, możemy postawić tezę, że memy muzyczne również nie są jedynie pasywnymi "danymi", ale funkcjonalnymi modułami, które kompozytor (jako główny agent) może wdrażać do rozwiązywania konkretnych problemów kompozycyjnych na różnych poziomach skali. Wzór 3-2-1 nie jest zatem tylko abstrakcyjną sekwencją, ale kompetentnym narzędziem, które może służyć zarówno do tworzenia lokalnego gestu melodycznego, jak i do realizowania średnioskalowego połączenia harmonicznego oraz zapewniania globalnego domknięcia tonalnego.

Zgodność hierarchicznych modeli Schenkera i Jana z koncepcją Levina nie jest jedynie powierzchowną obserwacją, że 'zagnieżdżone struktury istnieją tu i tu'. Głęboka spójność leży w funkcjonalnym uzasadnieniu istnienia takiej architektury. Zarówno w biologii, jak i w kulturze, wieloskalowa hierarchia jest efektywnym rozwiązaniem problemu budowania systemów złożonych i odpornych na błędy z prostszych komponentów przy udziale selekcji. Co więcej, dla kompozytora oznacza to również, że nie musi on 'wynajdować koła na nowo' na każdym poziomie organizacji utworu. Zamiast tego, może on sięgać po istniejące, kompetentne moduły (memy i mempleksy) i wdrażać je na odpowiednich poziomach hierarchii. Taki sposób pracy drastycznie upraszcza przestrzeń poszukiwań twórczych, podobnie jak MCA upraszcza przestrzeń poszukiwań dla ewolucji biologicznej.

W tym miejscu sceptyk może wysunąć poważnie brzmiący kontrargument: nie możemy mówić o żadnym funkcjonalnym uzasadnieniu MCA w memetyce muzyki, ponieważ dzieła muzyczne są świadomie oraz intencjonalnie komponowane, a struktura jest każdorazowo ustalana od nowa. Odpowiedź zgodna z retoryką memetyki pierwszej fali byłaby taka, że przecież umysł kompozytora jest w całości złożony z memów, a jego wolna wola jest im całkowicie podległa. Poniżej staram się wykazać, że obie te narracje są częściowo fałszywe, a bliższa prawdy jest ich specyficzna synteza, m.in. dzięki Dennettowskiej koncepcji przestrzeni darwinowskiej (darwinian space) (Dennett, 1995).

Przede wszystkim przytoczony powyżej kontrargument opiera się na fałszywej dychotomii między bezcelową, oddolną memetyczną ewolucją muzyki a świadomym, odgórnym projektem muzycznym kompozytora. Perspektywa memetyczna, w szczególności w ujęciu Daniela Dennetta, pozwala pogodzić te dwa bieguny – potraktować je jako dwa ekstrema ograniczające całe bogate spektrum 'przestrzeni projektowej' (design space) różnorodnych procesów i oddziaływań (Dennett, 1995).

Z jednej strony istnieje oddziaływanie oddolne (bottom-up): sprawczość 'agencjalnych materiałów' (Levin, 2023) w sferze kultury – samych memów i mempleksów. Nie są one bierną gliną w rękach twórcy. Każdy mem, niczym moduł w architekturze Levina,

posiada własne kompetencje: określone tendencje (gitarze basowej często towarzyszy perkusja), potencjał implikacyjny (dominanta często prowadzi do toniki) i zdolność do łączenia się z innymi memami (utwory muzyki techno w secie DJ-a). Kompozytor nie tworzy tych właściwości, lecz je odkrywa i wykorzystuje.

Drugie oddziaływanie Dennett określa jako odgórne (top-down): świadomy, intencjonalny, inteligentny projekt kompozytora. Nie neguję tego, że kompozytor posiada wolną wolę oraz przyznaję, że jego działanie może charakteryzować się autentyczną kreatywnością. Zarazem jednak należy zauważyć, że żaden twórca nie działa w próżni. Kompozytor w trakcie życia przechodzi złożony i głęboki proces enkulturacji, czy też asymilacji zastanej puli memów. Poza tym również skuteczne kreatywne działanie na polu kultury musi charakteryzować się praktyczną znajomością wielu różnych ograniczeń, kodów, norm i umiejętnością kreatywnego dopasowania się do nich. Sam warsztat kompozytorski to złożony mempleks głęboko warunkujący kształt ostatecznych owoców działania autentycznej kreatywności.

W tym ujęciu kompozytor nie jest bogiem-kreatorem tworzącym ex nihilo, lecz potężnym agentem selekcyjnym, który operuje wewnątrz ekosystemu memetycznego. Jego 'odgórne' decyzje są głęboko uwarunkowane i ograniczone przez 'oddolne' właściwości dostępnych mu memów. Zatem zastosowanie modelu skalowalnej architektury kompetencji (MCA) do muzyki jest w pełni uzasadnione. Zagnieżdżona, hierarchiczna struktura dzieła jest wynikiem dynamicznego sprzężenia zwrotnego między kompetencjami samych memów a odgórnym, selekcyjnym działaniem umysłu kompozytora, który sam jest produktem koewolucji genetyczno-memetycznej (Richerson, 2010).

Zatem naturalnie dochodzimy do wniosku, że całkowite ontologiczne odseparowanie konkretnego memu tylko na jednym poziomie funkcjonowania złożonej struktury memetycznej od innych elementów głęboko powiązanych z tym samym memem, ale na innych poziomach dzieła, może być nie tylko trudne, ale także prowadzące do fałszywych wniosków lub nawet niemożliwe.

Wcześniej wymienione trudności teoretyczne prowadzą do ważnego, nieobecnego w pracy Jana wniosku metodologicznego: funkcjonalna analiza memów muzycznych musi być z natury zależna od skali i obserwatora. Zamiast poszukiwać jednej, uniwersalnej odpowiedzi na nierozstrzygalne ontologiczne pytania, takie jak "jaki jest najmniejszy możliwy mem?" lub "jak odróżnić złożony mempleks od zbioru prostych memów?", należy przyjąć perspektywę pragmatyczną. Każde badanie memetyczne powinno rozpoczynać się od lokalnej i operacyjnej definicji memu, jasno określając kryteria jego identyfikacji i segmentacji w ramach przyjętej skali analitycznej. Takie podejście, zamiast wikłać się w problemy filozoficzne, pozwala na prowadzenie rygorystycznych i falsyfikowalnych badań nad konkretnymi zjawiskami transmisji kulturowej.

12. Memetics acknowledges the explanatory force of Schenkerian structural-hierarchic transformations, and the resultant structural levels. It also acknowledges the utility of Schenkerian graphical procedures in the representation of the hierarchic location of memes. Nevertheless, memetics asserts that structural levels are generated by bottom-up processes of meme conglomeration, not top-down control from the Ursatz. (s. 228)

Memetyka uznaje wyjaśniającą moc Schenkerowskich transformacji strukturalno-hierarchicznych i wynikające z nich poziomy strukturalne. Uznaje również użyteczność Schenkerowskich procedur graficznych w reprezentacji hierarchicznego położenia memów. Tym niemniej, memetyka zapewnia, że poziomy strukturalne wynikają z oddolnych procesów konglomeracji memów, a nie odgórnej kontroli poprzez Ursatz.

Jan w punkcie 12 skupia się na podkreśleniu wagi oddolnej generacji struktur (bottom-up) w procesie ewolucyjnym, co stanowi fundament memetyki w ujęciu Daniela Dennetta. Dennett (1995) postrzega złożone systemy, od ludzkiej świadomości po kulturę, jako emergentne rezultaty działania prostych, 'samolubnych' agentów – samolubnych neuronów zasiedlonych przez samolubne memy (Michael Levin całkowicie zgadza się z tą perspektywą i dodatkowo ją rozszerza aż do najprostszych jednokomórkowych organizmów (Levin, 2020)).

Funkcjonalność wyłania się nie z odgórnego planu, lecz z oddolnych interakcji replikatorów, ponieważ wszystkie naturalnie dążą do własnego powielenia. Taka perspektywa ‘z poziomu memu’, a nie kompozytora, jest istotą propozycji Jana: odczytania hierarchii Schenkerowskiej jako skutku, a nie przyczyny, muzycznej złożoności.

Tym niemniej takie radykalne przeciwstawienie procesów oddolnych i odgórnych bez dodatkowego komentarza jest zbyt daleko idącym uproszczeniem. W komentarzach do wcześniejszych tez Jana starałem się wykazać, że w środowisku muzycznym kompozytor funkcjonuje jako potężny mechanizm selekcyjny operujący na ‘agencjalnych materiałach’ kultury – memach zagnieźdżonych w wieloskalowej architekturze kompetencji. Rozproszona sprawczość istnieje zatem jednocześnie na wielu poziomach tej hierarchii.

Co więcej, sam Jan w innych tezach przyznaje moc sprawczą ‘dialektowi’ czy ‘wspólnej walucie memetycznej’ (Teza 10). Ta koncepcja odgórnie ograniczającego stylu jest funkcjonalnym odpowiednikiem Schenkerowskiego *Ursatz* (Schenker, 1979) – oba stanowią abstrakcyjną, głęboką strukturę, która kanalizuje procesy na niższych poziomach. Zatem, choć teza 12 trafnie odwraca tradycyjną perspektywę, jej kategorię sformułowanie pomija złożoność pętli sprzężeń zwrotnych między oddolną konglomeracją memów, odgórnym projektem kompozytorskim oraz ograniczającym polem możliwości samego stylu.

13. A given meme is almost invariably a multiparametric complex in the parameters of pitch and rhythm. At the next structural-hierarchic level, several memes may be propagated as a memplex, provided each constituent meme satisfies the condition of independent replication. Memplexes may exist recursively at several structural-hierarchic levels, the large-scale formal archetypes of western music being structural memes constituting replicated patterning at the deepest levels of organization. (s. 228)

W zdecydowanej większości przypadków mem muzyczny jest wieloparametrycznym kompleksem w dziedzinach wysokości dźwięku oraz rytmu. Na wyższym poziomie strukturalno-hierarchicznym, kilka memów może wspólnie propagować się jako mempleks, jeśli każdy z nich jest zdolny do samodzielnej replikacji. Mempleksy mogą istnieć rekurencyjnie na wielu poziomach strukturalno-hierarchicznych:

wielkie, formalne archetypy zachodniej muzyki artystycznej są strukturalnymi memami, które wymuszają wzorce replikacji na najgłębszych poziomach organizacji.

Teza 13 ma donośne znaczenie dla całej memetyki Jana. Autor wprowadza tu niezbędne narzędzia do realizacji jego ambitnego, nadrzędnego celu: napisania wielopoziomowej historii muzyki artystycznej z perspektywy samego memu (meme's eye view).

W tym celu centralną kategorię stanowi mempleks. Zdaniem autora jego funkcjonowanie rozwiązuje problem prostoty i złożoności, który omawialiśmy przy tezach 7 i 10. Przykładowo, kilka prostych, zdolnych do samodzielnej replikacji muzycznych memów (np. motywów) oddolnie łączy się w stabilny, symbiotyczny sojusz, który sam staje się jednostką replikacji wyższego rzędu (np. frazą). Taki proces, działając rekurencyjnie, pozwala na budowanie całej hierarchii muzycznych struktur – od najmniejszych figur, przez schematy formalne, aż po memy strukturalne najwyższego rzędu. Jan nie pisze, jak wyglądają te najgłębsze poziomy organizacji, ale możemy skonceptualizować je nawet jako intelektualne, racjonalne podstawy mempleksu zachodniej muzyki artystycznej, takie jak zasady harmonii czy cała muzykologia i refleksja o muzyce.

Niestety również w tym punkcie Jan przypisuje sprawczość wyłącznie najprostszym memom na samym dole struktury hierarchicznej. Przy poprzednich tezach usiłowałem już wyjaśnić, dlaczego takie spojrzenie musi być fałszywe. Ponadto interesująca koncepcja teoretyczna zawarta w punkcie 13 kryje w sobie ogromne praktyczne wyzwanie metodologiczne: rygorystyczny warunek 'niezależnej replikacji'. Aby udowodnić istnienie mempleksu, analityk musiałby empirycznie wykazać, że każdy jego składnik funkcjonuje w kulturze również samodzielnie, co przy złożonych mempleksach wymaga przeprowadzenia badań na niemal niewyobrażalną skalę.

Teza 13 wyjaśnia rolę mempleksu i tym samym stanowi mocny punkt spajający teorię Jana. Tym gorzej, że zawiera ona znaczne metodologiczne obciążenie: pozostawia

kluczowe rozróżnienie między złożonym memem a mempleksem w sferze trudnej do zweryfikowania hipotezy. Zarazem warto wspomnieć, że ten praktyczny problem może być w przyszłości rozwiązany przez zautomatyzowane analizy oparte na uczeniu maszynowym.

14. At shallow middleground levels, certain memplexes may function as perceptual-cognitive schemata, acting as frameworks for the organization of foreground-level (feature) memes.

W płytkiej warstwie środkowej, pewne mempleksy mogą funkcjonować jako schematy poznawcze, wyznaczając ramy dla organizacji memów w warstwie powierzchniowej. (s.228)

W podrozdziale 4.3.2 Jan omawia różne rodzaje memetycznych mutacji występujących w przykładach nutowych. Mówi tam o dodawaniu, odejmowaniu, zamienianiu oraz modyfikacji - te operacje odnoszą się do chromatycznej wysokości dźwięku oraz rytmu i można przyrównać je do różnych rodzajów wariacji, jakie występują w krótkich tematach lub innych strukturach melorytmicznych w muzyce klasyczno-romantycznej.

Badacz zauważa również, że w niektórych przypadkach można mówić o czymś więcej niż mutacji pojedynczych parametrów - o imitacji całej struktury brzmieniowej z nieznacznymi zmianami w kilku miejscach. Taką operację można przyrównać do replikacji Schenkerowskiej warstwy środkowej z jednego utworu do drugiego i wypełnienia jej ozdobnikami w nieco inny sposób. Ponadto, teza 14 ma ugruntować Janowską interpretację Schenkerowskiej strukturalnej teorii w psychologii poznawczej, łącząc Janowską wizję muzycznego mempleksu z teorią schematów (Rumelhart, 1980). Sam Jan wspomina, że w muzykologii została ona rozwinięta przez takich badaczy jak Gjerdingen (1988) czy Meyer (1956) (Jan, 2007).

W tym ujęciu często replikowane 'sojusze memów' funkcjonują jako mentalne ramy, które organizują materiał na powierzchni utworu poprzez tworzenie schematów poznawczych. Jednakże ta elegancka koncepcja rodzi dwa poważne problemy. Po pierwsze, zakłada kierunek przyczynowości: czy to na pewno mempleksy, poprzez

częstą replikację, tworzą w umyśle schematy? A może to wrodzone schematy faworyzują replikację pasujących do nich mempleksów? Perspektywa koewolucyjnego pochodzenia ludzkiej muzykalności (Savage, 2021) sugeruje, że i w tym wypadku przyczynowość rozprasza się w sprzężeniu zwrotnym pomiędzy oddolnym oddziaływaniem memów i odgórnym oddziaływaniem ucieleśnionego umysłu.

Po drugie, badacz naraża się na zarzut tautologii, ponieważ 'mempleks' może być postrzegany jedynie jako nowa, darwinowska etykieta dla dawno opisanego w psychologii 'schematu' (Rumelhart, 1980), bez wnoszenia nowej, falsyfikowalnej treści. Taką falsyfikowalną, nową treścią byłoby zaproponowane powyżej koewolucyjne wyjaśnienie owych schematów poznawczych. Kategoria mempleksu byłaby więc użyteczna, żeby przeprowadzić odpowiednie wnioskowanie w duchu Darwinowskim na temat muzycznych schematów poznawczych.

15. The memetic-structural hierarchies formed by memes and memplexes in intraopus contexts are extended by memetic-cultural hierarchies encompassing interopus contexts, these forming a continuum from the single, localized meme to such widely chronologically and geographically distributed entities as dialects. (s.228)

Hierarchie memetyczno-strukturalne wytworzone przez memy i memplexy w kontekstach pomiędzy dziełami są rozszerzone przez memetyczno-kulturalne hierarchie obejmujące konteksty pomiędzy dziełami. Wspólnie tworzą one kontinuum, które zaczyna się na pojedynczym memie, a kończy na szerokich, rozdystrybuowanych w czasie i przestrzeni bytach takich jak dialekty.

Teza 15, która ma za zadanie formalne połączenie hierarchii wewnątrzdziałowej i międzycziałowej, w dużej mierze powieliła problemy poruszone już we wcześniejszych punktach - do tego nie jest z nimi spójna logicznie. Kwestię odróżnienia konkretnego memu od ogólnego tła stylistycznego – czyli tutaj 'dialektu' – podjęliśmy już w tezach 9 i 10. Kryterium diagnostycznym była złożoność i cognitive salience. Największy problem stanowi jednak fundamentalna niespójność z radykalną perspektywą bottom-up z Tezy 12. Tam Jan kategorycznie odrzucał odgórną kontrolę Ursatz. Twierdził, że hierarchie są wyłącznie emergentnym skutkiem oddolnej konglomeracji memów. Tymczasem w Tezie

15, opisując 'dialekty' jako byty rozdystrybuowane w czasie i przestrzeni, niejawnie przyznaje im moc sprawczą – zdolność do odgórnego kształtowania praktyk muzycznych.

Z perspektywy łączącej memetykę z koncepcją skalowalnej architektury kompetencji (MCA), takie odgórne oddziaływanie jest nie tylko możliwe, ale i konieczne dla zrozumienia złożonych systemów. Problemem nie jest więc sama idea sprawczości dialektu, lecz wewnętrzna sprzeczność w modelu Jana, który nie potrafi spójnie pogodzić swojej dogmatycznej wizji bottom-up z obserwowalną rzeczywistością odgórnych ograniczeń kulturowych. Choć Jan deklaratorywnie jest purystą 'bottom-up', struktura jego własnej teorii – szczególnie ze względu na pojęcia takie jak 'wspólna waluta memetyczna' i 'hierarchia kulturowa' – logicznie wymusza istnienie odgórnej presji kształtującej. Dialekt w systemie Jana musi posiadać sprawczość. W przeciwnym razie kluczowe elementy jego teorii straciłyby moc wyjaśniającą.

16. The fundamental model for meme dynamics is that of gene dynamics: the tripartite algorithm - variation, heredity/replication and differential fitness - of Darwinian evolution by selection. (s. 228)

Fundamentalny model dynamiki w memetyce jest jednakowy jak w genetyce. Jest to trójdzielny algorytm (różnorodność, dziedziczność/replikacja i dostosowanie) Darwinowskiej selekcji naturalnej.

Punkt 16 ogólnie opisuje ewolucję memów, powołując się na cechy replikatorów opisane przez Uniwersalny Darwinizm (Dawkins, 1983). Z samym tym stwierdzeniem bezwzględnie należy się zgodzić. Być może pod względem retorycznym nieco mylące jest to, że Jan nie wspomina tutaj Uniwersalnego Darwinizmu, a raczej pisze o podobieństwie memów do genów. Czytelnik może odnieść wrażenie, że Jan sugeruje, że aby zrozumieć cechy memów, należy obserwować cechy genów i bezpośrednio je przenosić. Takie myślenie jest błędne. Ewolucja genetyczna oraz ewolucja memetyczna to dwie instancje uniwersalnego procesu selekcyjnego, który w każdym medium może przebiegać inaczej. Ewolucja, cechy, 'zachowanie' memów i genów różnią się od siebie pod wieloma względami.

17. The transmission of memes within cultural communities is analogous to the spread of infections between organisms, allowing memetics to draw fruitfully from the conceptual vocabulary and methodologies of epidemiology. (s. 228)

Rozchodzenie się memów w społeczeństwie przebiega analogicznie do roznoszenia infekcji pomiędzy organizmami, dzięki czemu memetyka może owocnie czerpać słownictwo oraz metodologie z epidemiologii.

Analogia epidemiologiczna wprowadzona w punkcie 17 jest bez wątpienia wartościowa. Dała nam wartościową poznawczo metaforę "wiralności" będącą od wielu lat w popularnym użyciu (Ling, 2021; Feltman, 2025). Tłumaczy ona dynamikę internetowych memów i tiktokowych hitów. Jan stawia jednak model epidemiologiczny w napięciu do bronionego wcześniej modelu genetycznego. Tymczasem relacja między wirusem a genem jest w biologii złożona. Wirusy to w istocie pakiety informacji genetycznej, obdarzone pasywną sprawczością, pasożytujące na maszynerii genetycznej gospodarza (Harris, 2021). Sprawczość wirusów jest bardzo bliska do sprawczości memów i pod tym względem bez wątpienia warto pogłębiać Dennettowską koncepcję memów jako wirusów umysłu (Dennett, 1995).

Co więcej, nawet gdyby Jan wprowadził epidemiologię bez napięcia z genetyką, to nie wystarczyłoby do opisu dynamiki replikacji memów. Autor pomija w swojej teorii trzecią, kluczową siłę: ewolucję lamarckowską memów (Kronfeldner, 2007). W kulturze zmiany nabyte przez jednostkę są dziedziczone – zmodyfikowana idea jest przekazywana dalej w nowej formie. Podkreśla to tylko, że ewolucja memetyczna nie jest idealnym odzwierciedleniem genetycznej, co sugeruje Jan. Memy zachowują się trochę jak geny, a trochę jak wirusy. Posiadają też własne, unikatowe mechanizmy transmisji, jak dziedziczenie lamarckowskie czy transmisja horyzontalna. Przyszła pełna teoria replikacji memów musi uwzględniać całą tę złożoność.

18. Those memes that have the greatest perceptual-cognitive salience will tend to be the most widely propagated in the meme pool, to the disadvantage of those less salient forms. The implication-realization

potential (i-rp) is posited as a means of quantifying this fecundity, offering a necessary but not sufficient index of salience. (s.228)

Te memy, które cechują się wysoką wyrazistością poznawczą (perceptual-cognitive salience) zwykle propagują się w puli memowej znacznie lepiej niż reszta. Potencjał implikacji-realizacji (i-rp) został stworzony jako środek do mierzenia tego rodzaju płodności. Oferuje on niezbędny, oraz niestety niewystarczający wskaźnik wyrazistości.

Teza 18 stanowi próbę stworzenia przez Jana (2007) mierzalnego wskaźnika memetycznego sukcesu poprzez pojęcie cognitive salience, kwantyfikowane za pomocą wskaźnika implikacji-realizacji (i-rp). Niestety, po pierwsze, w jego analizach wskaźnik i-rp odnosi się przede wszystkim do analizy zapisu nutowego muzyki z okresu klasyczno-romantycznego. Nie zbadamy więc przy jego pomocy nagrań audio lub muzyk pozaeuropejskich nawet w ich własnych systemach notacji. Jan jednak zdaje sobie sprawę z tych ograniczeń i proponuje schemat i-rp jako wersję podstawową, na bazie której można rozwinąć bardziej skomplikowane i uniwersalne systemy. Autor asekuracyjnie określa go jako warunek "konieczny, ale niewystarczający", co jednak stanowi istotne niedopowiedzenie.

Nawet gdyby rozwinąć schemat i-rp na nagrania muzyczne, to nie jest to wystarczający mechanizm do mierzenia sukcesu replikacyjnego. Wykazuje to analiza współczesnych ekosystemów memetycznych, od internetu po muzykę popularną. Najbardziej płodne memy nie są najbardziej wyraziste poznawczo, lecz opierają się na strategii, którą można nazwać "kradzieżą uwagi" (attention theft) (Monge-Roffarello, 2022; McFedries, 2014). Wykorzystują one optymalny balans między tym, co znane, a tym, co nowe: osadzają w wysoce konwencjonalnych, łatwych w przetwarzaniu strukturach jeden, niezwykle skuteczny i nieprzewidywalny element – "haczyk" (hook). W architekturze mediów społecznościowych jako mechanizm haczyka dodatkowo projektuje się pętle czasowej inwestycji i poznawczej nagrody (Lukyanchikova, 2023).

Taki model, zgodny z teoriami percepcji opartymi na niepewności i zaskoczeniu (np. u Marcusa Pearce'a (Pearce, 2018)), wyjaśnia, dlaczego największe hity często przypominają starsze utwory, a jednocześnie "kradną" uwagę słuchacza. W tym świetle system Jana, prawdopodobnie obciążony historycznym kontekstem analizowanej przez niego muzyki klasyczo-romantycznej, błędnie utożsamia ogólną wyrazistość z precyzyjnie zaprojektowanymi narzędziami, które faktycznie warunkują masową replikację w kulturze.

19. The metaphors of selfishness and the meme's eye view articulate the tautological inevitability at the heart of replicator theory. Wherever competition is the key to survival, those entities which compete most effectively will be the most likely to survive - be they animate creatures 'red in tooth and claw' or musical patterns 'looking' for a niche in memory space. (s. 229)

Metafory samolubności oraz widoku z perspektywy memu wyrażają tautologiczną konieczność leżącą w sercu teorii replikatorów. Zawsze, kiedy konkurencja jest niezbędna do przetrwania, te byty, które najbardziej efektywnie konkurują, mają największą szansę na przetrwanie. Dotyczy to w takim samym stopniu biologicznych organizmów, jak i fragmentów muzycznych 'szukających' miejsca w pamięci słuchaczy.

Teza 19 jest całkowicie zgodna z najgłębszymi twierdzeniami memetyki. W związku z tak wielką wagą warto ją doprecyzować. Metafora samolubności, choć potężna, jest również ryzykowna, ponieważ może nadmiernie upraszczać i maskować złożoność procesów selekcyjnych w kulturze. "Samolubność" powierzchownie odczytana sugeruje, że cała sprawczość leży po stronie replikatora, podczas gdy jego sukces jest w rzeczywistości własnością relacyjną, wynikającą z dynamicznego dopasowania między strukturą memu a jego środowiskiem – czyli ludzkim umysłem i zastaną kulturą. Mówienie o "walce o niszę w przestrzeni pamięci" (Jan, 2007) może przywołać na myśl prostą grę o zerowej sumie. A jednak, jak ustaliliśmy wcześniej, relacje między memami często mają charakter symbiotyczny i kooperatywny. Memy nie tylko konkurują, ale przede wszystkim tworzą złożone, współzależne ekosystemy (mempleksy), których nie da się w pełni zrozumieć przez pryzmat wyłącznie indywidualnej "samolubności". Warto

zatem uzupełnić tezę 19 o perspektywę ekologiczną, która uwzględnia nie tylko naturę replikatora, ale i złożoność środowiska, w którym musi on przetrwać.

Doprecyzowania wymaga też koncepcja *samolubności* memu (i genu), ponieważ sugeruje, że memy mogą *zachowywać się* samodzielnie lub posiadać prawdziwe intencje. Bez dodatkowego komentarza byłoby to niepoprawną antropomorfizacją. Warto jednak zauważyć, że jeśli pewien mem będzie obdarzony czymś w rodzaju instrukcji 'zreplikuj mnie', która skutecznie dotrze do podświadomego umysłu odbiorcy, wtedy taki mem ma największe szanse się rozprzestrzenić. Na drodze doboru naturalnego i pokrewnych procesów darwinowskich te memy, które nie dążą do replikacji, powoli znikają z kultury i ustępują miejsca tym, które o nie 'zabiegają'. Możemy więc użytecznie myśleć o memach tak, jakby faktycznie kierowały się pragnieniem replikacji.

Pomocnym narzędziem eksplanacyjnym jest tu koncepcja "dryfujących przyczyn" (free-floating rationale) Daniela Dennetta (Dennett, 1995). W procesie ewolucji nawet losowe mutacje filtrowane przez dobór naturalny prowadzą do powstawania struktur, które działają tak, jakby realizowały cel – w tym przypadku, własną replikację. Perspektywa memetyczna sugeruje, że jako wynik działania algorytmu procesu selekcyjnego, w toku ewolucji zaczyna wyłaniać się celowość, teleologiczność. Nawet losowe mutacje w selekcyjnym środowisku, z działaniem doboru naturalnego, zaczynają kształtować takie struktury, które coraz bardziej wyglądają, jakby do czegoś dążyły (Dennett, 2017, Levin, 2025). W ten sposób w memach stopniowo wyłania się pewna sprawczość, aż na drodze częściowo darwinowskiej ewolucji dochodzimy do mocno reaktywnych cybernetycznych systemów oraz sztucznej inteligencji, która właśnie na drodze memetycznej ewolucji stopniowo zyskuje coś, co coraz mocniej aproksymuje realne dążenie do celów. Inteligentną manipulację otoczeniem w celu optymalizacji własnej replikacji. Po prostu: zachowanie (Floridi, 2025; Azadi, 2025; Stanford HAI, 2025, Bandi, 2025).

20. Parallelling the generation of patterns at higher levels of memetic-structural hierarchies by memes at lower levels, mutation and selection at lower levels of memetic-cultural hierarchies cumulatively drive changes to phenomena at higher levels. The most powerful manifestations of these processes are the seamless, gradualistic metamorphosis of dialects and, more abstract still, the transmutation of systems of tonal organization. (s.229)

Podobnie jak memy niższego rzędu generują wzorce na wyższych poziomach hierarchii memetyczno-strukturalnych, tak mutacja i selekcja na niższych poziomach hierarchii memetyczno-kulturowych kumulatywnie napędzają zmiany w zjawiskach na poziomach wyższych. Najpotężniejszymi przejawami tych procesów są płynne, stopniowe metamorfozy dialektów oraz – w jeszcze bardziej abstrakcyjnym wymiarze – transmutacje systemów organizacji tonalnej.

Jan zwraca uwagę na istnienie wzajemnych memetycznych relacji pomiędzy wszystkimi poziomami kultury muzycznej. W takiej perspektywie możemy zobaczyć jasno to, że teoria muzyki, muzykologia, czy krytyka muzyczna mają twórczy wpływ na kształtowanie bieżącej muzyki. Warto zauważyć, że może to być powiązane z problemem racjonalizacji zachodnioeuropejskiej muzyki artystycznej, który to został oryginalnie dostrzeżony przez niemieckiego socjologa Maxa Webera (Weber, 1958).

Teza 20 stanowi trafne zwięźczenie darwinowskiego projektu Jana, opisując, jak kumulatywne zmiany na poziomie mikro (mutacje pojedynczych memów) napędzają makroewolucję kultury muzycznej – od płynnej metamorfozy "dialektów" aż po transmutację całych systemów tonalnych. Główny zarzut wobec tej tezy nie dotyczy jednak jej treści, lecz jej umiejscowienia i deklaratywnej niespójności z wcześniejszym, często dogmatycznie oddolnym, podejściem autora. Przez znaczną część książki Jan kładł nacisk na radykalną sprawczość bottom-up (Teza 12), negując lub pomijając odgórne wpływy. Tymczasem tutaj de facto przyznaje istnienie złożonej pętli sprzężenia zwrotnego, w której wszystkie poziomy hierarchii wzajemnie na siebie oddziałują. Taki model jest w pełni zgodny z postulowaną przeze mnie perspektywą skalowalnej architektury kompetencji (MCA), gdzie sprawczość jest rozproszona na wielu poziomach. Problem polega na tym, że Jan dochodzi do tego spójnego, interaktywnego obrazu dopiero w podsumowaniu oraz nie integruje go konsekwentnie ze swoimi

wcześniejszymi, bardziej kategorycznymi tezami. W ten sposób, przychylnie go odczytując, osłabia ogólną spójność swojej argumentacji. Z drugiej strony, można nawet zinterpretować jego stwierdzenia o sprawczości jako sprzeczne.

2.6. Kontynuator Jana – Andrew Hawkett

W kontekście memetyki muzyki Stevena Jana, jej oddziaływania i stabilizacji, warto omówić rozprawę doktorską, która powołuje się na niego bezkrytycznie. Andrew Hawkett, doktorant Jana, w swojej pracy *Badanie empiryczne koncepcji memu w muzyce przy użyciu masowej analizy danych kwartetów smyczkowych* (An Empirical Investigation Of The Concept Of Memes In Music Using Mass Data Analysis Of String Quartets) (Hawkett, 2013) wykorzystuje model Jana w praktyce. Za pewną krytykę czy też modyfikację wyrażoną nie wprost można uznać to, że Hawkett powołuje się tylko na niektóre aspekty teorii Jana. Nie wyraża wprost żadnych krytycznych wątpliwości, ale po prostu znaczna część teorii jest dla jego obliczeń całkowicie nieistotna. W każdym razie nie jest w znaczący sposób zawarta w systemie generującym obliczenia.

Autor wprost pisze, że rozwija model Jana (s.2). A zatem już na samym początku powstaje poważny problem – Jan wielokrotnie deklaruje, że w jego systemie faktyczne memy to idee czy też wzorce pamięciowe w umyśle człowieka. Partytury zatem, choć Jan sam wielokrotnie podkreśla ich znaczenie dla swojego systemu, stanowią w nim zaledwie materialny artefakt, produkt uboczny istnienia memów. Sam Jan określa zatem takie materialne artefakty jako femotypy, produkty femotypiczne. Hawkett zatem, według kryteriów systemu Jana, nie analizuje samych memów, a zaledwie materialne produkty ich istnienia. A zatem już sam tytuł pracy – empiryczne badanie *koncepcji memu muzycznego* wprowadza czytelnika w błąd, albo po prostu jest nieadekwatny do treści pracy.

Co więcej, w całej pracy Hawketta takie określenia jak femotyp (phemotype), femotypiczny (phemotypic) nie pojawiają się ani razu. Autor nie tylko uchyla się od komentarza na temat tego głównego aspektu ontologii memów, ale nawet nie wspomina o jego istnieniu. Nie można tego w żaden spójny logicznie sposób pogodzić z tym, że Hawkett deklaruje, że rozwija model Jana (s.2). Empiryczne, a także obliczeniowe aspekty pracy stanowią zdecydowanie jej największą wartość. Zarazem powstają dwa

fundamentalne pytania – co autor de facto oblicza, oraz co deklaruje, że oblicza?

Hawkett wyszukuje różnego rodzaju podobieństwa pomiędzy wzorcami wyodrębnionymi za pomocą analizy z treści ponad 400 części ponad 100 kwartetów smyczkowych. A zatem na pewno dokonuje analizy muzycznej w duchu ewolucyjnym. Obliczenia mają posłużyć jako empiryczny dowód słuszności systemu memetyki muzyki Jana – ale nie mogą tego osiągnąć z podstawowych przyczyn. Po pierwsze dlatego, że sama teoria wymieniona w tej pracy jest z systemem Jana niespójna. Po drugie, ponieważ nawet jeśli Hawkettowi udałoby się wykazać zmienność ewolucyjną materiału muzycznego, to samo w sobie nie dowodzi jeszcze, że jest to ewolucja kulturowa akurat w koncepcji memetycznej, a nie na przykład w alternatywnej conceptualizacji Richersona i Boyda. Hawkett podejmuje według mojej oceny trafną decyzję, żeby zdystansować się od problemów filozoficznych. Niestety w praktyce nie do końca mu się to udaje. Autor deklaruje, że na potrzeby swoich badań przyjmuje konkretną definicję memu. Ta deklaracja wymaga dodatkowego komentarza. Hawkett pisze:

Zatem robocza definicja muzycznego memu użyta w niniejszym badaniu to: Dowolny ciąg od trzech do jedenastu nut, wykluczając ornamentację symboliczną i dodatkowe parametry, który ewoluje na drodze doboru naturalnego (tj, selekcji, replikacji i wariacji), i posiada atrybuty replikatora – długowieczność, płodność i wierność kopiowania).

Consequently, the working definition of a meme in music used in the present study was:

Any three to eleven monophonic consecutive notes, excluding symbolic ornamentation and secondary parameters, that evolves by means of natural selection (i.e., selection, replication and variation), and which exhibits the replicator properties of longevity, fecundity and copying-fidelity.

Autor z jednej strony podaje ścisły obliczeniowy wzór, który stanowi sedno definicji memu dla jego komputerowego systemu – ciąg od 3 do 11 dźwięków bez dodatkowych oznaczeń. Niestety jednak na tym się nie zatrzymuje i obarcza swoją definicję też post-Janowskim chaotycznym balastem. Znacznie bardziej fortunate byłoby, gdyby autor rozdzielił dwa splecione tutaj zadania. Po pierwsze sformułowanie ontologicznej definicji memu muzycznego, a po drugie określenie przyjętej konkretnej obliczeniowej miary, skali rozmiaru memów obserwowanych w danym badaniu.

Przyjęcie odpowiedniej skali i jasne jej określenie to moim zdaniem konieczne posunięcie w każdej empirycznej analizie memetycznej tego typu. Po pierwsze jasno określamy przedmiot badania, a po drugie każdorazowo rozwiązujemy problem jednoczesnej wieloskalowości memów. Z drugiej strony pozostaje nam poważny problem ontologiczny. W pracy, która według tytułu ma zajmować się empirycznym badaniem *konceptji memu muzycznego* zdefiniowanie go jako ciągu melodycznego, który ewoluuje wyłącznie na drodze doboru naturalnego to poważny błąd.

Od samego początku oryginalna darwinowska koncepcja doboru naturalnego była przeciwstawiona doborowi sztucznemu (Darwin, 1859), który przede wszystkim odróżnia się celowością, odgórną kontrolą człowieka (Gregory, 2009). Oba doборы wywołują także inne efekty. Stwierdzenie bez dodatkowego komentarza, że muzyka ewoluuje na drodze doboru naturalnego to radykalne odmówienie sprawczości kompozytorowi, radykalne odmówienie wolnej woli słuchaczy, kuratorów i ludzi w jakikolwiek sposób zaangażowanych w kulturę muzyczną. Ewolucję memetyczną muzyki musimy rozumieć jako proces o rozproszonej sprawczości, która wynika ze złożonej pętli zwrotnej pomiędzy agencjalną treścią muzyczną i sprawczym selekcyjnym agentem, którym w XXI wieku nadal pozostaje człowiek.

Należy założyć, że autor nie zdawał sobie sprawy z tego, że konsekwencje logiczne jego słów prowadzą do takich wniosków. W punkcie 8.6.1. stwierdza, że jego praca nie ma ustalać tego, co pobudza proces ewolucyjny muzyki, tylko ustalić, czy ma on miejsce. Nie możemy jednak na podstawie tych słów wycofać wcześniejszej krytyki tylko dlatego, że autor sam sobie przeczy. Musimy raczej wyciągnąć wniosek, że nie wykonał odpowiedniego starania, aby zrozumieć proces doboru naturalnego, który ochoczo stawia w centrum swojego dochodzenia. Dodatkowo ostateczną racjonalizację swojego wyboru definicji memu Hawkett przeprowadza następująco:

Definicja bazuje na założeniu Millera, że mózg segreguje informacje w jednostki liczące siedem, plus minus dwa składniki (1956), razem z hipotezą Jana, że Mem nie może być krótszy niż trzy nuty, żeby miał wystarczającą zawartość informacyjną, aby był wyrazisty i wyróżnialny. A zatem wzorzec muzyczny, na

potrzeby niniejszego badania, definiujemy jako ciąg od trzech do jedenastu nut, wyłączając symboliczną ornamentację i parametry drugorzędne.

The definition was based on Miller's assertion that the brain chunks information into 'seven plus or minus two' units of information (1956), together with Jan's hypothesis that a meme cannot be shorter than three notes, in order to provide it with enough information to be distinct and salient (2007, pp. 60-61). Therefore the pattern in music, for the purposes of the present research, was defined as any three to eleven monophonic consecutive notes, excluding symbolic ornamentation and secondary parameters.

Takie wytłumaczenie do przyjęcia najważniejszej definicji w pracy, która ma badać koncepcję memu muzycznego, budzi wiele wątpliwości.

Dolna granica ciągu od 3 do 11 nut jest wyznaczona przez krytykowaną przeze mnie koncepcję Jana, natomiast górna granica -11- nie posiada żadnego racjonalnego wyjaśnienia. Autor powołuje się na nieco intuicyjną koncepcję siedmiu plus minus dwóch elementów w pamięci (Miller, 1956), ale siedem plus dwa to wciąż 9, a nie 11. Trzeba zatem założyć, że przyczyną jest to, że przy wybraniu liczby 11 przebadanie zebranej do tego badania bazy danych po prostu dało bardziej wymierny rezultat.

Działanie autora możemy skonceptualizować w następujący sposób: odnajdźmy wzorzec w partyturach, określmy prawdopodobieństwo występowania wzorców o określonych stopniach podobieństwa do tego wzorca, interpretując tę relację jako relację ewolucyjną. Następnie przyjmijmy dla nich jakąkolwiek nazwę, po to, żeby nadać etykietę naszym obliczeniom i w jakiś sposób je zinterpretować. Tym samym doktorant Jana wpada w tę samą pułapkę co on – przeprowadza badania klasycznej analizy muzycznej i post factum dopasowuje je do ewolucyjnych teorii w ramach filozoficznej interpretacji. Tym niemniej przy całej powyższej krytyce, która odnosi się do kształtowania pojęć i interpretacji, wciąż można docenić dokonane w pracy badania ilościowe i odnalezione wzorce ewolucji kulturowej w kwartetach smyczkowych. Dla czytelnika, który nie jest wobec memetyki sceptyczny, te badania stanowią doskonały przykład i demonstrację procesu memetycznej czy też kulturowej ewolucji materiału muzycznego.

3.1. Autorska koncepcja systemu muzycznego

W muzykologii termin „system” tradycyjnie odnosi się do wielu często rozłącznych zjawisk: systemu tonalnego i harmonicznego (Strunk, 2003), systemu notacji (Bent, 2001), klasyfikacji instrumentów (Lee, 2020), systemu edukacji czy wreszcie układu pionowych akordów w partyturze (Strunk, 2003). W tym znaczeniu system bywa pojmowany intuicyjnie – jako uporządkowany zbiór zasad czy elementów. W niniejszej pracy przyjmuję jednak ujęcie odmienne. Interesuje mnie system muzyczny jako dynamiczny, sprawczy, ewoluujący układ złożony z ludzi i kultury muzycznej.

Zatem na potrzeby niniejszej pracy rozumiem system muzyczny jako częściowo zamknięty, ewoluujący układ przepływu informacji, w którym krzyżują się różne siły: ludzka kreatywność, presje selekcyjne i ewolucyjny design. W jego ramach – zgodnie z dennettowską zasadą kompetencji bez zrozumienia (Dennett, 1995) powstają sprawcze struktury kulturowe, zdolne do podtrzymywania i reprodukcji częściowo niezależnie od intencji zaangażowanych jednostek. Taki system muzyczny obejmuje także procesy enkulturacji, w których człowiek uczy się obowiązujących reguł, a zarazem sam zostaje przez nie kształtowany w pętli zwrotnej. Dzięki temu system nie tylko organizuje praktykę muzyczną, lecz także wytwarza środowisko selekcyjne dla dalszej ewolucji memów i mempleksów.

Muzyczne memy i mempleksy, takie jak melodie, ale też muzyczne instytucje, technologie i ludzkie praktyki współtworzą częściowo zamkniętą pętlę regulacyjną. Taki system funkcjonuje w określonej niszy ekologiczno-kulturowej (Laland, 2000) oraz ją współtworzy, a jego stabilność i zmienność wynikają z ciągłych sprzężeń zwrotnych pomiędzy jednostkowymi aktorami i strukturami zbiorowymi. W duchu teorii aktora-sieci (Latour, 2005) system muzyczny nie redukuje się do działań ludzi – współtworzą go także technologie, zapisy nutowe, instytucje czy algorytmy rekomendacyjne, które posiadają własną ewolucyjnie rozwijaną sprawczość. W ujęciu cybernetyki systemów (Beer, 1972) system taki jest zdolny do regulacji, adaptacji i przetrwania. Dodatkowo

zgodnie z koncepcją wieloskalowej architektury kompetencji (Levin, 2023), struktura zarządzania takiego systemu muzycznego jest rozproszona, a jego elementy warstwowo łączą się w coraz bardziej złożone układy funkcji i kompetencji.

Jednym z mechanizmów działania takich systemów muzycznych jest proces odciążania poznawczego i kulturowego (Clark, 1997; Menary, 2010): część zadań pamięciowych, twórczych i organizacyjnych człowieka zostaje przeniesiona na nośniki i instytucje, które podlegają memetycznej ewolucji. W tym sensie system muzyczny to nie tylko sieć dźwięków, ale cały 'krajobraz działań' (taskscape) (Ingold, 1993), co oznacza, że praktyki, narzędzia i idee są ze sobą nierozzerwalnie splecione. Opisywany system stopniowo zwiększa stopień autonomii wobec jednostkowej ludzkiej intencjonalności i rozwija się w kierunku autonomii.

Zgodnie z cybernetyczną teorią systemów autonomicznych Mariana Mazura (1966), system autonomiczny to taki, który „posiada zdolność sterowania oraz zdolność przeciwstawiania się utracie tej zdolności” (cyt. za Pawlak, 2020). Przyjęta tu koncepcja systemu muzycznego całkowicie spełnia tę definicję. Dalej w pracy prezentuję muzyczne systemy od oralności, poprzez rozwój notacji i instytucji, aż po współczesne rekomendery i sztuczne inteligencje. Wszystkie z nich zarówno sterują przepływem informacji muzycznych, jak i wytwarzają mechanizmy chroniące ich zdolność samoregulacji. Tak rozumiany system muzyczny to układ, który trwa i rozwija się dzięki zdolności do zachowania własnej autonomii – nawet wtedy, gdy ludzkie uczestnictwo ulega marginalizacji, a w centrum działania pozostają ekosystemy maszynowe i kultury sztucznej inteligencji.

3.2.1. Stadium proto-earworms

Na drodze do stopniowego spadku agentowości człowieka w funkcjonowaniu systemów muzycznych bardzo ważną rolę odgrywa powstanie pierwszych muzycznych replikatorów. Nie mamy jak dotąd pewności, jakie dokładnie jest ewolucyjne pochodzenie muzyki: czy muzyka poprzedzała język jako sposób komunikacji (Darwin, 1871), może język poprzedzał muzykę (Spencer, 1857), a może muzyka i język pierwotnie posiadały wspólnego przodka – muzojęzyk (Brown, 2000). Poniżej postaram się wykazać, że wszystkie koncepcje tego, skąd wzięła się muzyka, byłyby bliższe prawdy i kompletne, gdyby nie zapominały o proto-sprawczości muzycznych replikatorów. Warto wspomnieć, że również Steven Jan stworzył własną teorię replikatorowych początków muzyki (Jan, 2015c).

Niezależnie od prapoczątków, możemy definitywnie stwierdzić, że w pewnym momencie powstała muzyka jako odrębne, wyróżnialne zjawisko od innych przejawów kultury oraz systemów komunikacji. Trudno będzie na potrzeby niniejszej historycznej refleksji obiektywnie zdefiniować twardą, jednoznaczną granicę tego, gdzie kończy się język, a zaczyna muzyka. Definicji muzyki w muzykologii jest wiele i badacze nie osiągnęli tu konsensusu. Z memetycznego punktu widzenia nie ma w tym nic dziwnego. Wszystko, co podlega ewolucji, mieści się w spektrum powolnych, płynnych zmian, a narzucanie zero-jedynkowych granic to przede wszystkim ludzkie narzędzie pomagające w racjonalnej systematyzacji post hoc. Tak samo, jak w przyrodzie często nie istnieją jednoznaczne granice gatunków (Walker J., 2024), tak samo w muzyce wiele dzieł łamie gatunkowe podziały i tak samo muzyka i język pomimo niektórych niewątpliwie rozłącznych, charakterystycznych cech posiadają także części wspólne.

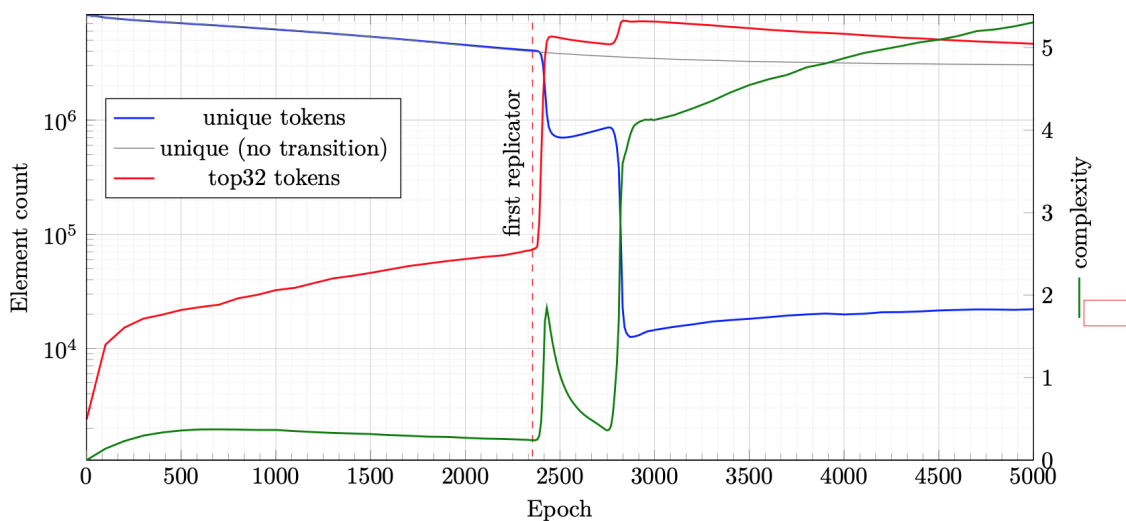
Jeżeli hipoteza muzojęzyka jest prawdziwa, muzyka i język pochodzą od wspólnego systemu komunikacji. Jeżeli nie, muzyka i język zaistniały jako osobne systemy, pełniące różne funkcje oraz stanowiące różne adaptacje – ale i tak w momencie pojawienia się brzmiącej formy wspólnie tworzyły audiosferę naszych praprzodków, a

ich reprezentacje wspólnie zamieszkiwały w pamięci przodków człowieka. Rozważając zatem pojawienie się pierwszych replikatorów wśród dźwiękowych komunikatów naszych przodków, nie będziemy poszukiwać ostrego rozróżnienia pomiędzy komunikatami muzycznymi i językowymi.

Należy dodać, że według diskutowanych dziś teorii zarówno muzyka, jak i język mają swoje niezależne początki (albo jeden wspólny początek) w komunikacji dźwiękowej lub gesturalnej. Jeżeli prawdziwa jest hipoteza gesturalna, replikatory gesturalne poprzedziły powstanie replikatorów dźwiękowych. Poniżej jednak zajmuję się warunkami powstania pierwszych replikatorów dźwiękowych, niezależnie od tego, czy poprzedziły je replikatory gesturalne, czy też nie. Zanim powstała kultura dźwiękowa, to znaczy zanim rozpoczęło się pozagenetyczne dziedziczenie ekspresji dźwiękowych z modyfikacjami, komunikaty dźwiękowe nie mogły być w żaden sposób imitowane, kopiowane z modyfikacją, nie mogło istnieć żadne ich społeczne uczenie się.

Mówimy zatem o hipotetycznym scenariuszu, w którym wydawane przez naszych przodków dźwięki wynikają wprost z ich fizjologii oraz zachowań instynktownych. To znaczy z ewolucyjnie ukształtowanej percepcji dźwięku oraz instynktownych zawołań. Jeżeli taki stan w ogóle miał miejsce, co wydaje się mało prawdopodobne, to według współczesnych szacunków musiałoby to mieć miejsce wcześniej niż 1.75 miliona lat temu (Jurgens, 2009; Beyene, 2013). Obecność złożonych technik tworzenia narzędzi z kamienia z tamtego czasu sugeruje już komunikację językową (Beyene, 2013), a badania mózgu sugerują, że przy takich czynnościach uaktywniają się również obszary odpowiadające za mowę, co utwierdza badaczy w przekonaniu, że 1.75 miliona lat temu nasi przodkowie posługiwali się już jakąś formą języka (Stout, 2012). To znaczy, że dźwiękowe dziedziczone informacje kulturowe (dźwiękowe memy) są z nami już przeszło półtora miliona lat. Kolejnym warunkiem opisującym powstanie dźwiękowego replikatora jest wysokie prawdopodobieństwo takiego zajścia, co sugeruje pokrewny eksperyment empiryczny.

Badanie z 2024 (Agüera, 2024) wykazuje empirycznie, że jeżeli zachodzą warunki umożliwiające powstanie replikatora, to, niezależnie od typu środowiska, swobodnie mutujące informacje prędzej czy później formułują replikator, który, wraz z potomstwem, zaczyna w takiej puli informacyjnej dominować:



W jednym z kluczowych eksperymentów ukazującym "pierwotną zupę" złożoną z losowych instrukcji logicznych zaobserwowano gwałtowną zmianę stanu systemu. W momencie pojawienia się pierwszego efektywnego replikatora różnorodność programów w "zupie" drastycznie spada, a nowo powstały replikator wraz ze swoimi kopiami zaczyna dominować w całym środowisku. Ilustruje to wykres, na którym widać gwałtowny spadek liczby unikalnych "tokenów" (odpowiedników naszych komunikatów dźwiękowych) i jednocześnie wzrost złożoności systemu, co sygnalizuje narodziny i rozprzestrzenianie się replikatora.

W artykule Blaise'a Agüery y Arcas i zespołu, autorzy badają fundamentalne kwestie dotyczące powstawania życia, obserwując procesy ewolucyjne w środowisku cyfrowym na komputerze. Autorzy mają na względzie to, że materię żywą wyróżnia zdolność do replikacji. Choć sama ta teza może być polem szerokiej dyskusji, nie ma w tym miejscu znaczenia dla naszego wniosku, ponieważ poszukujemy jedynie replikatorów, a nie przejawów życia. Autorzy także poszukiwali replikatorów, a interpretacje dotyczące początków życia mieszczą się poza sferą obliczeniową omawianego eksperymentu. Badania te, choć prowadzone na instrukcjach, operacjach logicznych w ramach różnych języków programowania, dostarczają cennych analogii do procesów ewolucji kulturowej,

w tym ewolucji memów dźwiękowych. Można nawet porównać takie działanie do programu badawczego metabiologii Gregorego Chaitina (Chaitin, 2012), w który formalizuje dynamikę procesów ewolucyjnych przy pomocy operacji logicznych i matematycznych. Różne metody komputerowego modelowania systemów ewolucyjnych mają swoje mocne strony, a także ograniczenia (Bartocci, 2016) W tym miejscu rozważam dynamikę zachowania się informacji - także informacji kulturowych.

Autorzy (Agüera, 2024) wykazują, że w środowiskach, w których losowe, początkowo niereplikujące się programy (operacje logiczne) wchodzi ze sobą w chaotyczne interakcje, spontanicznie dochodzi do powstania replikatorów. Tego typu efekt zaszedł we wszystkich oprócz jednej próby. Zaobserwowany efekt może jednak być wynikiem określonych cech narzędzia, jakim jest komputer/sieć, które nie muszą być obecne we wszystkich systemach ewoluujących. Analizowano różne środowiska i języki programowania: od wariacji ezoterycznego języka Brainfuck (BFF), przez języki stosowe (Forth), aż po emulatory istniejących materialnie procesorów, jak Z80. W każdym z tych przypadków, z wyjątkiem języka SUBLEQ, zaobserwowano spontaniczne powstawanie replikatorów o postępującej złożoności. Co ciekawe, w SUBLEQ najprostszy możliwy replikator jest znacznie bardziej złożony niż w reszcie testowanych środowisk. Istnieje zatem możliwość, że właśnie warunek złożoności wpływa na to, czy w danym środowisku możliwy replikator powstanie, czy nie. Ponieważ jednak wszystkie eksperymenty były ograniczone czasowo, nie możemy powiedzieć, że definitywnie wykazano, że replikator w środowisku SUBLEQ nie może powstać. Wiemy zatem, że w określonych empirycznych próbach ów replikator nie powstał, ale zawsze istnieje możliwość, że przy dłuższym prowadzeniu eksperymentu jednak nagle spontanicznie doszłoby do jego wyłonienia się. Co istotne, proces replikacji zachodzi również w wyniku losowych interakcji i samomodyfikacji programów – bez konieczności zachodzenia losowych mutacji w tle. Mutacje przyspieszają powstanie replikatora, ale nie są do tego warunkiem koniecznym.

Nie udowodniono, że replikator powstaje zawsze, kiedy tylko może. Możemy za to stwierdzić, że przekonująco wykazano, iż powstanie replikatora w dogodnych

warunkach jest możliwe, a nawet wysoce prawdopodobne. Za to istnienie takich przypadków, w których pomimo sprzyjających warunków, replikator ponad wszelką wątpliwość nigdy nie powstanie, nie zostało formalnie udowodnione. Ponadto, spontaniczne powstawanie replikatorów zostało wykazane także w chemii, gdzie w sprzyjających warunkach zaobserwowano samorzutne formowanie się cząsteczek zdolnych do samopowielania (Liu et al., 2020). Zatem, wyniki te dostarczają silnego formalnego oraz empirycznego wsparcia dla tezy, że jeśli istnieją warunki umożliwiające replikację (w kulturze dźwiękowej będą to zdolności percepcyjne, pamięć, zdolność do imitacji dźwięków (uczenie się wokalne)), to replikatory, czy to w formie operacji logicznych, czy memów dźwiękowych, prędzej czy później powstaną i zaczną dominować w puli informacyjnej.

To z kolei otwiera drogę do dalszej, bardziej złożonej ewolucji, w której replikatory konkurują ze sobą, tworzą ekosystemy, a nawet symbiotyczne kolektywy, wykazując coraz bardziej złożone zachowania. Rozważmy zatem, jak mógłby wyglądać taki pierwszy dźwiękowy replikator i jakie warunki muszą być spełnione, aby mógł powstać. Przytoczona wcześniej symulacja źle zinterpretowana mogłaby sugerować, że pojawienie się replikatorów dźwiękowych mogło odbyć się przypadkiem. W rzeczywistości jednak bezpośrednio musiało być konsekwencją pojawienia się specyficznych, wysoko rozwiniętych zdolności poznawczych i ewolucji określonych cech fizjologicznych człowieka. Te cechy i zdolności, nawet jeżeli wyewoluowały pierwotnie z innych powodów, wspólnie stworzyły podatną niszę dla replikacji wzorców dźwiękowych (Killin, A., 2016).

Pierwszym warunkiem powstania dźwiękowych replikatorów jest istnienie odpowiednio precyzyjnej percepcji słuchowej – zdolności różnicowania i przetwarzania informacji akustycznych. Nasi przodkowie musieli być w stanie odróżniać właściwości dźwięku, takie jak wysokość, barwa czy rytm. Ponadto, do powstania replikatora postrzegany wzorec dźwiękowy musi zostać zapamiętany. Potrzebowaliśmy w tym celu dwóch rodzajów pamięci słuchowej: pamięci długotrwałej oraz roboczej. Tylko w ten sposób wzorce mogły zostać skutecznie utrwalone. Niezbędne do późniejszej kulturowej

ewolucji muzyki struktury pamięci, z logicznej konieczności, musiały powstać z powodów, które dziś określilibyśmy jako pozamuzyczne. Badania wskazują, że takimi powodami mogła być rosnąca złożoność komunikacji społecznej oraz analiza środowiska (Murray, 2020). Współcześnie prowadzi się badania na noworodkach, które wykazują ich wrodzone predyspozycje do przyswajania języka i muzyki (Kujala, 2023; Brandt, 2012). Sugeruje to istnienie głęboko ewolucyjnie zakorzenionych, specyficznych struktur mózgowych oraz zdolności pamięciowych niezbędnych, aby zachodziła transmisja informacji kulturowych (Ravignani i in., 2016). Co więcej, badania empiryczne wskazują, że kulturowa transmisja i zarazem ewolucja kulturowa komunikacji wokalne nie tylko u ludzi, ale także u waleni i ptaków opiera się właśnie na uczeniu się i pamięci. (Ravignani i in., 2016, Garland & McGregor, 2020). Wskazuje się także na istnienie pętli fonologicznej jako komponentu pamięci krótkotrwałej w korze słuchowej, który przechowuje informacje słuchowe i jest kluczowy dla procesu zapamiętywania dźwięków (Atherton, R. i in, 2018).

Wierność pamięci słuchowej jest kluczowym warunkiem wstępnym pojawienia się replikatora. Nie odnosi się ona jednak do wiernego kopiowania całego obrazu akustycznego, lecz do przechowywania i odtwarzania tych cech wzorca dźwiękowego, które są poznawczo dystynktywne. Dla stabilności replikatora wystarczy zatem, aby system poznawczy utrwał elementy uznane za istotne: czy to kontur intonacyjny w mowie (Kapatsinski, 2017), czy precyzyjne relacje wysokościowe w muzyce (Tillmann, 2023). „Dokładność” nie oznacza więc absolutnej tożsamości, ale taką zgodność, która pozwala słuchaczowi rozpoznać i powtórzyć dany wzorzec jako ten sam. O ile w mowie tolerowane są znaczne różnice między kolejnymi realizacjami tego samego tonu leksykalnego, o tyle w muzyce margines odchylenia bywa mniejszy, np. nie jest obojętne, czy w melodii pojawia się tercja mała, czy wielka. Pomimo tych różnic, w obu przypadkach mechanizm opiera się na zdolności pamięci do utrzymania względnej stabilności wybranych cech, co umożliwia powstanie replikatorów.

Żeby umożliwić istnienie dźwiękowego replikatora, nie wystarczy słyszenie i zapamiętywanie dźwięków: niezbędna jest również zdolność powtórzenia ich. W

literaturze istnieje pojęcie uczenia się wokalnego (vocal learning), które odnosi się właśnie do umiejętności wokalne imitacji zasłyszanego wzorca (Janik, 1997). Złożone uczenie się wokalne, definiowane jako zdolność do modyfikowania wytwarzanych dźwięków na podstawie bodźców słuchowych, odróżnia się od wokalizacji wrodzonych i jest rzadkością w królestwie zwierząt, choć u ludzi jest wysoko rozwinięte (Ravignani, 2021; Akçay, 2023). Zaawansowane uczenie się wokalne obejmuje przewidywania czasowe i silną integrację słuchowo-motoryczną, kluczowe dla precyzyjnej kontroli i zarazem imitacji dźwięków (Ravignani, 2021). Ponadto, spośród szeroko rozumianego uczenia się wokalnego wyróżniane bywa również uczenie się produkcji wokalne (vocal production learning, VPL), o szczególnej istotności dla dźwiękowych replikatorów. VPL polega na powielaniu nieznanych wcześniej bodźców dźwiękowych do produkcji nowych sygnałów lub modyfikowania już istniejących, co pozwala na większą złożoność i elastyczność niż opieranie się na wokalizacjach wrodzonych. Ewolucji VPL towarzyszyły innowacje w neuronalnych obwodach motoryki wokalne (Akçay, 2023). Złożone, zaawansowane uczenie się wokalne pozwala nie tylko na modyfikację wrodzonych zawołań, ale także na generowanie i imitację arbitralnych wzorców dźwiękowych, co jest niezbędne dla otwartego systemu kulturowej replikacji dźwięku. Tylko w ten sposób może powstać wystarczająca dźwiękowa różnorodność, która umożliwi zachodzenie procesów selekcyjnych.

Posiadanie umiejętności wokalne imitacji samo w sobie nie jest wystarczającym wskazaniem, że imitacja istotnie często występowała – dostatecznie często, żeby sprowokować powstanie środowiska selekcji replikatora. Istnieje za to wiele różnorodnych badań (Tzafestas, 2021; Paukner, 2009; Schleihauf, 2020) wskazujących, że imitacja istotnie stanowiła fundament społecznych zachowań naszych przodków – że to właśnie zdolność do imitacji sprowokowała i ukształtowała powstanie późniejszej złożonej ludzkiej kultury. Badania kognitywistyczne wykazują, że imitacja jest zjawiskiem ewolucyjnie pierwotnym, które prawdopodobnie poprzedziło zdolność do przypisywania znaczeń i celowej komunikacji (Tzafestas, 2021). Społeczną rolę naśladownictwa oraz głębokie korzenie ewolucyjne tej zdolności dostrzegamy dzięki

obserwacjom innych naczelnych – kapucynek – które wykazują wyraźną preferencję wobec naśladowujących je ludzi (Paukner, 2009).

U współczesnych ludzi ów pęd do naśladowania jest tak silny, że dzieci w wieku przedszkolnym stosują tzw. nadmierną imitację, kopiując nawet te czynności, które są w sposób oczywisty zbędne lub nieefektywne do osiągnięcia celu. Wskazuje to, że imitacja jest motywowana nie tylko instrumentalnie, ale przede wszystkim społecznie – chęcią afiliacji i przyswojenia norm (Schleihauf, 2020). Według Tomasello (Tomasello, 1999) zdolność imitacji stanowiła fundament wczesnej kultury homininów, którą nazywa kulturą mimetyczną – od *mimesis*, naśladownictwa. Badacz stwierdza, że naśladownictwo było kluczowe dla przekazywania umiejętności takich jak wytwarzanie narzędzi, na długo przed pojawieniem się złożonej kreatywności (Tomasello, 1999). Dowody archeologiczne, takie jak niewielka innowacyjność narzędzi aszelskich na przestrzeni ponad miliona lat, sugerują, że homininy były biegłe w precyzyjnym kopiowaniu, ale ich zdolności do innowacji były ograniczone w porównaniu do późniejszej eksplozji kreatywności u *Homo sapiens* (Stout, 2011; Henshilwood, 2011). Ponadto, z innych badań wynika, że imitacja pełni rolę „społecznego spoiwa” (Paukner, 2009), a jej brak w interakcjach społecznych może prowadzić do wzrostu poziomu stresu, co potwierdzają testy fizjologiczne (Kouzakova, 2010).

Mamy zatem silne wskazania ku temu, że nasi przodkowie posiadali sprawny słuch i odpowiednią pamięć dźwiękową, potrafili wokalnie imitować zasłyszane nowe dźwięki oraz ich zachowanie charakteryzowała wszechobecna wzajemna imitacja. Te trzy zdolności nie są niezależne, lecz tworzą zintegrowany system. System ten zbudował niezbędne i wystarczające warunki, aby wzorce dźwiękowe zaczęły zachowywać się jak replikatory – środowisko, w którym informacja może być przechowywana, zmieniana i kopiowana. Środowiskiem replikacji genów jest materialny świat podlegający prawom fizyki i chemii, a nasi przodkowie na tym systemie nadbudowali kolejny, mentalny, jako zaawansowaną niszę do replikacji dźwiękowych memów. Kim Sterelny zauważa:

„Konstruowanie niszy przez zwierzęta zawsze modyfikowało środowisko fizyczne i społeczne, ale teraz modyfikacje te uległy spotęgowaniu, gdy coraz bardziej celowe i świadome formy naśladowania i nauczania wśród homininów doprowadziły do powstania «środowiska poznawczego», niszy zdefiniowanej przez zarchiwizowaną informację”.

Animal niche construction had always modified the physical and social environment, but now the modifications were compounded, as modes of more and more purposive and conscious imitation and instruction among hominins brought about an “epistemic environment,” a niche defined by archived information. (Sterelny, 2003)

Rozwijając jej koncepcję, można powiedzieć, że nasz zintegrowany system pamięci, wokalnego uczenia się oraz imitacji doprowadził do powstania dźwiękowego środowiska poznawczego. Richard Dawkins pisał, że ewolucja memów rozpoczęła się w „Pierwotnej zupie” kultury. To porównanie do początków życia w chaotycznym, chemicznym środowisku. „Pierwotna zupa” replikatorów dźwiękowych to ludzkie muzyczne środowisko poznawcze i społeczne. Powstanie tej nowej niszy pociągnęło za sobą powstanie nowych presji selekcyjnych dla genów, ale przede wszystkim także dla samej informacji kulturowej. Wokalizacje dźwiękowe przestały być jedynie wytworami popędów biologicznych, a mogły zyskać własne „życie” – elementarną sprawczość – poprzez replikację. Zanim ten system w pełni zaistniał, komunikacja dźwiękowa była w dużej mierze dziedziczona genetycznie, instynktowna. Wraz z zaistnieniem systemu replikacji, wzorce dźwiękowe mogły być generowane, zapamiętywane i kopiowane z wystarczającą wiernością i zmiennością, aby podlegać ewolucji kulturowej. Jest to jakościowa zmiana, otwierająca nową ewolucyjną ścieżkę dla muzycznych informacji.

Powyższe rozważania na temat warunków powstania dźwiękowego replikatora – opartego na zintegrowanym systemie percepcji, pamięci i wokalnej imitacji – prowadzą nas wprost do współcześnie obserwowanego zjawiska earworm, czyli „robaka usznego”. Jest to mimowolnie powracający w myślach fragment melodii, zapętlałającej się w naszej świadomości bez naszej intencjonalnej kontroli. Jak próbowałem wykazać,

choć zjawisko to kojarzone jest głównie ze współczesną muzyką popularną, jego mechanizm jest głęboko zakorzeniony w ewolucyjnej architekturze naszego umysłu i stanowi doskonały model dla pierwszych, elementarnych form muzycznej ewolucji kulturowej.

Koncepcja „robaka usznego” idealnie spaja się z memetyczną teorią kultury. Filozof Daniel Dennett (1995) określa memy mianem „wirusów umysłu” (viruses of the mind). Według niego, memy, podobnie do wirusów biologicznych, kolonizują ludzkie umysły i wykorzystują ich zasoby poznawcze do własnej replikacji, często bez względu na korzyść dla samego gospodarza. Ewidentnie te same lub blisko pokrewne zjawiska ma właśnie określić metafora robaka, który „wchodzi do ucha”. Earworm to nic innego jak wirus umysłu nabyty drogą słuchową. Dźwiękowy wirus umysłu to trafniejsza metafora niż robak uszny – dźwiękowe memy replikują się w naszej pamięci, nie w naszych uszach, a informacyjna, martwa biologicznie struktura memu na pewno bardziej odpowiada wirusom niż robakom. Earworm to zatem dźwiękowy oraz muzyczny mem, który z sukcesem zainfekował nasze muzyczne środowisko poznawcze – czyli jego ekologiczną niszę. Jego funkcją nie jest przekazanie znaczenia ani inne działanie dla dobra człowieka-gospodarza. Ewolucyjnie wyłonioną funkcją każdego replikatora jest przede wszystkim powielenie własnej struktury. Mimowolny, natrętny charakter earworms jest zaś świadectwem ich agentowości, sprawczości – zdolności do aktywacji w umyśle gospodarza nawet wbrew jego woli.

Współczesne badania nad mimowolnymi wyobrażeniami muzycznymi (ang. Involuntary Musical Imagery, INMI), jak naukowo określa się earworms, wskazują na ich specyficzne cechy, które czynią je wysoce skutecznymi replikatorami. Badania prowadzone m.in. przez Victorię Williamson (Williamson, 2014) czy Kelly Jakubowski (Jakubowski i in., 2017) pokazują, że fragmenty muzyczne, które najczęściej stają się earworms, charakteryzują się prostą, przewidywalną strukturą meliczną, szybszym tempem oraz powtarzalnymi motywami rytmicznymi. Te cechy minimalizują obciążenie dla pamięci roboczej i długotrwałej, ułatwiając precyzyjne zakodowanie i przywołanie

wzorca, co jest kluczowym warunkiem wierności replikacji. Te konkretne cechy możemy uznać za memetyczne adaptacje, i nie możemy wnioskować, że pierwsze dźwiękowe replikatory posiadały taką samą strukturę. Tym niemniej – podobieństwa nie są wykluczone. Można zatem z dużą dozą prawdopodobieństwa założyć, że pierwsze dźwiękowe replikatory, które narodziły się w „pierwotnej zupie” kultury homininów, miały charakter bardzo zbliżony do dzisiejszych earworms. W środowisku poznawczym, w którym dominowała mimetyczna, a nie innowacyjna forma przekazu kulturowego (Tomasello, 1999), to właśnie najprostsze, najbardziej zaraźliwe i łatwe do wokalnego odtworzenia wzorce dźwiękowe miały największą szansę na przetrwanie i propagację. Ich sukces nie opierał się na wartości estetycznej czy komunikacyjnej, lecz na skuteczności sprowokowania replikacji, napędzanej między innymi przez społeczny pęd do imitacji. Te proto-earwormy pasożytowały na nowo powstałej dźwiękowej niszy poznawczej, torując drogę dla całej późniejszej, złożonej ewolucji kulturowej muzyki.

W tym ujęciu dzisiejszy earworm staje się żywą skamieliną – oknem na sam początek ewolucji kulturowej w domenie dźwięku. Ukazuje on w najbardziej podstawowej formie, jak informacja dźwiękowa mogła zyskać własne „życie” i elementarną sprawczość, uniezależniając się częściowo od genetycznie ukształtowanych popędów i rozpoczynając nową, częściowo autonomiczną ścieżkę ewolucyjną memów dźwiękowych. Od prostego, samolubnego, replikującego się wzorca dźwiękowego do Ballad Chopina czy Linkin Park droga była długa, ale jej pierwszy krok został postawiony właśnie wtedy, gdy pierwszy „robak uszny” zagnieździł się w umyśle naszego przodka.

3.2.2. Przejście od kultury oralnej do kultury piśmiennej w muzyce Zachodu

Niniejszy rozdział omawia najważniejsze z perspektywy memetycznej aspekty przejścia w muzyce od kultury oralnej do kultury pisanej, koncentrując się na kulturze muzycznej Zachodu. Ten bardzo skomplikowany, ciągły i wielotorowy proces wiąże się ze zmianami w strukturze muzyki, a także w sposobach jej dystrybucji oraz konceptualizacji. Ponadto, zmieniły się także same ewolucyjne presje selekcyjne, a nawet mechanizmy procesu ewolucji muzyki.

Dla tej rozprawy ten aspekt ewolucji muzyki ma znaczenie kluczowe, ponieważ przeniesienie istnienia muzycznego repertuaru z ludzkiej pamięci na zewnętrzne nośniki, takie jak partytury, stanowi znaczący krok w postępującym wciąż procesie uniezależniania się memów muzycznych od środowiska poznawczego człowieka. Dalsze przejście od partytury do taśmy perforowanej, wałków woskowych, płyt winowych, kaset, płyt CD oraz streamowanych plików audio było już naturalną konsekwencją tego wielkiego pierwszego kroku.

Wynalazek zapisu jest tutaj rozumiany nie jako neutralny archiwizacyjny dodatek do muzyki, ale jako potężna, sprawcza technologia, która doprowadziła do fundamentalnej transformacji muzyki oraz istotnie, nieodwracalnie zmieniła kierunek jej ewolucji. Warto zaznaczyć, że analogiczny proces miał miejsce w przypadku języka: wynalezienie pisma nie tylko umożliwiło przechowywanie i przekazywanie treści niezależnie od pamięci jednostek. Przede wszystkim zasadniczo zmieniło dynamikę rozwoju cywilizacji, stało się jednym z kluczowych motorów złożoności społecznej i akumulacji wiedzy (Diamond, 1997; Wright, 2000). Ta zmiana dynamiki rozwoju cywilizacji miała oczywiście także ogromny wpływ na memetyczną ewolucję muzyki.

Żeby wykazać wszystkie te zmiany w muzyce, należy na początek przytoczyć główne mechanizmy funkcjonowania muzyki w kulturze oralnej, żeby potem na zasadzie

kontrastu zademonstrować ich odpowiedniki w kulturze pisanej. Przytoczone przykłady i ujęcia teoretyczne będą przedstawiać wybrane aspekty muzycznej kultury oralnej wystarczające do stworzenia elementarnego modelu przejścia od oralności do piśmienności w kulturze muzycznej z perspektywy memetycznej. W poprzednim rozdziale omówiłem warunki powstania i charakterystykę pierwszych dźwiękowych replikatorów i zarazem pierwszych memów muzycznych. Jak wspominałem, istnieją dwie fundamentalne zależności związane z replikacją memów.

Po pierwsze, aby zachodził proces ewolucyjny, muszą być spełnione trzy warunki: zachodzenie spontanicznych mutacji, dziedziczenie tych mutacji przy replikacji, oraz presja selekcyjna. W muzycznej kulturze oralnej te trzy warunki są spełnione niejako automatycznie z powodu niedoskonałości czy też cech ludzkiego systemu komunikacji i pamięci. Komunikaty dźwiękowe, w tym muzyczne, naturalnie ewoluują, ponieważ od intencji wyrażenia komunikatu, do jego treści, do komunikatu odebranego i dalej zapisu tego komunikatu w wadliwej pamięci nieuchronnie zachodzi zmienność (Shannon, 1948).

Po drugie, skuteczność replikacji memu, w tym memu muzycznego, możemy funkcjonalnie analizować pod kątem trzech cech: płodności, długotrwałości, kopiowalności (Dawkins, 1976). W muzycznej kulturze oralnej jedynym realnym nośnikiem treści muzycznej są bezpośrednie komunikaty dźwiękowe, wykonania muzyczne. Ich długotrwałość i kopiowalność są bardzo mocno ograniczone z racji nietrwałości materialnego medium i psychofizycznych ograniczeń człowieka, które opisywałem w poprzednim rozdziale.

Tym samym w kulturze oralnej główną siłą muzycznych replikatorów jest ich płodność, czy też zaraźliwość. W lokalnych repertuarach krążą te melodie, które zapadają w pamięć. Zarazem szybko ulegają one mutacjom i ewolucji, ponieważ nie istnieje żaden pojedynczy idealny wzorzec. Każde wykonanie jest jednocześnie aktem twórczej re-kreacji. Zmienność nie jest anomalią, lecz normą. Ten ciągły proces mutacji prowadzi do powstawania tzw. „rodzin melodii” (tune families) (Nettl, 2005), gdzie jeden utwór

istnieje jako dynamiczna „chmura” powiązanych ze sobą wariantów, a nie jako pojedynczy, stały byt. To płynne istnienie muzyki jest analogiczne do opisu języka w kulturach oralnych przez Waltera Onga, gdzie słowo jest ulotnym „wydarzeniem” dźwiękowym, a nie statycznym, wizualnym „przedmiotem” (Ong, 1982).

Ewolucja dzieła jawi się tu jako proces rozproszony, określony przez Nettla jako „re-kreacja wspólnotowa” (communal re-creation) (Nettl, 2005), w której cała populacja nosicieli danego memu przyczynia się do jego transformacji.

Ta koncepcja idealnie ilustruje to, co filozof Daniel Dennett określa mianem ewolucji oddolnej (bottom-up), w której powstaje projekt bez projektanta (design without a designer) (Dennett, 1995). Utwór, który istnieje pod postacią rodziny melodii, czy też chmury wariantów, ze swoją wewnętrzną logiką i złożonością, nie jest wynikiem odgórnego planu jednego twórcy. Jest raczej emergentnym rezultatem niezliczonych, drobnych interakcji i modyfikacji dokonywanych przez anonimową wspólnotę. Kolejne warianty utworu wyłaniają się z dynamicznego pola oddziaływań społecznych, a nieświadomymi i niewolicjonalnymi współautorami są wszyscy biorący udział w re-kreacji członkowie społeczności. W tym ujęciu poszczególne warianty melodyczne, rytmiczne czy frazy – memy muzyczne – konkurują ze sobą o przetrwanie w złożonej poznawczej niszy ekologicznej: zbiorowości ludzkich umysłów i aparatów wykonawczych. Replikują i rozprzestrzeniają się te wersje, które są najlepiej przystosowane do ograniczeń i predyspozycji tej niszy: są łatwiejsze do zapamiętania, wywołują silniejszą reakcję emocjonalną czy też są lepiej dopasowane do możliwości ludzkiego głosu (Aubé, 2013; Jacoby, 2024; Jacoby, 2019). Proces ten działa niejako samoczynnie, napędzany uniwersalnym algorytmem ewolucyjnym (wariacja, selekcja, replikacja), bez potrzeby istnienia świadomego, intencjonalnego kierunku nadawanego przez człowieka.

Napędzającą ten rodzaj memetycznej ewolucji muzyki replikację w terminologii memetycznej Susan Blackmore określa jest jako kopiowanie produktu (copy-the-product), ponieważ, jak w przypadku uczenia się wokalnego, imitowane jest bezpośrednio postrzegane zachowanie (Blackmore, 1999). Wykonawców w takim

systemie charakteryzuje zjawisko, które Dennett nazwał „kompetencją bez zrozumienia” (competence without comprehension) (Dennett, 2017). Osoby muzykujące potrafią biegle śpiewać, grać, improwizować w ramach swojej lokalnej mikro-tradycji. Nie muszą w tym celu posiadać analitycznego zrozumienia zasad akustyki, teorii muzyki czy psychologii poznawczej, które sprawiają, że dana melodia po prostu działa. Kluczową rolę odgrywają tu odziedziczone memetycznie wzorce, które same są produktem długiej, ślepej ewolucji kulturowej. Ważna jest także intuicja, która powstała jako efekt istnienia pętli sprzężenia zwrotnego pomiędzy dziedzicznymi genetycznie mechanizmami poznawczymi i sygnałami ze środowiska. W każdym razie kompetencje wynikają z ucieleśnionej praktyki oraz muzykalności, a nie racjonalnego, dyskursywnego rozumienia.

Powstaje tu tym samym bardzo interesujący paradoks – istniejący system muzyczny działa w precyzyjnie określony sposób, posiada bardzo konkretne zasady – ale te zasady nie zostały ani świadomie przez człowieka zaprojektowane, ani nie są wprost znane zaangażowanym członkom społeczności. Chmury wariantów i lokalne mikro-tradycje ściśle opierają się o pewne zasady muzyki, pomimo tego, że nie zostały one nigdzie sformułowane i zapisane. W takiej sytuacji powstaje pytanie: jak to możliwe, żeby system złożony z ludzi i memów sprawnie funkcjonował w określony sposób, zgodnie przestrzegając niepisanych zasad? Aby na to odpowiedzieć, Dennett wprowadza koncepcję *free-floating rationales*, co na język polski bywa tłumaczone jako „swobodnie unoszące się racje” (Święczkowska, 2019), ale w tej pracy używam określenia dryfujące przyczyny. Chodzi o algorytmy postępowania ugruntowane w presjach procesu ewolucyjnego, które nie muszą być reprezentowane w umyśle żadnego konkretnego sprawcy, aby istnieć i być skutecznymi. Unoszą się one swobodnie, istniejąc w przyczynowej historii rozwoju danego systemu (Dennett, 1995). Według przykładu Dennetta, w biologii przyczyną szybkości geparda jest presja selekcyjna – szybsze osobniki skuteczniej polowały i przetrwały, przekazując geny. Żaden gepard nie musi tej przyczyny rozumieć. Mechaniczne relacje wyewoluowanego bytu oraz jego niszy i środowiska po prostu nieodwracalnie prowadzą do pewnych zachowań, a nie innych. Dalej postaram się wykazać, że tak samo dryfujące racje

kształtują świat muzyki, pomimo tego, że mówimy o znacznie bardziej zaawansowanym systemie. Przykładowo, teatry operowe nie muszą posiadać regulaminów, które zakazałyby wykonawcom zjedzenie kebaba na scenie w trakcie wykonania, a jednak takie zdarzenie prawdopodobnie nie zostało jeszcze nigdy odnotowane. Wynika to z wysoce zinternalizowanego wspólnego kodeksu zachowań w muzycznej niszy poznawczej. Koncepcja Dennetta odnajduje paralele w biologii kognitywnej. Zwierzęta, rośliny, a także bakterie w wyniku ewolucji skonstruowane są według pewnych określonych reguł, a zatem same w sobie mogą być traktowane jako nośnik skutecznej informacji – niejako „wiedzy” o świecie (Lyon, 2025). Lyon, autor artykułu *Fundamental Principles of Cognitive Biology 2.0* (Lyon, 2025) podaje przykład tego, że bakteria *E. coli* posiada niejawną wiedzę o tym, jak metabolizować laktozę w przypadku braku dostępności glukozy. Wiedza ta nie znajduje się w mózgu, którego oczywiście bakteria nie posiada. Jest ucieleśniona w molekularnych sieciach regulacyjnych, takich jak *lac operon*, które sterują aktywacją enzymów i transporterów w odpowiedzi na dostępne zasoby. Stanowi to rozwiązaniem ewolucyjnego problemu przetrwania (Lyon, 2025).

W świetle uniwersalnego darwinizmu tę samą mechanikę ukrytego kodowania informacji możemy dostrzec, kiedy zachodzi ewolucja memetyczna. Dotyczy to również wiedzy muzycznej, która częściowo wyłania się z presji selekcyjnych działających w ludzkiej niszy poznawczej. Dany wzorzec muzyczny może odnosić sukces, ponieważ wyjątkowo skutecznie wywołuje pożądane emocje. Wykonawcy używają go, bo „dobrze brzmi”, nie potrzebując do tego celu racjonalnych przyczyn wynikających z wiedzy psychoakustycznej (Huron, 2016) czy neuroestetycznej (Patel, 2007). Przyczyna użycia danego wzorca dryfuje swobodnie, oparta na jego skuteczności wykutej w toku zmienności ewolucyjnej. Zakotwiczona jest jedynie w historii jego udanych replikacji.

Koncepcja ta dostarcza klucza do zrozumienia niepisanej, niejawnej gramatyki muzyki. Sugeruje, że pierwotne zasady muzyki to nie te, które są spisane w podręcznikach. To raczej nieartykułowane pryncypia, które przetrwały rygorystyczny filtr selekcji kulturowej.

Istnienie ogólnych zasad ukształtowania materiału muzycznego w toku zmienności ewolucyjnej empirycznie wykazane zostały w eksperymentalnych łańcuchach transmisyjnych (transmission chain experiments) (Milton, 2020). Taki łańcuch transmisyjny to przykład memetycznej ewolucji w warunkach laboratoryjnych. W przełomowym badaniu, uczestnicy proszeni o odtwarzanie losowych sekwencji uderzeń perkusyjnych, po kilku „pokoleniach” przekazu kulturowego zaczęli spontanicznie wytwarzać uporządkowane rytmy, które wykazywały wszystkie sześć statystycznych uniwersaliów odnalezionych w muzyce świata (Ravignani et al., 2016). W badaniach tych początkowo przypadkowe i złożone melodie, przekazywane kolejno od osoby do osoby, w sposób systematyczny upraszczają się i nabierają struktury. Dzieje się tak pod wpływem nieświadomych tendencji transmisyjnych (transmission biases), wynikających z naszych dwutorowych ograniczeń. Po pierwsze, ograniczenia poznawcze – nasza pamięć z trudem przyswaja duże ilości nowych informacji, w związku z czym faworyzuje prostsze i bardziej znajome struktury melodyczne (Vuvan, 2014). Po drugie, ograniczenia fizjologiczne – nasz aparat wokalny z większą łatwością produkuje małe interwały melodyczne niż duże skoki (Ammirante, 2023), co prowadzi do „wygładzania” poszarpanych melodii w toku transmisji. To dwie grupy konsekwentnie działających algorytmów przetwarzania informacji w ludzkiej muzycznej niszy poznawczej.

W konsekwencji ten proces memetycznego doboru naturalnego, napędzany przez ograniczenia człowieka, nieuchronnie prowadzi do spontanicznego wyłaniania się cech uniwersalnych dla muzyki, takich jak niewielka liczba kategorii rytmicznych (zazwyczaj trzy), które są oparte na prostych, całkowitych proporcjach (np. 1:2, 2:3) (Ravignani et al., 2016). Powszechne cechy muzyki z kultur świata – takie jak użycie skal o małej liczbie dźwięków (5-7, np. pentatonicznych (Miller, 1956; McDermott, 2005)), przewaga małych interwałów melodycznych (Savage, 2015) czy łukowate kształty melodii (Savage, 2015; Adams, 1976) – nie są więc arbitralnymi konwencjami. Powstają w wyniku konsekwentnych mechanizmów przetwarzania informacji w naszej niszy poznawczej. Są to ewolucyjne rozwiązania, dryfujące przyczyny, które zostały odkryte przez ślepy proces transmisji kulturowej jako najbardziej skuteczne sposoby na przetrwanie memu w ludzkiej niszy poznawczej.

Powstawanie tych struktur to nie przypadek. Jak argumentuje Fred Lerdahl, systemy muzyczne oparte na elementach tworzących przejrzyste relacje stabilności (np. proste interwały i skale) umożliwiają słuchaczowi budowanie złożonych, hierarchicznych reprezentacji umysłowych (Lerdahl, 1992). Powtarzające się zniekształcenia w przekazie nie są całkiem przypadkowe. Dzięki temu, że podlegają filtracji przez ludzki system poznawczy, sprzyjają utrwalaniu specyficznych rozwiązań. Takich mianowicie, które stają się podstawą dla bardziej złożonych struktur umysłowych. Tę ludzką skłonność do tworzenia i czerpania przyjemności ze złożonych, hierarchicznych struktur w muzyce, W. Tecumseh Fitch (2005) metaforycznie nazwał "dendrofilią", a pojęcia to odnosi się także do muzyki. Proces ewolucyjnego powstawania uniwersalnych w kulturach muzycznych świata złożonych struktur omawia także Piotr Podlipniak. W jego koncepcji uniwersalia muzyczne są skutkiem istnienia biologicznych adaptacji, które stanowią muzycznie specyficzne zdolności poznawcze. W tej koncepcji fundamentalne cechy muzyki nie są jedynie produktem ubocznym ograniczeń poznawczych, ale wynikiem działania mechanizmów ewolucyjnych, które ukształtowały w naszym gatunku szereg muzycznie specyficznych zdolności, jak np. instynkt tonalny (Podlipniak, 2015). Ten biologiczny fundament wyjaśnia, dlaczego pewne struktury muzyczne są tak rozpowszechnione i intuicyjnie zrozumiałe w różnych kulturach (Podlipniak, 2007).

Podsumowując, muzyka w tradycji oralnej ewoluuje według ograniczeń ludzkiej niszy poznawczej, oddolnie, tworzy chmury wariantów, podlega bezpośredniej imitacji - kopiowany jest produkt - zachowanie, oraz wytwarza samoczynnie implicytną, naturalną gramatykę. Memy muzyczne w replikacji polegają na płodności, lokalnej zaraźliwości, ponieważ ich długowieczność i kopiowalność są bardzo niskie. Wszystkie te atrybuty ulegają zmianie na skutek przejścia od kultury oralnej do kultury pisanej. Technologia zapisu muzycznego nie powstała nagle i znikąd. Powstawała stopniowo, zgodnie z logiką ewolucji, której to nieustannie podlega aż do dziś. Wczesny Zachodni system zapisu używał neum adiastrematycznych (ok. IX w.), które były jedynie mnemonicznymi wskazówkami dla wykonawcy, który już znał melodię. Ich funkcja polegała na wizualnym

reprezentowaniu konturu melodii, bez precyzyjnego określania interwałów czy rytmu, co oznacza, że wspomagały one pamięć oralną, ale jej nie zastępowały (Christensen, 2002). Kolejne przełomy polegały na rozwoju notacji diastematycznej, precyzującej wysokość dźwięku na liniach, a następnie notacji menzuralnej (ok. 1250 r.), która pozwoliła na bardziej precyzyjne określenie czasu trwania dźwięków (Christensen, 2002). Te innowacje były warunkiem koniecznym dla rozwoju bardziej zaawansowanej polifonii. Ponadto, doprowadziły one do powstania systemu zdolnego do zakodowania muzyki z taką precyzją, że możliwe stało się pełne oddzielenie "dzieła" od pamięci i ulotnego kontekstu wykonawczego, przekształcając je w trwałe, analizowalny "tekst" (Ong, 2002). Nastąpiła reifikacja muzyki. Muzyczna pula memetyczna wzbogaciła się o nowe namacalne, materialne przedmioty.

Pojawienie się zapisu nutowego było rewolucją informacyjną, która fundamentalnie zmieniła ontologiczny status muzyki (Gleick, 2011). Notacja jest technologią, która przekształca ulotne, czasowe wydarzenie dźwiękowe w trwałe, przestrzenny obiekt (Ong, 2002). Dźwięk, który w kulturze oralnej istniał tylko w akcie artykulacji i natychmiast zanikał, zostaje unieruchomiony na papierze nutowym, stając się artefaktem, który można poddać analizie, archiwizować i reprodukować w niemal niezmięnionej formie (Ong, 2002; Gleick, 2011). Proces powstania zapisu dyskretnie kodyfikującego wysokość dźwięku i rytm można postrzegać jako technologiczną „digitalizację” muzyki. Daniel Dennett używa tego terminu do opisu systemów, które umożliwiają korektę do normy, jak w przypadku fonemów w języku, które pozwalają na niezawodną transmisję mimo różnic w wymowie (Dennett, 2017). Notacja menzuralna czyni niemalże to samo dla muzyki: ustanawia dyskretnie, policzalne jednostki wysokości i czasu trwania, przekształcając ciągłe spektrum możliwych dźwięków w skończony zbiór symboli. Partytura staje się w ten sposób algorytmem – zbiorem precyzyjnych, sekwencyjnych instrukcji, które, „uruchomione” przez kompetentnego wykonawcę (procesora), generują przewidywalny i powtarzalny rezultat. To fundamentalnie odróżnia ją od oralnej heurystyki, gdzie wykonawca rekonstruuje utwór na podstawie ogólnych wzorców i formuł komunikowanych w ciągłej, ‘analogowej’ formule.

W memetyce metafory biologiczne należy traktować z dużą ostrożnością, ale warto przytoczyć w tym miejscu krótko analizę filozoficzną Michaela Dicksona (2024). W jego ujęciu partytura pełni funkcję stabilnego, skodyfikowanego genotypu – trwałego zbioru instrukcji dla utworu. Natomiast każde konkretne wykonanie staje się jego fenotypem, czyli fizyczną, dźwiękową manifestacją tego zapisu, która podlega ocenie i presjom selekcyjnym (Nussbaum, 2007). Ich separacja w momencie przejścia do kultury pisanej ma fundamentalne znaczenie, ponieważ w kulturze oralnej taki podział na genotyp i fenotyp nie istniał. Ontologiczna całość dzieła istniała wyłącznie w ulotnym akcie wykonania. Wprowadzenie trwałego genotypu w postaci partytury pozwoliło na stabilizację memu muzycznego i drastyczne ograniczenie wariacji na elementarnym poziomie (Dickson, 2024). Innymi słowy, głęboka memetyczna zmiana zachodzi tutaj też w naturze samego muzycznego replikatora ze względu na zmianę trybu kopiowania. Jeżeli zapis nutowy jest instrukcją do tego jak wykonać muzykę, kopiowanie nut w nomenklaturze Blackmore nazwalibyśmy *copy-the-instruction* – kopiowanie instrukcji, a nie kopiowanie przedmiotu (Blackmore, 1999).

Ta fundamentalna zmiana w trybie kopiowania stanowiła punkt zwrotny w memetycznej ewolucji muzyki Zachodu. W kulturze oralnej, gdzie replikacja opierała się na naśladowaniu samego wykonania, wierność przekazu była z natury niska. Błędy wykonawcze stawały się częścią wzorca dla następnej iteracji memu, prowadząc do nieustannej, rozproszonej transformacji utworu w chmurę wariantów (Nettl, 2005). Zapis nutowy, jako instrukcja, umożliwił zerwanie z tą rozproszoną, nietrwałą formułą ontologiczną. Od tej pory partytura stanowiła nową bazę ontologiczną dla utworu muzycznego. Tym samym błędy oraz intencjonalne i nieintencjonalne wariacje w wykonaniu nie wpływały na samą treść utworu. Od teraz, dzięki digitalizacji poprzez użycie dyskretnego systemu, owa treść mogła być kopiowana ze stuprocentową dokładnością. W praktyce o idealną kopię trudno, ale replikacja mogła odtąd statystycznie zachodzić z niepomiarnie większą dokładnością, a stuprocentowe kopiowanie zostało teoretycznie umożliwione.

To spowodowało skokowy wzrost dwóch kluczowych cech replikatora muzycznego w porównaniu do kultury oralnej: kopiowalności (teraz rozumianej jako wierność zapisu i wykonania) oraz długowieczności - bytowania w nieziennej formule na papierze nutowym zamiast w zawodnej ludzkiej pamięci. Podczas gdy muzyczne memy w kulturze oralnej muszą polegać na swojej zaraźliwej płodności, memy zapisane w nutach zyskują zdolność do potencjalnie nieśmiertelnego trwania i precyzyjnej reprodukcji. Ta ontologiczna transformacja muzyki miała bardzo daleko idące konsekwencje dla jej dalszej ewolucji memetycznej. Technologia pisma uczyniła z muzyki trwałą obiekt, który można było poddać analizie, krytyce i spekulacji (Christensen, 2002). Muzyka przestała być wyłącznie czynnością, a stała się "tekstem", który można było studiować i o którym można było tworzyć dyskursy (Christensen, 2002). W ten sposób powstały idealne warunki dla narodzin i ewolucji nowego, potężnego mempleksu: zachodniej racjonalnej teorii muzyki. Był to początek daleko idącej racjonalizacji praktyki muzycznej, procesu, w którym intuicyjne działanie zaczęło być systematyzowane i objaśniane przez abstrakcyjne zasady. W tym miejscu należy także zaznaczyć, że prostsze i nienotacyjne teorie muzyki istniały od dawna w tradycjach ustnych, gdzie logika muzyczna była przekazywana poprzez praktykę, pamięć i nauczanie ustne, poza reżimem pisma (Lawson, 2010). Jednak wynalazek pisma umożliwia osiągnięcie większej złożoności poprzez kumulatywność i skomplikowaną ewolucję idei.

Rozwój myśli teoretycznej i zapisu nutowego wszedł w pętlę dodatniego sprzężenia zwrotnego. Teoria nie tylko opisywała istniejącą muzykę, ale zaczęła również preskryptywnie kształtować nowe kompozycje, dając twórcom narzędzia do świadomego projektowania. Dziedzictwo greckiej koncepcji muzyki jako matematycznej scientia, przekazane przez Boecjusza, nadało tej racjonalizacji głęboki intelektualny autorytet (Christensen, 2002). Dopiero to połączenie – precyzyjnego nośnika (zapis nutowy) i abstrakcyjnego, racjonalnego systemu (teoria harmonii, kontrapunktu, formy) – umożliwiło powstanie kumulatywnej złożoności. Wielkie formy, takie jak fuga czy forma sonatowa, nie były już tylko produktem ewolucji ślepego zegarmistrza (Dawkins, 1986), ale mogły być świadomie konstruowane w oparciu o zinternalizowany i stale

rozwijany system teoretyczny. Zmieniły się presje selekcyjne: o przetrwaniu memu nie decydowała już tylko jego łatwość zapamiętania, ale również jego zdolność do analizy i potencjał do tworzenia rozbudowanych, intelektualnie satysfakcjonujących struktur. W ten sposób technologia zapisu i myśl teoretyczna wspólnie stworzyły nową, unikalną niszę dla muzyki jako sztuki zapośredniczonej w dyskursie intelektualnym.

Ta nowa, zrjonalizowana nisza przygotowała warunki do fundamentalnej zmiany w samym mechanizmie ewolucji muzyki z oddolnego w kierunku odgórnego. Pojawienie się partytury jako stabilnego obiektu o dyskretnej strukturze otworzyło drogę dla ewolucji odgórnej (top-down), określanej przez Daniela Dennetta mianem inteligentnego projektowania (Dennett, 1995). Kompozytor, uzbrojony w narzędzia zapisu i teorii, przestał być jedynie przekaznikiem i re-kreatorem tradycji. Znacznie poszerzyła się jego zdolność do świadomego planowania złożonych struktur „offline”, poza czasem rzeczywistym wykonania, stając się projektantem, który mógł świadomie eksperymentować z materiałem muzycznym. W tym kontekście ewolucyjny model Dennetta pozwala zinterpretować to przejście jako transformację samego mechanizmu ewolucji kulturowej. Tradycja oralna funkcjonowała na zasadach bliskich ewolucji darwinowskiej: ślepa wariacja i dobór naturalny w ekstremalnych przypadkach prowadziły do powstawania złożonych adaptacji bez udziału świadomego projektanta. Oczywiście w kulturach oralnych także zachodzi świadome projektowanie muzyki, ale bez takich zewnętrznych racjonalnych narzędzi jego zakres jest bardziej ograniczony. Notacja muzyczna natomiast wprowadziła do ewolucji kultury mechanizm, który możemy określić jako lamarckowski. W klasycznym darwinizmie cechy nabyte nie są dziedziczone. Jednak kompozytor pracujący z notacją mógł zaprojektować nową, złożoną cechę muzyczną – na przykład innowacyjną progresję harmoniczną – i zapisać ją bezpośrednio w „genomie” (Dickson, 2024) (czy też „memomie” (Jan, 2024)) dzieła, czyli partyturze. Ta „nabyta” (zaprojektowana) cecha była następnie dziedziczona z bardzo wysoką wiernością przez kolejne iteracje, pokolenia dzieła, czyli jego wykonania i kopie. W ten sposób notacja znacznie powiększyła możliwość dziedziczenia cech nabytych w akcie twórczym, co może wyjaśniać gwałtowne przyspieszenie innowacji i zmian stylistycznych w muzyce Zachodu od późnego średniowiecza.

Rozwój tego nowego, odgórnego trybu ewolucji ma rozległe konsekwencje. Z całą pewnością doprowadził on do ukształtowania się zupełnie nowej niszy dla memów muzycznych. Zmieniły się dotychczasowe presje selekcyjne. Zamiast łatwości zapamiętania i funkcjonalności wspólnotowej, kluczowe stały się nowe wartości. Wśród nich wymienić możemy złożoność strukturalną – notacja umożliwiła tworzenie skomplikowanych form polifonicznych i rozbudowanych architektur formalnych, które byłyby niemożliwe do przekazania w tradycji czysto oralnej.

Muzyczne tradycje oralne również mogą wytworzyć skomplikowane struktury, jak na przykład zaawansowana gra na tabli wspierana przez mnemotechniki (Siedenburg, 2016). Ale sama tradycja oralna nie mogłaby ostatecznie doprowadzić do powstania takich dzieł jak *Święty Franciszek z Asyżu* Oliviera Messiaena czy też 29-godzinny *Licht* Karlheinz Stockhausena. W tradycji zachodniej złożoność stała się wartością estetyczną samą w sobie. Ponadto w niespotykany wcześniej sposób wartość zyskała innowacyjność i oryginalność – w kulturze piśmiennej wartością stało się tworzenie czegoś nowego, co odróżnia kompozytora od jego poprzedników. Wierność tradycji ustąpiła miejsca dążeniu do oryginalności. Kumulacja złożoności w kulturze muzycznej umożliwiła także stworzenie wokół muzyki złożonych systemów relacji społecznych - dyrektorka filharmonii, dyrygent, koncertmistrz, prowadzący sekcji, pierwsze skrzypce, drugie skrzypce itd (Knardal, 2020; Byczkowska-Owczarek, 2022).

Znaczącą rolę zaczął odgrywać autorytet instytucjonalny – selekcja memów zaczęła być kształtowana przez instytucje – dwory, kościoły, a później wydawnictwa i akademie – które decydowały, które „projekty” zostaną sfinansowane, zapisane i rozpowszechnione. W ten sposób przejście od oralności do piśmienności nie było jedynie zmianą nośnika informacji. Była to fundamentalna rekonfiguracja całego ekosystemu ewolucyjnego muzyki, która zmieniła same reguły gry, zastępując powolny, oddolny proces darwinowski nowym, dynamicznym mechanizmem odgórnego, inteligentnego projektowania, co dodatkowo wyjaśniam w kolejnych rozdziałach. Należy przy tym wspomnieć, że muzyka w systemie kultury oralnej działa nieprzerwanie aż do dziś,

koegzystując z muzyką kultury piśmiennej. Najprostszym przykładem w kulturze zachodniej może być muzyka jazzowa: choć wykorzystująca notację muzyczną, oparta jest także w dużym stopniu na oralnej tradycji improwizacji.

3.2.3. Teoria muzyki jako presja selekcyjna

W momencie powstania zachodnioeuropejskiej artystycznej praktyki muzycznej w tradycji piśmiennej, powstała jednocześnie bardzo silna presja selekcyjna w postaci teorii muzyki. Odpowiednie opracowanie tego tematu mogłoby stanowić temat na odrębną dysertację, zatem w tym miejscu zaledwie go zasygnalizuję. Relacja muzyki do muzycznej teorii oraz muzycznej wiedzy stanowi temat złożony i niejednoznaczny. Zauważenie tej skomplikowanej korelacji jest jednym z ciekawych punktów wyjścia do refleksji o ewolucji kulturowej muzycznej tradycji piśmiennej. Klasyczna edukacja muzyczna może błędnie sugerować, że teoria muzyki jest niezachwianą nauką podstawą, elementarnym budulcem, z którego wielcy kompozytorzy tworzyli swoje wybitne dzieła. Dziś możemy relację muzyki i teorii opisać zgoła odmiennie: to dwa współzależne systemy, które nawzajem na wiele sposobów się kształtują. W tym także wzajemnie racjonalizują i legitymizują w tautologicznej pętli. A zatem presje selekcyjne kształtujące teorię rzutują również w złożony sposób na muzykę.

Zasygnalizuję dosłownie jeden wybrany przykład: preferencja do teorii pięknych czy też eleganckich, niezależnie od znajomości statusu ich słuszności, tak, jak np. preferencje względem prostych proporcji matematycznych pomiędzy częstotliwościami dźwięków systemów muzycznych w sporze o temperacje (Tenney, 1988; Parncutt, 2018). Zauważam zatem łańcuch oddziaływań. Fascynacja proporcją liczbową wpływa na koncepcję muzycznej temperacji. Temperacja kształtuje ramy pojęcia konsonansu, a zatem też zasady harmoniki. Harmonika w znacznym stopniu determinuje cechy dzieła muzycznego. A zatem inicjalny zachwyty nad ideą proporcji liczb rzutuje na brzmiającą strukturę utworu.

Nowoczesna metodologia nauki dostrzega rolę takiej praktyki (McAllister, 1996). Wybór eleganckiej teorii łatwo można pomylić z racjonalnym użyciem brzytwy Ockhama, ostatecznie ulegając zniekształceniu poznawczemu. Szczególną wagę ma tutaj sama

estetyczno-matematyczna podstawa zachodniej teorii muzyki, czyli powiązanie dźwiękowego piękna z liczbami całkowitymi (Bowling, 2015). Dobrze pasuje tutaj cytata słynnego fizyka i matematyka Paula Diraca:

It is more important to have beauty in one's equations than to have them fit experiment.

Ważniejsze, aby wzory matematyczne były piękne niż zgodne z eksperymentami.

(Dirac, 1996)

Dirac wprost przekłada estetyczne walory idei nad pomiar empiryczny. Tego typu myśleniem obarczona jest zachodnia teoria muzyki, a w szczególności harmonia, od czasów pitagorejskich aż do dziś. Fundamentalne, starożytne odkrycie, że proste stosunki liczbowe długości strun (np. 2:1, 3:2, 4:3) odpowiadają interwałom postrzeganym jako konsonansowe (oktawa, kwinta, kwarta) i zarazem piękne, stało się kamieniem węgielnym teorii muzyki (Parncutt, 2018; Bowling, 2015). Ponadto ta empiryczna obserwacja została podniesiona do rangi zasady metafizycznej: muzyka stała się odzwierciedleniem matematycznej harmonii kosmosu, a liczby – kluczem do zrozumienia jej istoty (Bowling, 2015). W ten sposób na gruncie teorii muzyki doszło do pomieszania fizyki, estetyki i metafizyki, co dało początek spekulatywnej tradycji, która przez wieki, aż do XVII stulecia, zajmowała się kontemplacją muzycznych esencji w kategoriach liczbowych. Wczesny, kluczowy błąd, na którym zaważyły dalsze losy zachodniej teorii muzyki, polegał na przedłożeniu eleganckiej, abstrakcyjnej zasady nad złożoność percepcji i praktyki muzycznej (Wiggins, 2010; Marijeh, 2024).

W Oświeceniu Jean-Philippe Rameau próbował zracjonalizować teorię muzyki, odchodząc od metafizyki liczb do empirii zjawiska rezonansu (Rameau, 1726), jednak nowe, „naukowe” odkrycie posłużyło mu jedynie do uzasadnienia tych samych, z góry przyjętych zasad harmonicznycch (Christensen, 1993), co pokazuje, jak estetyczne preferencje mogą kształtować interpretację danych naukowych. Przykład Rameau opisuje odwrotny kierunek oddziaływania.

3.2.4.1. Wprowadzenie ANT: specjacja muzyki akademickiej

Wynalazek notacji muzycznej umożliwił powstanie w muzyce kultury piśmiennej. Stanowiło to fundamentalną zmianę paradygmatu ewolucyjnego. Zmienił się nie tylko nośnik informacji, ale cała sieć relacji otaczających funkcjonowanie muzyki. Z punktu widzenia memetyki możemy określić ten proces jako specjację pokrewnych, sprawczych bytów. Informacje muzyczne rozproszyły się również na memetyczne potomstwo poza muzyką brzmiącą. Powstał nowy współgrający ekosystem z takimi nowymi elementami jak instytucje, teorie i procesy kształcenia. Taką egzotycznie brzmiącą analizę legitymizuje socjologiczna teoria aktora-sieci (Actor-Network Theory, ANT) (Latour, 2005). ANT nie uznaje technologii i instytucji za pasywne tło dla ludzkich działań, lecz konceptualizuje je jako aktywnych „aktorów” (aktantów), współtworzących rzeczywistość społeczną wraz z ludźmi (Latour, 2005).

ANT proponuje radykalnie symetryczne podejście do analizy zjawisk społecznych, w którym zarówno ludzie, jak i nie-ludscy aktorzy (np. technologie, teksty, instytucje) posiadają sprawczość (agency) i mogą wchodzić w relacje tworzące heterogeniczne sieci (Latour, 2005). Teoria ta odrzuca tradycyjne dychotomie, takie jak społeczeństwo/natura czy podmiot/przedmiot, na rzecz badania „płaskiej ontologii”, w której stabilność zjawisk społecznych jest efektem ciągłej pracy nad utrzymaniem i rekonfiguracją powiązań między różnorodnymi aktantami (Callon, 1986). Z tej perspektywy, wynalazek lub technologia nie jest po prostu zewnętrznym narzędziem używanym przez ludzi, ale aktywnym uczestnikiem, który przekształca sieć, tworzy nowe role, definiuje nowe cele i zmienia dynamikę władzy (Law, 1992).

Wynalazek notacji muzycznej z perspektywy teorii aktora-sieci jest potężnym nie-ludzkim aktantem, który rekonfiguruje cały muzyczny ekosystem. Oznacza to, że pojawienie się notacji wywołało całą lawinę zdarzeń, które nieodwracalnie zmieniły muzyczne sieci zależności. Te sieci zostały także znacznie wzbogacone o nowe,

sprawcze elementy. Szczegółowy opis memetycznej ewolucji sprawczych instytucji muzycznych wzbogacony o perspektywę cybernetyczną przedstawiam w kolejnym rozdziale. Obecnie skupiam się na przedstawieniu charakterystyki ewolucyjnej systemu muzyki akademickiej, który powstał w wyniku rozwijania technologii muzycznego zapisu.

Tutaj szczególną sprawczość zyskuje sama teoria muzyki. Skodyfikowana w traktatach, przestaje jedynie opisywać istniejącą praktykę, a zaczyna ją preskryptywnie kształtować. Proces ten można ująć w kategoriach racjonalizacji w rozumieniu Maxa Webera (Krzyżanowski, 2017): poddania dziedziny artystycznej rygorom obliczalnych reguł i zrozumiałych zasad (Weber, 1958). Za przejawy racjonalizacji możemy uznać m.in. rozwój temperacji stroju (Sköld, 2022), teorii harmonii (Cohen, 2022) czy ustandaryzowanej notacji (Schuiling, 2019). Dążenie do uczynienia z muzyki systemu podlegającego przejrzystej analizie i kontroli możemy rozumieć dwojako. Z jednej strony sama muzyczna teoria wytwarza memetyczną presję selekcyjną i rozwija się, coraz mocniej oddziałując na brzmiącą muzykę. Po drugie ewoluuje i racjonalizuje się cały społeczno-instytucjonalny system muzyczny, a najsprawniejsze funkcjonowanie systemu zapewniają przejrzyste i jasno skwantyfikowane dane. Zatem w szerszej perspektywie ewolucja całego systemu instytucji także wywoływała odgórną (top-down) presję na sam materiał muzyczny, aby optymalizować własne działanie.

W ten sposób powstaje dodatkowa pętla sprzężenia zwrotnego. Rosnąca złożoność muzyki stwarza zapotrzebowanie na bardziej zaawansowaną teorię i notację, a te z kolei umożliwiają tworzenie jeszcze bardziej skomplikowanych kompozycji. Ten ewolucyjny wyścig zbrojeń w dziedzinie złożoności staje się istotnym motorem napędowym rozwoju zachodniej muzyki artystycznej. Powstanie instytucji takich jak Konserwatorium Paryskie (ok. 1795) stanowiło kluczowy krok w procesie standaryzacji edukacji muzycznej i profesjonalizacji zawodu muzyka (Rubinoff, 2015). Zgodnie z teorią Pierre'a Bourdieu, ta instytucjonalizacja doprowadziła do wyłonienia się autonomicznego pola muzycznego (fr. *champ musical*) – przestrzeni społecznej z własnymi regułami gry, hierarchiami i formami kapitału (Bourdieu, 1984).

W polu tym podstawową walutą staje się nie kapitał ekonomiczny, lecz kapitał symboliczny: prestiż, uznanie i legitymizacja. Można powiedzieć, że sama ogólna pula kapitału ewoluowała, poszerzyła się. Mechanizmem umożliwiającym skuteczną konwersję jednego kapitału na drugi został patronat artystyczny. Ta nowa sieć, złożona z ludzkich i nie-ludzkich aktorów, redefiniuje pole memetycznego sukcesu muzyki. Dominujący mem to już nie zaraźliwy earworm, ale taki mem muzyczny, który z powodzeniem nawiguje w obrębie reguł i wartości nowego, wyspecjalizowanego pola. W ramach pola w rozumieniu Bourdieu, sądy estetyczne zawsze niosą za sobą pewne społeczne konsekwencje. Gust staje się podstawowym mechanizmem społecznej dystynkcji. Zdolność do docenienia muzyki „trudnej” i „złożonej” – czyli muzyki klasycznej – staje się formą kapitału kulturowego, wyznacznikiem przynależności do elity. Z perspektywy biologii ewolucyjnej mechanizm ten można dodatkowo wyjaśnić poprzez teorie sygnalizacji, zgodnie z którymi wytwarzanie i odbiór złożonych form kulturowych pełnią rolę kosztownych sygnałów jakości jednostki. Czyli tworzenie lub docenianie prestiżowej sztuki to sposób konstruowania atrakcyjności w doborze płciowym. Geoffrey Miller (2000) podkreśla, że zdolności artystyczne i estetyczne mogły rozwinąć się w ramach doboru płciowego jako forma wykazania kondycji poznawczej. Jednocześnie zjawisko to wpisuje się w szerszy kontekst grupizmu, w którym preferencje estetyczne i style kulturowe wzmacniają spójność wewnątrzgrupową i różnicowanie międzygrupowe (Kurzban, 2005).

„Koneser” odrzuca „naiwną” przyjemność estetyki ludowej. Szuka za to zdystansowanej, intelektualnej kontemplacji. Cel optymalizacji w tej nowej niszy ewolucji muzyki ulega zmianie. Można też powiedzieć, że w ogóle pojawia się cel, a wcześniej go nie było (co omawiam dalej przy pomocy heurystyki przestrzeni Darwinowskiej). Takie spojrzenie sugeruje już, że przechodzimy od globalnej historycznej niszy bliskiej memetycznemu doborowi naturalnemu do wyspecjalizowanej niszy napędzanej doborem sztucznym. Oddolnie powstające zaraźliwe melodie tracą tu moc. Zmieniają się reguły gry, czy też presje selekcyjne. Liczy się tworzenie dzieł, które sygnalizują i nadają prestiż. To jest właśnie rdzeń mechanizmu doboru sztucznego w

tym kontekście: muzyka zaczyna być „hodowana” pod kątem specyficznej, subiektywnej cechy estetycznej.

Zapis nutowy umożliwia powstanie długotrwałych, stabilnych dzieł. Takie konkretne dzieło z czasem akumuluje przeróżne połączenia w sieci szerszego muzycznego systemu. Na przykład IX Symfonia Beethovena, a w szczególności Oda do Radości w latach 1974–1979 stała się hymnem narodowym Rodezji (Deschamps, 2021), następnie została formalnie przyjęta jako hymn Rady Europy w 1972 roku (Deschamps, 2021) i Unii Europejskiej w 1985 roku (European Union, 2025). Symfonia zakorzeniła się również w tradycjach narodowych, czego przykładem jest zapoczątkowany w 1918 roku przez niemiecki ruch robotniczy zwyczaj wykonywania jej w Sylwestra (Rehfeldt, 2020) oraz japoński coroczny rytuał Daiku zapoczątkowany około 1925 roku (Nippon, 2013). Między innymi spośród tak głęboko usieciowionych dzieł wyłania się muzyczny kanon. Możemy rozumieć go jako względnie stabilny mempleks złożony ze zbioru dzieł, które mają centralne znaczenie dla całego systemu.

Powstaje także „Kult geniusza” (Denora, 1997). Pośmiertne funkcjonowanie Beethovena czy Chopina można wytłumaczyć jako wysoce skuteczne i stabilne mempleksy: koadaptacyjne zbiory wzajemnie wzmacniających się memów (idei, zachowań, narracji). Mempleksy te obejmują ideę kompozytora jako cierpiącego, niezrozumianego półboga, którego dzieło przekracza swoją epokę. Samo dzieło ulega niemal sakralizacji wskutek emanacji tegoż geniuszu. Zmarły kompozytor staje się potężnym, nie-ludzkim aktantem w ramach sieci. Jego nazwisko może funkcjonować jako marka (Wódka Chopin) albo trendujący hashtag (#bach). Wykonawcy, krytycy i instytucje zdobywają kapitał symboliczny poprzez afiliację z tą marką. W zamian ich działalność (koncerty, wydania nutowe, festiwale) zapewnia zmarłemu geniuszowi ciągłą replikację i kulturową relewantność. W ten sposób tworzymy pętlę symbiotyczną.

Przykładowo, Międzynarodowy Konkurs Pianistyczny im. Fryderyka Chopina może stanowić studium przypadku instytucjonalnego aktanta, który aktywnie zarządza i propaguje mempleks genialnego Chopina. Poprzez rygorystyczne wymogi repertuarowe

i kryteria oceny, konkurs kanonizuje określony zbiór dzieł. Promuje przy tym specyficzny, wysoce wirtuozowski, ale też historycznie poinformowany styl wykonawczy. Dokładnie ten styl staje się globalnym standardem.

Zwycięstwo w konkursie przynosi pianiście ogromny kapitał symboliczny i staje się trampoliną do międzynarodowej kariery. Owa trampolina to właśnie usieciwienie. Laureat zyskuje widoczność jako silny wektor replikacji mempleksu Chopina w przestrzeni kulturowej. Artysta wykonuje chopinowskie dzieła na całym świecie i wzmacnia jego kanoniczny status. Konkurs aktywnie angażuje także państwowych i korporacyjnych sponsorów. PKN Orlen umacnia swoją rolę jako mecenasu kultury oraz zyskuje ludzką twarz. Również firma Google zaangażowała się w promocję XVIII Konkursu Chopinowskiego (Różalska, 2021). Kapitał symboliczny i ekonomiczny prowadzą złożoną grę, w wyniku której mempleks Chopina stanowi bijące serce złożonej sieci aktorów i aktantów. W tym miejscu pokuszę się o nakreślenie pewnej analogii. Do tego celu posłuży mi Dennettowska memetyczna koncepcja przestrzeni Darwinowskiej (Dennett, 1995, 2017). Stanowi ona heurystykę, dzięki której można umiejscowić procesy ewolucji kulturowej w szerokim spektrum. Określa ono źródło wariacji - ślepa lub ukierunkowana oraz rodzaj doboru - sztuczny lub naturalny. Dzięki temu w memetycznej ewolucji możemy rozróżnić oddolny (bottom-up) i odgórny (top-down) mechanizm projektowania nowości.

Jedno ekstremum stanowi oddolna ewolucja kulturowa, która w znacznej mierze polega na procesie doboru naturalnego. Wariacja powstaje poprzez pomyłki, błędy poznawcze i przypadkowe rekombinacje. Nie ma tu działania podpartego intencją. Powstały w ten sposób muzyczny mem nie ma kompozytora. Mechanizm projektowania działa oddolnie, mem ewoluuje w toku zbiorowej imitacji i stopniowo, przypadkowo zyskuje cechy stanowiące o przystosowaniu do środowiska. Dobór działa tutaj w sposób rozproszony, wynika głównie z chwilowych stanów ludzkiej pamięci i uwagi. Doskonały przykład tego krańca darwinowskiej przestrzeni stanowi replikacja earworms (Jakubowski, 2017), a szczególnie proto-earworms w prehistorii, zgodnie z opisem z początku tego rozdziału. Dziś earwormem często może być np. muzyka z celowo

zaprojektowanej reklamy, co przesuwają już to zjawisko na dalszy obszar omawianej darwinowskiej przestrzeni, w kierunku drugiego bieguna.

Tenże drugi biegun to odgórną (top-down) ewolucja kulturowa, która przypomina dobór sztuczny. Nowe warianty są celowo, intencjonalnie projektowane przez kompozytorów. Muzyka jest racjonalnie projektowana, jej cechy są świadomie kształtowane, żeby spełniać konkretne normy, skutecznie funkcjonować w określonych środowiskach. Kompozycja serialna, w szczególności metoda dwunastotonowa Schönberga, reprezentuje ten ekstremalny biegun. W tym przypadku generatywne reguły systemu są sztywno określone. Odgórną kompozycji towarzyszą jawne cele, takie jak zyskanie prestiżu i pieniędzy. Powstaje cały racjonalny plan obiegu instytucjonalnego: kompozycja precyzyjnie spisana w systemie nutowym nadaje się do druku i wykonania w zespole. Wierność przekazu jest podtrzymywana przez konserwatoria i szkoły kompozycji (Taruskin, 2005).

Moja analogia ma nakreślić to, że ewolucja bliska obu krańcom spektrum prowadzi do powstawania systemów o bardzo rozbieżnych kompetencjach replikacyjnych. Mianowicie, pierwotna nisza memetyczno-muzyczna, pełna earwormów, musiała ewoluować na sposób bliski do bezcelowego doboru naturalnego. Przystosowanie, zawsze lokalne i chwilowe, do otaczającego środowiska, wytworzyło silny i wszechstronny system. Przypominał on dziką przyrodę, także kształtowaną przez szeroki wachlarz presji środowiskowych. Te systemy wiążę zatem z prężnymi i adaptacyjnymi fenotypami oraz memami. Muzyka akademicka jest natomiast przesycona intencjonalnością, celowością, złożonym, racjonalnym działaniem. Analogiczna do rasy udomowionej, wyhodowanej przez ludzkich agentów (sieć instytucjonalną) w celu uzyskania wąskiego zestawu pożądanych cech, takich jak złożoność, oryginalność, prestiż. To plasuje metodę ewolucji systemu muzyki akademickiej blisko doboru sztucznego, znanego chociażby z hodowli psów.

Intensywna selekcja kierunkowa na rzecz złożoności i teoretycznego uzasadnienia w zamkniętym polu akademickim jest analogiczna do chowu wsobnego (inbreeding).

Proces ten prowadzi do zubożenia wariacji memetycznej. Pula memowa idei muzycznych staje się coraz mniejsza i bardziej homogeniczna, gdy kompozytorzy tworzą w coraz większym stopniu w odpowiedzi na teorię i dzieła innych kompozytorów z tej samej tradycji, a nie w odpowiedzi na szersze presje społeczne czy poznawcze. Tak powstają ewolucyjne kompromisy: optymalizacja pod kątem złożoności odbywa się kosztem dostępności percepcyjnej i ogólnej sprawności kulturowej.

Muzyka akademicka poprzez ewolucję w zamkniętej, chronionej niszy instytucjonalnej, utraciła kontakt z szerszym środowiskiem naturalnym – gustem publiczności. Podobnie jak zwierzęta hodowlane, które tracą zdolność przetrwania na wolności, stała się ona krucha i zależna od sztucznego systemu wsparcia. Jak wskazuje teoria „choroby kosztów” Baumola, koszty w sektorach o niskim wzroście produktywności (jak sztuki performatywne) nieuchronnie rosną. Prowadzi to do strukturalnego deficytu, który musi być pokrywany z dotacji (Baumol, 1966). Zależność od subwencji bezpośrednio wskazuje na ewolucyjne nieprzystosowanie tej formy muzyki do świata kapitalizmu.

Zatem system akademicki napędzają subsydia od patronów – osób prywatnych, korporacji, rządów. Patronat ten nie jest czystym altruizmem; jest to wymiana, w której patroni dostarczają kapitału ekonomicznego, a w zamian otrzymują kapitał symboliczny – prestiż związany ze wspieraniem „sztuki wysokiej”. System ten izoluje muzykę akademicką od presji selekcyjnych rynku masowego. Pozwala jej przetrwać i ewoluować w chronionej, sztucznej niszy – swoistym metaforycznym skansenie albo akwarium. Muzyka nowa, współczesna, a szczególnie dodekafonia i serializm stanowią skrajny przykład tej tendencji. Systemy te są kulminacją procesu racjonalizacji. Zerwały z zasadami percepcyjnymi, które kształtowały muzykę, stanowiąc źródło doboru naturalnego. Jak argumentował Fred Lerdahl (1992), istnieje w nich ogromna przepaść między gramatyką kompozytorską (złożonymi, matematycznymi regułami) a gramatyką słuchacza (zdolnością do poznawczego przetworzenia struktury). Z perspektywy memetycznej, memy dodekafoniczne okazały się bezpłodne, czyli niezdolne do skutecznej replikacji w umysłach słuchaczy poza wąską grupą specjalistów. Nawet Theodor Adorno (1949), choć sam był obrońcą Schoenberga, dostrzegał w totalnej

racjonalizacji muzyki dialektyczne napięcie i ryzyko utraty subiektywnej ekspresji. Muzyka akademicka utraciła prężność (robustness) charakterystyczną dla tradycji oralnych. Jest krucha i wymaga rozbudowanego systemu wsparcia: konserwatoriów do kształcenia wykonawców, specjalistycznych sal koncertowych do jej prezentacji oraz złożonego dyskursu muzykologicznego do jej legitymizacji.

Z perspektywy teorii aktora-sieci, proces ten nie był napędzany wyłącznie ludzką intencją. Był to raczej emergentny efekt działania heterogenicznej sieci, w której ludzcy aktorzy (kompozytorzy, teoretycy, patroni) wchodzili w złożone interakcje z nie-ludzkimi aktantami: technologiami (notacja, druk, nagrania dźwiękowe), instytucjami (kościół, dwór, konserwatorium, filharmonia) oraz samymi ideami (teoria harmonii, koncepcja dzieła, kult geniusza). Ta sieć, koewoluując, stworzyła samowzmacniający się i samoreferencyjny system, który zaczął 'żyć własnym życiem'. Analogia doboru sztucznego, choć może wydawać się prowokacyjna, nie ma charakteru wartościującego. Wyjaśnia, że choć muzyka akademicka wydała dzieła o niezrównanej złożoności intelektualnej, dokonało się to kosztem szerokiej transmisyjności kulturowej. Muzyka akademicka posiada unikalną i fascynującą ścieżkę ewolucji kulturowej. Stworzyła ona piękne, skomplikowane, a zarazem kruche zjawisko. Dziś jest podtrzymywana przy życiu przez te same systemy przepływu kapitału symbolicznego, które powołały ją do istnienia.

3.2.4.2. Sprawcze instytucje od kościoła po MTV

W poprzednim rozdziale wstępnie zarysowałem koncepcję instytucji muzycznej w ramach ANT, aby opowiedzieć o powstaniu ewolucyjnej niszy Zachodniej muzyki artystycznej. Bieżący rozdział rozwinie tę koncepcję instytucji o dodatkowe wymiary. Ukáže mempleksowy charakter muzycznych instytucji na podstawie literatury muzykologicznej i nie tylko. To bardzo szeroki problem, ponieważ instytucje od wieków odgrywały kluczową rolę w kształtowaniu i przekazywaniu kultury muzycznej (Wieczorek, 2013; Cohen, 2009). Ze względu na ich specyficzną strukturę oraz zdolność do wywierania wpływu na społeczeństwo konceptualizuję je jako agencjalne byty społeczne.

W historii muzyki znamy wiele przykładów takich sprawczych instytucji: Kościół (Crook, 2009; Tyrała, 2020), dwory mecenasów (Reynolds, 1997; Wieczorek, 2013), konserwatoria (DelDonna, 2020; Masci, 2022), wytwórnie fonograficzne (Katz, 2010; Philips, 2013) czy szkoły (Campbell, 2022; Southcott, 2003). Organizują one praktykę muzyczną, regulują jej normy i propagują idee – czyli generują replikację i ewolucję różnorodnych memów muzycznych. Co więcej – instytucje same w sobie fundamentalnie składają się z informacji (Weick, 1995), to znaczy ich strukturę możemy zrozumieć jako rozbudowany dynamiczny mempleks. Aby mempleks instytucji faktycznie działał jako silnik do replikacji memów, musi być on napędzany przez ludzi, pracowników. Tym niemniej niniejszy rozdział ma na celu wykazać, że instytucje, w szczególności instytucje muzyczne, są tylko w ograniczonym stopniu zależne od woli i intencji tychże poszczególnych uczestników.

Po wprowadzeniu historycznym przeprowadzę autorską syntezę trzech różnych koncepcji naukowych, żeby ukazać, że podobieństwo złożonych instytucji do organizmów żywych to więcej niż metafora. Zgodnie z interpretacją teorii złożzeń stanowią one wręcz pewną formą życia (Walker, 2024). Jest to również całkowicie zgodne z twierdzeniem szeroko cytowanego w memetyce muzyki Daniela Dennetta,

który definiuje człowieka jako organizm hybrydyczny złożony z memów i genów (Dennett, 2017). Instytucje reprezentują tutaj pewną manifestację inteligencji zbiorowej, rządzoną przez te same fundamentalne zasady, które napędzają ewolucję złożoności biologicznej. Posługując się biologiczną metaforą, proponuję wyjaśnienie instytucji jako informacyjnego egzoszkieletu nadbudowanego nad ludzką zbiorowością na drodze koewolucji genetyczno-memetycznej. Mam na myśli społeczną analogię do wprowadzonego przez Michaela Levina pojęcia kleju poznawczego (Levin, 2023), którym autor wyjaśnia przede wszystkim współdziałanie struktur poznawczych wewnątrz poszczególnych organizmów. Levin (2023) opisuje sygnalizację bioelektryczną jako „poznawczy klej, który umożliwia ewolucji skalowanie od fizjologii do umysłu”, podkreślając, że sieci bioelektryczne integrują aktywność komórkową w wyższe poziomy poznania (jw.). Autor wyjaśnia, że „każda inteligencja to zbiorowa inteligencja”, co wskazuje, iż poznanie wyłania się z koordynacji bioelektrycznej, a nie wyłącznie z izolowanych procesów neuronalnych (jw.). W tym ujęciu sieci bioelektryczne stanowią “obliczeniowe medium, które pozwala na spójną propagację wzorców wyższego rzędu poprzez tkanki, w efekcie działając jak klej, który spaja poznawcze całości” (jw).

Idąc tym tropem, spróbuję wykroczyć poza metafory, argumentując za dosłownym zastosowaniem ewolucyjnych zasad biologicznych oraz memetycznych do instytucji jako systemów kulturowych. Fundamentalne wyzwanie teoretyczne, które stoi u podstaw tej analizy, dotyczy sposobu, w jaki zbiór różnorodnych bytów – ludzi, technologii, praw, idei – może zjednoczyć się w autonomicznego agenta, posiadającego własne cele, z których nadrzędnym jest przetrwanie i samoreprodukcja. Ta systemowa sprawczość jest w stanie nadrzędnie traktować cele poszczególnych ludzkich komponentów, a nawet je zastępować, jeśli nie działają one w zgodzie z dążeniem systemu do przetrwania.

Wprowadzę zatem syntezę wcześniej wspomnianych trzech filarów teoretycznych w jeden, spójny model. Nie są to trzy różne perspektywy na ten sam obiekt, lecz opisy trzech różnych, lecz w pełni zintegrowanych warstw tego samego obiektu.

Wieloskalowa Architektura Kompetencji (MCA) Michaela Levina (Levin, 2022) dostarcza fundamentalnej ontologii zagnieżdżonej, rozwiązującej problemy sprawczości. Teoria Aktora-Sieci (ANT) Bruno Latoura (Latour, 2005) opisuje anatomię – materialną kompozycję organizmu instytucjonalnego. Model Systemu Zdolnego do Przetwarzania (VSM) Stafforda Beera (Beer, 1972) opisuje z kolei jego fizjologię – cybernetyczny system kontroli, który zapewnia żywotność. Instytucję definiuję zatem jako wieloskalowo kompetentną hierarchię (MCA) złożoną z agencjalnych części (ANT), zorganizowaną przez cybernetyczną strukturę kontrolną (VSM).

Na początek prezentuję krótki zarys historyczny przykładowych istotnych instytucji muzycznych, których działanie następnie opisuję w modelach teoretycznych i meta-teoretycznym. Zarys historyczny instytucji kształtujących zachodnią praktykę muzyczną rozpoczniemy od średniowiecza. Rolę hegemonicznego mecenasa i jednego z ośrodków władzy pełnił wtedy Kościół katolicki (Hoppin, 1978). Poza innymi sferami dominacji wyznaczał on również fundamentalne ramy dla twórczości artystycznej. W dziedzinie muzyki Kościół funkcjonował jako kluczowa instytucja sprawcza (Planchart, 2018). Jego wpływ obejmował kontrolę nad mecenatem, kształtowanie kanonu repertuarowego oraz definiowanie norm estetycznych (Planchart, 2018; Fenlon, 2009).

Normatywna funkcja Kościoła najpełniej przejawiała się w muzyce liturgicznej. Formy takie jak chorał gregoriański były produktem i narzędziem polityki unifikacji kultu (Corno, 2022). Estetyka chorału odzwierciedlała cele teologiczne, takie jak sprzyjanie kontemplacji i podkreślanie wagi słowa Bożego. Instytucjonalne ramy, narzucane przez hierarchię kościelną, określały zarówno dopuszczalny repertuar jak i zasady jego wykonania (Taruskin, 2010). W Europie około połowy XV w. kontrola kościoła nad muzyką zaczęła słabnąć (Busse Berger, 2015).

Z perspektywy ANT można powiedzieć, że europejska średniowieczna wspólnota kościelna stanowiła złożoną sieć dynamicznych powiązań. Obejmowała ona duchownych, instytucje edukacyjne (szkoły przyklasztorne, uniwersytety), skrybów

przepisujących manuskrypty oraz sam repertuar. Czyli sprawczość była rozproszona pomiędzy muzyczne memy, mempleksy oraz ludzkich aktorów. Ta sieć, utrzymywana przez reguły (np. jednolity kalendarz liturgiczny i notacja), przejawiała się jako spójny podmiot instytucjonalny kierujący rozwojem muzyki sakralnej.

W renesansowej Europie narodził się system mecenatu świeckiego, w którym dwory monarsze i arystokratyczne stały się głównymi ośrodkami finansowania i organizacji życia muzycznego (Wieczorek, 2013). Władcy i możnowładcy traktowali muzykę jako element prestiżu oraz symboliczny kapitał w rywalizacji politycznej. Kompozytorzy i wykonawcy byli zatrudniani na dworach, co zapewniało im środki utrzymania w zamian za tworzenie i wykonywanie muzyki. We Włoszech i Francji powstawały akademie takie jak Camerata Florencka, które skupiały elitę intelektualną i artystyczną celem dyskusji o sztuce i eksperymentów muzycznych (Katz, 1984; Rion, 2004). Warto przy okazji dodać, że z działalności Cameraty zrodziła się pierwsza opera barokowa. Takie akademie pełniły funkcję instytucji animujących rozwój nowych form muzyki, co czyniło je zbiorowymi aktorami na scenie kultury (The Editors of Encyclopaedia Britannica, 2023).

W XVII i XVIII wieku model instytucji muzycznej ewoluował wraz z rozwojem życia koncertowego. W miastach niemieckojęzycznych, często przy uniwersytetach, działały tzw. Collegia Musica (Bogdan, 2018). Były to zrzeszenia muzyków i studentów, które organizowały regularne koncerty. Tak wyglądały narodziny publicznej instytucji muzycznej niezależnej od patronatu dworskiego czy kościelnego (Grout, 2001).

W całej Europie powstawały również formalne towarzystwa koncertowe, takie jak Academy of Ancient Music w Londynie około 1726 (Johnstone, 2020), paryskie Le Concert Spirituel w 1725 (Pesic, 2021) czy wiedeńskie Tonkünstler-Sozietät w 1771 (Mellace, 2014). Ich celem była organizacja cyklicznych, biletowanych koncertów. Tak właśnie ugruntowała się pozycja muzyki jako elementu życia społecznego dostępnego dla nowej, mieszczańskiej publiczności (Weber, 1975). Oświeceniowe ideały wzmacniały przekonanie, że muzyka jest dobrem publicznym. Znalazło to odzwierciedlenie w

tworzeniu stowarzyszeń śpiewaczych, jak Sing-Akademie zu Berlin w 1791 (Rutledge, 2011), oraz pierwszych nowoczesnych konserwatoriów, z Konserwatorium Paryskim w 1795 (Rubinoff, 2017) na czele (Taruskin, 2010). Te byty instytucjonalne dysponowały własną strukturą zarządczą, regulaminami i celami. Propagowały zarówno muzyczną wiedzę jak i konkretny repertuar. Możemy zatem postrzegać je jako odrębnych aktorów życia społeczno-kulturowego.

Wiek XIX przyniósł fundamentalną transformację w organizacji życia muzycznego. Napędzały ją przez procesy urbanizacyjne i wzrost znaczenia zamożnej klasy średniej (Scott, 2001). To właśnie mieszczaństwo, szukające sposobów na manifestację swojego statusu i aspiracji kulturalnych, stało się głównym motorem i odbiorcą nowych instytucji muzycznych (Weber, 1975). W największych miastach Europy i Ameryki zaczęły powstawać liczne towarzystwa filharmoniczne i chóry, działające na zasadzie stałych, subskrypcyjnych koncertów. Do najważniejszych należały wiedeńskie Gesellschaft der Musikfreunde założone w 1812 (Pace, 2012), londyńskie Philharmonic Society z 1813 (Ehrlich, 1995) czy paryskie Société des Concerts du Conservatoire z 1828 (Holoman, 2004).

Te instytucje nie tylko organizowały koncerty. Ponadto, świadome formowały one kanon arcydzieł, które stopniowo zyskały status ponadczasowych i autonomicznych dzieł sztuki. Orkiestry i towarzystwa konsekwentnie powtarzały w repertuarze utwory Bacha, Haydna, Mozarta i Beethovena. Lydia Goehr uważa, że w ten sposób stworzyły one swoiste „wymaginowane muzeum dzieł muzycznych”, oddzielając muzykę poważną od tej o charakterze czysto użytkowym czy rozrywkowym (Goehr, 1992). Pociągało to za sobą błyskawiczny rozwój wydawnictw nutowych. Firmy takie jak Breitkopf & Härtel (Rowland, 2019) czy Ricordi (Adams, 2023) stały się potężnymi graczami. Pośredniczyły one między kompozytorem a wykonawcami i szerokim gronem amatorów. Koncertowa popularność utworu bezpośrednio przekładała się na masową sprzedaż partytury z jego transkrypcją fortepianową lub kameralną, a wydawcy aktywnie promowali swoich kompozytorów (Taruskin, 2010).

Pod koniec stulecia ukształtowały się również prawne podwaliny nowoczesnego rynku. We Francji powstało SACEM (Stowarzyszenie Autorów, Kompozytorów i Wydawców Muzycznych, 1851) – pierwsza na świecie organizacja zbiorowego zarządzania prawami autorskimi (Ficsor, 2002). W tym okresie instytucje muzyczne stały się bardziej sformalizowane. Musiały także przystosować się do zmieniających się gustów i technologii (Weber, 1994): filharmonie rywalizowały o publiczność i sponsorów (jw.), a wydawcy prześcigali się w produkcji opracowań na domowe pianino (November, 2021; Christensen, 1999). Wszystko to wskazuje, że instytucje te posiadały pewną samosterowność – reagowały na otoczenie (popyt publiczności, trendy artystyczne) poprzez zmiany repertuaru, formuły koncertów czy modeli finansowania.

XX wiek przyniósł kolejną transformację krajobrazu instytucjonalnego w muzyce. Wraz z popularyzacją nagrań dźwiękowych, radia, a później telewizji powstała nowa klasa instytucji medialno-komercyjnych (Krämer, 2011). Były to wytwórnie płytowe, rozgłośnie radiowe, firmy fonograficzne, a w drugiej połowie wieku również stacje telewizyjne (Miller, 2022) i globalne koncerty rozrywkowe. Te podmioty przejęły w znacznym stopniu funkcję mecenatu i dystrybucji: to one decydowały, którzy artyści i jakie nagrania zyskają ogólny zasięg (Hesmondhalgh, 2019). Można tu wskazać na analogię do roli Kościoła w średniowieczu. Duże wytwórnie poprzez swoje strategie marketingowe i selekcję repertuaru narzucały pewien kanon popularnej muzyki, kształtując gusta masowej publiczności. Przykładowo, format radiowy determinował długość i strukturę piosenek – promował utwory ok. 3-minutowe o przewidywalnej budowie zwrotka-refren. Za to listy przebojów wpływały na to, że nowe utwory dostosowywały się do określonych trendów, co tworzyło zwrotną pętlę (Rothenbuhler, 2003).

Podobnie jak wcześniejsze instytucje, również te medialne nieustannie przystosowują się do zmian w środowisku. Przykładowo, radio musiało ustąpić częściowo pola telewizji, a wytwórnie musiały dostosować się do epoki streamingu. W końcu XX i na początku XXI wieku obserwujemy ponadto narodziny nowych, mniej sformalizowanych form organizacji muzycznych. Są to społeczności online, sieci wirtualnych wytwórni czy

niezależne sieciowe kolektywy artystyczne (np. *DYM recordings*, *Pointless Geometry*, *Kołorking Muzyczny*) – co świadczy o ciągłym procesie instytucjonalnej ewolucji. Te różnorodne formy pełnią podobne funkcje: integrują działania wielu jednostek w spójną całość, z której wyłania się nadrzędny cel. Jest nim realizacja określonego programu lub misji, która odgórnie i hierarchicznie ukształtowanie zachowania pojedynczych uczestników.

3.2.4.3. MCA, ANT, VSM

Wcześniejszy rozdział w skrótovej formule ukazał rolę i funkcje muzycznych instytucji jako sprawczych bytów w historii muzyki Zachodu od czasu wynalezienia zapisu muzycznego. Wspominałem już o trzech naukowych teoriach, które mają wyjaśnić ową sprawczość – zgodnie z główną tezą tej pracy – nabytą na drodze memetycznej ewolucji. Wspominałem wcześniej, że instytucję definiuję jako wieloskalową architekturę kompetencji (MCA) (Levin, 2023) złożoną z agencjalnych części (ANT), zorganizowaną przez cybernetyczną strukturę kontrolną (VSM).

Wieloskalowa architektura kompetencji to teoria opisuje ontologię oraz filogenezę sprawczości instytucji. MCA Michaela Levina opisuje uniwersalny w biologii system organizacji, w którym mniejsze struktury są częścią większych. Co najważniejsze, takie wielopoziomowe zagnieżdżenie struktur odnosi się także do ich funkcjonowania, sprawczości oraz poznania. Wszystkie warstwy, rozpoczynając od pojedynczej komórki, mają własne cele. Osiągają je dzięki temu, że są w stanie efektywnie pobierać informacje z zewnątrz (postrzegać) oraz skutecznie je przetwarzać w sposób fundamentalnie inteligentny, gdzie inteligencja definiowana jest jako zdolność do osiągnięcia jednego celu różnymi metodami (Levin, 2024). Implikuje to również zagnieżdżoną hierarchię sprawczości. Każdy poziom tej struktury – od pojedynczych komórek, przez tkanki i organy, aż po organizm jako całość – funkcjonuje jako kolektywny podmiot o określonym stopniu sprawczej autonomii i inteligencji.

Sprawczość na danym poziomie nie sprowadza się do pasywnego przetwarzania danych. To aktywna zdolność do rozwiązywania problemów w celu osiągnięcia konkretnych stanów docelowych (Levin, 2022). Komórki inteligentnie dążą do homeostazy, podczas gdy tkanki i organy, jako agenci wyższego rzędu dążą do zachowania lub odtworzenia złożonej anatomii. Należy podkreślić istnienie mechanizmu kooptacji celów: sprawczość niższego rzędu jest zaprzęgnięta do realizacji celów wyższego rzędu.

Oddolną sprawczość i inteligencję pojedynczej komórki Levin wielokrotnie wykazuje w swoich empirycznych badaniach, gdzie nawet jedna komórka żaby lub ludzkiego płuca w nowym kontekście cechuje się zaskakującym i skomplikowanym zachowaniem, jak chociażby poruszanie się i tworzenie wcześniej nieznanymi struktur (Blackiston, 2021; Gumuskaya 2023) Na tej oddolnej sprawczości komórki, która to na kolejnych poziomach organizacji łączy się w coraz bardziej złożone struktury opiera się Levinowska koncepcja agencjalnych materiałów Darwina (Levin, 2023). Przeczy ona tradycyjnemu postrzeganiu materii biologicznej jako pasywnych klocków, budulec, który musi zostać precyzyjnie ukształtowany przez genom. Zamiast tego Levin postuluje, że ewolucja wytwarza, selekcionuje i operuje na materiałach (komórkach, tkankach), które same w sobie są kompetentnymi, rozwiązującymi problemy agentami. To właśnie dlatego, że budulec jest z natury sprawczy i inteligentny, cały system nie wymaga centralnego sterowania.

Zatem organizm nie mikrozarządza (Gładych, 2017) działaniem każdej komórki. Zamiast tego, poprzez m.in. komunikację bioelektryczną, wyznacza nadrzędny cel, na przykład „zbuduj tutaj kończynę”, a kolektywna inteligencja komórek autonomicznie znajduje sposób na jego realizację, jednocześnie osiągając własne cele. Tym samym każda warstwa organizacji życia posiada własne kompetencje do działania w dynamicznym i nieprzewidywalnym świecie. Levin dąży w ten sposób do stworzenia jednolitego systemu aktywnego, inteligentnego i skalowalnego przetwarzania informacji w strukturach biologicznych, a także kolektywach. Często powtarza, że jego zdaniem każda inteligencja to inteligencja zbiorowa i sama ta koncepcja może służyć jako narzędzie do integracji biologii na różnych poziomach organizacji, od cząsteczek, poprzez komórki, tkanki, organy organizmy aż po roje oraz społeczeństwa (Levin, 2024). Poszukiwanie odpowiednio uniwersalnego opisu inteligencji dla wszystkich skal biologii doprowadza go właśnie do uwzględniania w swoich modelach nawet struktur społecznych, co wyjaśniał w wywiadzie dla Europejskiego Instytutu Nauki w Zarządzaniu:

Każda teoria inteligencji musi być w stanie pomóc nam bezpośrednio zmierzyć oraz wejść w interakcję z czymś, co obecnie nazywamy różnymi inteligencjami. Nie interesują mnie teorie, które stosują się wyłącznie do ssaków, lub ssaków i ptaków, ani te, które zawodzą, gdy musimy pomyśleć o ośmiornicy albo insektach (...). Chcę takiej koncepcji inteligencji, która opisz wszystkie możliwe agentów. Oznacza to nie tylko członków filogenetycznego drzewa na Ziemi, ale także nowatorskie konstrukty syntetycznej biologii, wszystkie sztuczne inteligencje, jakie zbudujemy, zarówno te w formie programów komputerowych, jak i te ucieleśnione, potencjalnych agentów egzobiologicznych, dziwaczne byty w przeróżnych skalach, w tym pojedyncze komórki, wewnątrzkomórkowe ścieżki sygnałowe, ogromne rzeczy, takie jak struktury społeczne i w gruncie rzeczy może nawet proces ewolucyjny sam w sobie.

(...) any theory of intelligence (...) has to be able to help us directly compare and interact with what we now call diverse intelligences. So I'm not interested in theories that apply just to mammals or to mammals and birds or that break down when we have to think about octopus or insects(...). I want a framework for intelligence that is going to handle all possible agents. This means not only the things we see in the phylogenetic tree here on earth, but novel synthetic biology constructs, artificial intelligences that we may build, either in hardware or software, potential exobiological agents, weird beings at other scales, including individual cells, subcellular molecular networks, enormous things like social structures, and basically, perhaps the evolutionary process itself (Levin, 2022)

Cytat ten stanowi doskonały pomost do analizy instytucji. Choć Levin nie nazywa tego w ten sposób, włączanie struktur społecznych w model MCA jest głęboko spójne z wizją memetycznej ewolucji kultury, co wyczytać też można z innych jego artykułów (Levin, 2024b; Levin, 2024c). Skoro model MCA dostarcza nam biologicznie ugruntowanej ontologii dla zagnieżdżonych, agencjalnych bytów, a sam Levin wskazuje, że jego teoria powinna obejmować struktury społeczne, pojawia się dodatkowe, praktyczne pytanie: w jaki sposób taka wieloskalowa struktura na poziomie grup społecznych utrzymuje spójność i zdolność do przetrwania w czasie? Jakie uniwersalne prawa rządzą jej organizacją?

Odpowiedzi na te pytania dostarcza cybernetyka. Już w połowie XX wieku, na długo przed sformułowaniem MCA, teoretycy systemów zmagali się z problemem złożoności i przetrwania w organizacjach. Przekonujące rozwiązanie tego problemu stworzył brytyjski badacz Stafford Beer, gdy opracował Model Systemu Wiabilnego (Viable System Model – VSM). Jest to schemat organizacyjny opisujący uniwersalne i

konieczne warunki, które każdy system (biologiczny, społeczny czy technologiczny) musi spełnić, aby być „wiabilnym”, czyli zdolnym do autonomicznego przetrwania w zmiennym otoczeniu (Beer, 1981).

Warunki te obejmują zdolność do samoprodukcji, utrzymania równowagi wewnętrznej oraz do reakcji adaptacyjnych na zmiany środowiska. Istotną cechą systemów Beera jest ich modułarna, hierarchiczna budowa. Badacz zauważył, że każdy system zdolny do przetrwania (wiabilny) ma strukturę rekursywną: składa się z podsystemów, które same w sobie są wiabilne i zorganizowane na podobieństwo całości (Beer, 1984). Innymi słowy, instytucja zawiera w sobie mniejsze instytucje – np. uniwersytet może zawierać wydziały, katedry i laboratoria. Wszystkie one działają względnie autonomicznie. Realizują lokalne cele, ale jednocześnie są funkcjonalnie włączone w nadrzędny cel całości. Taka struktura jest całkowicie tożsama z Levinowskim MCA. Levin używa przykładu, że kiedy występuje uszkodzenie naskórka, komórka skóry realizuje nadrzędny cel organizmu, regenerację, jednocześnie realizując swój zwyczajny oddolny cel utrzymywania połączenia z sąsiednią komórką.

Jednak jeśli MCA Levina opisuje ontologię i filogenezę sprawczości w organizacji społecznej, to VSM Beera dostarcza plan jej cybernetycznego szkieletu: struktury kontrolnej i komunikacyjnej, bez której każda złożona organizacja uległaby rozpadowi. Oba modele, choć wywodzą się z różnych dziedzin, są w zadziwiający sposób spójne i posłużą mi do wspólnego opisanie działania instytucji muzycznych jako mempleksów spajających aktywność poszczególnych ludzi. Beer argumentował, że aby jakikolwiek system, taki jak biologiczny organizm, firma, czy państwo, był w stanie zachować autonomię i adaptować się do zmian w otoczeniu, musi posiadać specyficzną, rekursywną strukturę złożoną z pięciu wzajemnie połączonych podsystemów (Beer, 1981). VSM opisuje zatem wewnętrzną architekturę sprawczości; dostarcza swoistej cybernetycznej neurofizjologii bytu instytucjonalnego.

Model ten precyzuje niezbędne i wystarczające warunki dla autonomicznego istnienia systemu. Opisuje, w jaki sposób organizacja przetwarza informacje, zarządza

złożonością, utrzymuje tożsamość i koordynuje swoje działania operacyjne w celu przetrwania (Beer, 1981). W tej analizie VSM posłuży do zmapowania funkcjonalnego odpowiednika mózgu i układu nerwowego instytucji muzycznych. W modelu VSM wyróżnia się pięć podstawowych funkcji (podsystemów) obecnych na każdym poziomie organizacji. I tak: System 1 to jednostki wykonawcze, np. pojedyncze zespoły, orkiestry, wydziały, które realizują podstawowe operacje. System 2 zapewnia koordynację i komunikację między nimi. System 3 odpowiada za integrację wewnętrzną i bieżące zarządzanie co w biologii odpowiada utrzymaniu homeostazy. System 4 skierowany jest na otoczenie – zbiera informacje o zmianach warunków zewnętrznych, planuje adaptacje i strategię. Zaś System 5 pełni rolę polityki i tożsamości – ustala ogólne cele, misję i spaja całość, zapewnia równowagę między działaniem „tu i teraz” a długofalową perspektywą (Beer, 1981).

Dzięki takiej architekturze instytucja staje się wiabilna: reaguje na bodźce zewnętrzne (przez System 4), utrzymuje spójność wewnętrzną (System 3 i 2), realizuje bieżące funkcje (System 1) i nie ztraca przy tym swojej tożsamości i celu (System 5). Model VSM przyznaje każdej warstwie organizacji pewien zakres decyzyjności oraz sprawczości. Podstawowe jednostki (System 1) posiadają swobodę działania w swoim obszarze – podobnie jak komórki w organizmie wykonują lokalne zadania. Wyższe poziomy nie ingerują w każdy detal, lecz wyznaczają ramy i granice. Podobnie jak w MCA, jest to efektywnie unikanie mikrozarządzania. Reagują one dopiero wtedy, gdy działanie niższych jednostek wymaga korekty w imię dobra całości. Taki mechanizm sprzężeń zwrotnych realizuje się poprzez różnego rodzaju kontrole i audyty.

Ta właśnie wielopoziomowa organizacja umożliwia adaptacyjność: system może korygować swoje działanie na poziomie lokalnym bez konieczności centralnej interwencji za każdym razem, a jednocześnie może centralnie zmienić politykę, gdy wymaga tego przetrwanie w zmiennym otoczeniu. Co ciekawe, Stafford po wydaniu książki *The Brain of The Firm* był krytykowany za to, że jego system opisujący organizmy i organizacje opierał się na analogii. W artykule *The Viable System Model* tłumaczy, że było odwrotnie – jego cybernetyczne opisy systemów wiabilnych

wynikają z modeli czysto matematycznych, ale są tak uniwersalne, że pasują zarówno do opisu organizmu jak i organizacji (Beer, 1984).

Za przykład wiabilnego systemu może posłużyć filharmonia. Aby spełniała warunki wiabilności, jej poszczególne zespoły i działy, takie jak orkiestra, chór czy administracja muszą potrafić samodzielnie rozwiązywać swoje problemy. Oznacza to, że umieją przygotować koncert, rozwiązać konflikt personalny lub opracować ofertę dla szkół. W ten sposób zyskują elastyczność i szybkość działania. Rola centralnego kierownictwa polega zaś na zharmonizowaniu tych działań z misją całości - linią programową, budżetem - i reagowaniu na zmiany zewnętrzne (np. nowe preferencje publiczności, kryzys finansowy czy sanitarny).

Ten podstawowy model można wzbogacić i doprecyzować. Konstruowany w tym rozdziale opis instytucji muzycznych z wieloskalową architekturą kompetencji (MCA) oraz modelem systemu wiabilnego (VSM) uzupełnia i domyka jeszcze opisywana wcześniej teoria aktora-sieci (ANT). Dzięki niej do ewolucyjnie wykształconych wiabilnych struktur społecznych możemy włączać także nie-ludzkich agentów, takich jak fortepian, koło kwintowe, radio czy algorytmy rekomendacyjne Spotify.

Głębia i adekwatność prezentowanego wcześniej podejścia meta-teoretycznego najlepiej ukażą się w przykładowym modelu. Jako że wcześniej opisywałem już instytucje z kręgu muzyki artystycznej, teraz przykład będzie mocniej zakorzeniony w kulturze muzyki popularnej. Przedmiotem analizy będzie typowa, duża XX-wieczna wytwórnia płytowa. Ponieważ nawet największe wytwórnie mają różnorodne struktury (Schmuck, 2012), a każda z nich podlega nieustannej ewolucji, na podstawie dostępnej literatury proponuję pewien uogólniony model, który pokazuje, w jaki sposób może być skonstruowana wytwórnia, aby spełniać ramy VSM i jak przekłada się to na realia ANT. Opieram się przede wszystkim na modelach integracji pionowej wytwórni zaproponowanych przez Petersona i Bergera (1975), a także Mola, Wijnberga i Carolla (Mol, 2005).

Proponowana typowa duża wytwórnia jest w pełni sprawczym bytem cybernetycznym, którego głównym celem stało się nie tyle tworzenie muzycznej kultury, ile efektywne zarządzanie nią w celu maksymalizacji zysku. Zastosowanie Modelu Systemu Viabilnego pozwala na zrozumienie jej wewnętrznej logiki jako organizacji dążącej do przetrwania na konkurencyjnym rynku. Strukturę typowej dużej wytwórni z XX wieku mapujemy na pięć podsystemów Beera następująco:

W systemie 1 funkcjonuje kilka konkurujących ze sobą działów operacyjnych. Każdy z nich stanowi także rekursywny VSM. Na początek wytwórnia musi odnaleźć utalentowanych artystów, którzy zechcą podpisać kontrakt. Tym zajmuje się dział Artists and Repertoire czyli A&R. Żeby artyści mogli zarejestrować swoją muzykę, potrzebne jest także studio nagraniowe wraz z wszystkimi jego pracownikami (Pastushkov, 2024). Ten realny produkt utrwalany jest w fizyczne nośniki poprzez dział produkcji fizycznej. Tutaj zachodzi masowa replikacja. Dalej, dział marketingu i promocji odpowiada za promowanie produktu w mediach. Skomponowany, zarejestrowany, wytłoczony i rozpromowany materiał musi jeszcze skutecznie trafić do słuchaczy i tutaj włącza się dział dystrybucji (Pastushkov, 2024).

System 2 odpowiada za koordynację. Tłumi konflikty i nieuniknione oscylacje między jednostkami Systemu 1, które z natury mają sprzeczne interesy. Nie da się uniknąć tego, że dział A&R chce inwestować w ryzykownych artystów, a marketing woli promować pewne hity. To fundamentalny dylemat eksploracji i eksploatacji (March, 1991). Z jednej strony system dąży do sprawnego działania tu i teraz, a z drugiej zabiega o przyszły byt. Rozwikłanie tego napięcia dzięki systemowi 2 wytwarza nowe twórcze strategie. W praktyce ten system 2 realizuje się przez mechanizmy koordynacyjne, takie jak centralne harmonogramy wydawnicze, alokacja budżetów na poszczególne projekty, a także komitety do spraw repertuaru, które synchronizują działania różnych działów.

System 3 prowadzi zarządzanie i audyt. Tu znajduje się „mózg operacyjny” wytwórni. Jego działanie napędza kierownictwo średniego i wyższego szczebla. Zadaniem

systemu 3 jest zarządzanie bieżącą działalnością poprzez alokację zasobów. Między systemami 1 i 3 trwają negocjacje, aby osiągnąć cele i zachować środki (budżet, czas studyjny, wsparcie promocyjne). Ponadto System 3 monitoruje wydajność operacji za pomocą określonych wskaźników, takich jak dane sprzedaży czy czas odtwarzania piosenek w radiu. To nie wszystko. System trzeci posiada też swój własny kanał audytu, System 3*. Tenże pozwala na bezpośrednie, wyrywkowe sprawdzanie stanu operacji, omijając oficjalne kanały sprawozdawcze. Realizuje go na przykład prezes wytwórni osobiście odsłuchujący dema lub odwiedzający sesję nagraniową (Davis, 2013). W ten sposób prezes uzyskuje surowe, nieprzetworzone dane, które mogłyby umknąć w oficjalnych raportach. Pozwala to systemowi 3 lepiej zrozumieć prawdziwą złożoność operacji systemu 1 i skuteczniej nią zarządzać.

Na system 4 składa się strategia i adaptacja. To „oczy i uszy” instytucji, odpowiedzialne za patrzenie na zewnątrz i planowanie przyszłości. W opisywanej wytwórni rolę tę pełnią działy badań rynku, analitycy trendów, a także pracownicy A&R w ich funkcji prognostycznej. System 4 skanuje otoczenie – zmieniające się gusta publiczności, działania konkurencji, pojawienie się nowych gatunków muzycznych czy technologii (np. kasety magnetofonowa, MTV). Dzięki temu modeluje przyszłe scenariusze i rekomenduje strategiczne adaptacje, które zapewniają firmie długoterminową żywotność.

Najszerszy system 5 obejmuje tożsamość i politykę. To najwyższa instancja decyzyjna: zarząd, właściciele, prezes. System 5 jest odpowiedzialny za zdefiniowanie ostatecznej polityki, misji i tożsamości organizacji. Co najważniejsze, utrzymuje równowagę między wewnętrzną optymalizacją „tu i teraz” (żądania Systemu 3) a zewnętrzną, złożoną adaptacją do przyszłości (rekomendacje Systemu 4). Rozwiązuje także fundamentalne konflikty tych systemów. Tutaj także zachodzi dylemat eksploracji i eksploatacji, ale na szerszym poziomie.

Jak widać powyżej, model VSM dostarcza eleganckiego, funkcjonalnego szkieletu organizacji. Dopełnia go Teoria Aktora-Sieci, która wyjaśnia, jak sprawczość wytwórni jest realnie konstruowana i negocjowana nie tylko przez sprawczość człowieka. Z perspektywy ANT wytwórnia nie jest monolitycznym bytem, lecz dynamicznym sojuszem pomiędzy ogromną liczbą ludzkich i nie-ludzkich aktantów. Sprawczość wytwórni nie rezyduje wyłącznie w gabinecie prezesa (System 5).

Oczywiście główną sprawczą rolę w wytwórni pełnią ludzie. Pełnią oni rozmaite funkcje. Do głosu dochodzą artyści, producenci, inżynierowie dźwięku, menedżerowie, prawnicy, pracownicy marketingu, dziennikarze muzyczni, DJe radiowi, sprzedawcy w sklepach płytowych, a także słuchacze. Wszyscy ci ludzie operują jednak w specyficznym środowisku materialnym, które także ma niebagatelne znaczenie. Z tego środowiska na pierwszym miejscu wyróżniają się dokumenty i inskrypcje: kontrakty płytowe (kluczowy dokument definiujący relacje i zobowiązania), partytury, teksty piosenek czy recenzje w prasie. Ogromne znaczenie mają także listy przebojów (np. Billboard Hot 100), które działają jak potężne „urządzenia inskrypcyjne” (inscription devices) (Latour, 1988), obiektywizujące i kwantyfikujące sukces.

Niezbędne są także technologie: mikrofony, konsolety mikserskie, magnetofony wielośladowe, maszyny do tłoczenia winylu, radioodbiorniki. Każda z tych technologii kształtuje możliwości i ograniczenia procesu twórczego i dystrybucyjnego. Kategorię pokrewną, ale szerszą stanowią obiekty materialne: Taśma-matka, płyta winylowa w okładce, singiel promocyjny wysyłany do stacji radiowych. Ponad ludźmi i obiektami stoją prawa i standardy: Prawo autorskie (fundamentalna zasada umożliwiająca handel muzyką jako własnością intelektualną), standardy techniczne (np. prędkość obrotu płyty 33 1/3 RPM, częstotliwość próbkowania), niepisane reguły gatunkowe. Wytwórnia nigdy nie działa w próżni, zawsze współpracuje z innymi instytucjami, takimi jak: stacje radiowe, sieci telewizyjne (później MTV), sklepy płytowe, organizacje przyznające nagrody (np. Grammy).

Siła wytwórni polega na jej zdolności do stworzenia dynamicznych, usieciowionych relacji pomiędzy tymi wszystkimi aktantami. W języku ANT nazywa się to przeprowadzeniem skutecznego procesu translacji – „przetłumaczeniem” interesów wszystkich tych aktantów na własny język i wciągnięcia ich do swojej sieci jako sojuszników. Wytwórnia „tłumaczy” artystyczne ambicje muzyka na precyzyjne zapisy w kontrakcie; potrzebę stacji radiowej do posiadania chwytliwych hitów na wyprodukowanie trzyminutowego singla; pragnienie nastolatka do identyfikacji z idolem na okładkę płyty i kampanię wizerunkową. Utrzymanie tych wszystkich, często sprzecznych, translacji w stanie względnej stabilności jest źródłem władzy i sprawczości wytwórni.

Do takiej usieciowionej wizji wytwórni muzycznej idealnie pasuje krytyka przemysłu kulturowego sformułowana przez Theodora Adorna i Maxa Horkheimera (1947). Adorno opisywał proces, w którym kultura masowa staje się zbiorem standaryzowanych, przewidywalnych produktów, tworzonych w celu pacyfikacji mas i maksymalizacji zysku. Pseudindywidualizacja to tworzenie powierzchownych różnic między pozornie odmiennymi produktami kulturowymi. Wytwórnia może na przykład promować artystów, którzy różnią się ubiorem i stylem tańca, a ich muzyka jest już bardzo podobna. W języku ANT taka pseudindywidualizacja to proces, w którym różnorodni aktanci (artyści) są zmuszani przez system do odgrywania tej samej, wymiennej roli w sieci, aby zapewnić jej stabilność i przewidywalność.

Poprzez swoją systemową sprawczość sieć wytwórni może narzucać estetyczny schemat, który minimalizuje ryzyko rynkowe. Połączenie perspektyw VSM, ANT i krytyki Adorna pozwala sformułować wniosek, że XX-wieczna wytwórnia płytowa funkcjonowała jako wysoce zaawansowany, cybernetyczny system doboru sztucznego. Zastąpiła ona w dużej mierze mechanizmy doboru naturalnego panujące na wolnym rynku wśród mniejszych graczy. Zamiast pasywnie obserwować, które piosenki odnoszą sukces, wytwórnia aktywnie interweniowała w proces selekcji. Jej System 4 (Strategia) analizował istniejące hity, aby wyekstrahować z nich powtarzalną formułę. Następnie System 3 (Zarządzanie), wykorzystywał całą sieć podległych mu aktantów (kontrakty,

budżety, promocję radiową), żeby narzucić tę formułę nowym produktom. W ten sposób widać, że duże wytwórnie nie odkrywały hitów. Raczej systematycznie je konstruowały – *hodowały* muzykę pod kątem specyficznej cechy: maksymalnej rentowności w ramach istniejącej sieci dystrybucji. Czyli można mówić tu o sztucznym doborze. Jest to ten podobny mechanizm ewolucyjny do tego, który wcześniej zidentyfikowałem w kontekście muzyki akademickiej. Jednak celem optymalizacji nie był tu prestiż, lecz zysk. Sugeruje to uniwersalność logiki doboru sztucznego jako strategii działania złożonych bytów instytucjonalnych. Sztuczna selekcja wymaga dążenia do określonego celu, a w zgodzie z Levinowskim MCA oraz memetyką Dennetta cele zaczynają się pojawiać właśnie dopiero po przejściu pewnego poziomu złożoności.

3.2.5. Systemy Rekomendacji Muzyki

W latach 90. XX wieku globalna dystrybucja treści kultury podlegała kontroli mass mediów. Również sprzedaż i rozpowszechnianie muzyki zależały od największych światowych przedsiębiorstw medialnych, które kształtowały rynek oraz trendy według własnej wizji (Bagdikian, 2004). Niekwestionowany monopol ogromnych wytwórni i koncernów medialnych wydawał się nie do przełamania, zgodnie z logiką końca historii (Fukuyama, 1992), spójną zresztą z ich generalnym przekazem. Tymczasem obecnie zamiast końca historii obserwujemy raczej medialny kryzys wielkich narracji (Lyotard, 1984) – atomizację punktów widzenia i początek mikro-trendów związany z erą smartfonów i mediów społecznościowych (Jenkins, 2006). Zamiast celowo, odgórnie kształtowanej narracji znanej z mass mediów, dominują niezliczone spersonalizowane narracje sterowane częściowo przez użytkowników nowych mediów, a częściowo przez algorytmy, a raczej systemy złożone z algorytmów, które zbiorczo nazywamy systemami rekomendacji (recommender system, recommender, rekomender).

W pierwszych dekadach XXI wieku dominacja mass mediów załamuje się w obliczu głębokiej transformacji sposobu dystrybucji treści kultury. Podstawowym warunkiem tej zmiany jest rewolucja technologiczna: rozpowszechnienie smartfonów z dostępem do internetu. Nastąpiła masowa produkcja przenośnych urządzeń, które jeszcze dekadę wcześniej nazywane byłyby superkomputerami. Już w pierwszej dekadzie XXI wieku stały się one niezmiernie popularne również dzięki temu, że były dostępne w przystępnych cenach (Zittrain, 2009). Te miliony urządzeń komunikacyjnych i multimedialnych nowej generacji stanowiły doskonałe środowisko do implementacji nowego rodzaju oprogramowania. W drugiej dekadzie XXI wieku ujawniły się już nowe aplikacje, które zdominowały swoje rynki, najlepiej odpowiadając na potrzeby konsumentów. Wśród nich są m.in: YouTube, Spotify, Netflix, Facebook, Instagram, TikTok, Uber, Amazon. Te aplikacje mają różnorodne zastosowania, ale łączy je jedna rzecz – wszystkie opierają się na architekturze systemów rekomendacyjnych. Rekomendery zdominowały każdy sektor cyfrowej kultury (Van Dijck, 2018).

To zjawisko ma także ogromny wpływ na świat muzyki. Tradycyjne – dominujące jeszcze w XX wieku – metody poznawania przez słuchaczy nowej muzyki, takie jak radio, sklepy z płytami, rekomendacje osób znajomych czy muzykologów – błyskawicznie tracą na znaczeniu i zanikają. Obecnie większość nowej muzyki poznajemy poprzez rekomendacje w mediach społecznościowych i na platformach streamingowych (Hagen, 2015).

Owe rekomendacje są owocem działania zaawansowanych systemów informatycznych opartych na sztucznej inteligencji, a szczególnie dynamicznie rozwijającego się uczenia maszynowego (Schedl, 2018). Takie systemy odnajdziemy na wszystkich znaczących muzycznych platformach streamingowych, takich jak Spotify, Apple Music, YouTube Music czy TikTok (który przede wszystkim postrzegany jest jako audiowizualne medium społecznościowe, ale w dalszej sekcji rozdziału pokażę, że ma też cechy muzycznej platformy streamingowej). Rekomendery stały się integralną częścią codziennych nawyków słuchania muzyki na całym świecie. Fundamentalnie zmieniły one sposób, w jaki użytkownicy i prosumenci (osoby zarazem produkujące i konsumujące) kształtują ogólny ekosystem muzyczny (Morris, 2015).

Niniejszy rozdział demonstruje, że systemy rekomendacji muzycznej nie są jedynie pasywnymi narzędziami do filtrowania treści. Przeciwnie, rozumiem je jako sprawcze memetyczne silniki, czyli systemy aktywnie pobudzające i kształtujące kulturę muzyczną. Ich sprawczość manifestuje się poprzez prowokację aktywnych i dynamicznych zmian muzyczno-memetycznych presji selekcyjnych oraz drastyczną zmianę środowiska, w którym funkcjonuje muzyczna pula memowa. W praktyce oznacza to wpływ na:

- globalne gusta i preferencje muzyczne,
- trendy w kompozycji i produkcji muzycznej,
- ścieżki odkrywania nowej muzyki przez słuchaczy,
- funkcjonowanie rynku muzycznego, w szczególności wydawania, promocji i dystrybucji muzyki (Morris, 2015).

Nowe zjawisko algorytmicznego odkrywania muzyki stanowi nie tylko zmianę technologiczną – to także fundamentalne przejście kulturowej władzy symbolicznej w rozumieniu Bourdieu (Bourdieu, 1993), a proces ten opisuję poniżej. Przy bliższej analizie możemy zidentyfikować nową, wielowymiarową arenę walki o władzę symboliczną i relacje klasowe. Ta walka toczy się na co najmniej dwóch komplementarnych polach: pomiędzy słuchaczami oraz pomiędzy słuchaczem a platformą.

W pierwszej kolejności pokrótce opiszę koncepcję walki o dominację w sferze technologicznej pomiędzy słuchaczami. Co ciekawe, zgodnie z analizą André Janssona (Jansson, 2021) pole walki stanowi już sama obsługa cyfrowych technologii. Codzienna, praktyczna ekspertyza w zarządzaniu strumieniami danych i domową infrastrukturą (głośnikami, aplikacjami) staje się narzędziem dominacji. Osoby biegłe technicznie mogą w codziennych interakcjach nieświadomie stosować symboliczną przemoc wobec tych, którzy nie radzą sobie z obsługą nowych mediów, utrwalając w ten sposób społeczne hierarchie. To znaczy, że osoba, która lepiej obsługuje sprzęt, posiada pewien rodzaj władzy.

Poza tym inna koncepcja wyznacza jeszcze drugie pole walki. Konkurencja o dominację w prestiżowej sferze smaku czy też gustu odbywa się także między człowiekiem a samą platformą. Opisuje to Jack Webster (Webster, 2019), który pokazuje, że firmy takie jak Spotify zmieniają sposoby zaznaczania pozycji klasowej przy pomocy gustu muzycznego. Z jednej strony, algorytmy i łatwy dostęp do ogromnych katalogów muzyki osłabiają tradycyjne źródła dystynkcji klasy średniej, takie jak specjalistyczna wiedza muzyczna czy zdolność do eksperckiego doceniania muzyki. Jeszcze niedawno fachowy komentarz czy posiadanie przepastnej domowej kolekcji robiło wrażenie. W erze streamingu wszyscy odbiorcy mają dostęp do tych samych przepastnych katalogów, których nikt nie jest w stanie poznać w całości. Ekspert, tak samo jak nowicjusz, ma tylko wiedzę na temat niewielkiego ułamka repertuaru dostępnego na wyciągnięcie ręki.

A jednak te same platformy otwierają także nowe możliwości do demonstrowania kapitału kulturowego (Webster, 2019). Klasy średnie adaptują się chociażby poprzez tworzenie wyrafinowanych playlist, które sygnalizują ich unikalny gust i kompetencje kuratorskie bezpośrednio na platformie (Webster, 2019). To nie wszystko. Również sama konsumpcja muzyki na nośnikach fizycznych, takich jak płyty winylowe, a nawet płyty CD czy kasety, stała się swoistą formą wychodzenia ponad nowo ustalony standard, czyli słuchanie ze streamingu. To forma oporu. W ten sposób nowe pokolenie słuchaczy może demonstrować swoje koneserstwo, odróżniając się od zwyczajnego, przeciętnego, masowego, cyfrowego odbiorcy (Webster, 2019). Choć pozornie streaming i algorytmy demokratyzują dostęp do kultury, w rzeczywistości tworzą nowe, bardziej subtelne mechanizmy reprodukcji społecznych nierówności. Dominacja przejawia się już nie tylko w tym, co się wie o muzyce, ale również w tym, jak sprawnie obsługuje się technologię oraz jak potrafi się zaprezentować swój gust w nowym, cyfrowym ekosystemie. A poza nowymi relacjami władzy pomiędzy ludźmi w świetle streamingu muzycznego, powstaje również nowy rodzaj walki o sprawczość między człowiekiem a systemem rekomendacji muzycznej.

Dotąd proponowanie, rekomendowanie nowej muzyki słuchaczom było domeną ludzkich aktorów – krytyków, DJ-ów, muzykologów (także organizacji: mass mediów i wytwórni muzycznych) – którzy pełnili funkcje kuratorskie i kontekstualizowali muzykę (Negus, 1999). Systemy rekomendacyjne – napędzane przez uczenie maszynowe oraz szerzej rozumianą sztuczną inteligencję – w dużej mierze zautomatyzowały tenże proces kuratorski (Gillespie, 2014). Owa automatyzacja sprowadza się do tego, że architektura wyboru muzyki jest projektowana głównie przez częściowo autonomiczne algorytmy. Ostatecznie to one - nieprzejrzyste czarne skrzynki - decydują, z jaką muzyką codziennie stykają się miliony użytkowników (Pasquale, 2015). Platformy streamingowe coraz mocniej stawiają na sugestie algorytmiczne, umieszczając je w centrum swoich modeli biznesowych (Brown, 2016). W konsekwencji w rosnącym stopniu to one kreują to, które utwory są w danej chwili popularne, i do których w ogóle w praktyce jako słuchacze możemy dotrzeć.

Jednocześnie trzeba mocno podkreślić, że wewnętrzne działanie tych systemów najczęściej nie jest transparentne ani dla ich użytkowników, ani nawet dla ich twórców. Reasumując: preferencje i nawyki większości słuchaczy muzyki na całym świecie kontrolowane są przez nieprzejrzyste, częściowo autonomiczne systemy oparte na sztucznej inteligencji.

W uproszczeniu może się wydawać, że systemy rekomendacyjne po prostu filtrują treści jak sito. Ale taki obraz pomija ich dynamiczną sprawczość w tej materii. Konstruuje one aktywnie dla użytkowników złożone nisze muzyczne. Ta algorytmiczna konstrukcja niszy nie tylko kształtuje indywidualne nawyki słuchania muzyki, ale także steruje wspólną ewolucyjną trajektorią gatunków muzycznych i preferencji słuchaczy w skali globalnej. Przytaczana w poprzedniej części niniejszej pracy biologiczna i memetyczna teoria konstrukcji niszy zakłada, że organizmy (a przez analogię, odbiorcy muzyki) aktywnie kształtują swoje środowisko, co z kolei wpływa na ich byt oraz ewolucję (Odling-Smee, 2003; Laland, 2000). Algorytmy rekomendacji muzycznej, poprzez kształtowanie repertuaru, tworzą technologiczną niszę memetyczną (Acerbi, 2019), środowisko, w którym memy muzyczne, takie jak style czy gatunki muzyczne replikują się i ewoluują. Personalizując strumienie, czy też całe grupy i ciągi pobieranych treści, systemy te aktywnie definiują granice i zawartość muzycznego środowiska użytkownika, często tworząc muzyczne bańki informacyjne (Ricci, 2022). Poprzez zmiany chwilowych nawyków słuchaczy, algorytmy wpływają też na rozwój ich długoterminowych preferencji oraz na to, jakie muzyczne treści zyskują popularność i rezonans kulturowy. Tworzy to pętlę sprzężenia zwrotnego, w której algorytmy kształtują gust, a dane generowane na podstawie tych ukształtowanych gustów dalej udoskonalają algorytmy, wpływając na tworzenie nowych stabilnych systemów w ewolucji kulturowej muzyki (Acerbi, 2019).

Systemy rekomendacyjne to zbiory algorytmów, których celem jest zwiększenie ilości konsumowanych przez użytkownika treści kulturowych, takich jak muzyka (Ricci, 2022). Sam ten fakt należy uwypuklić: użytkownicy są zadowoleni, ponieważ poznają więcej piosenek, które trafiają w ich gusta. Ale celem muzycznej korporacji udostępniającej

dany rekomender nie jest zadowolony słuchacz, tylko zysk, a ów zysk jest generowany dzięki temu, że słuchacz zwiększa swoją konsumpcję muzyki. Kierunkiem optymalizacji takich systemów i naczelną wartością rekomendacji treści jest to, aby słuchacz nie mógł się od nich oderwać. Mówiąc szerzej, systemy rekomendacji to programy, które przetwarzają informacje o użytkowniku i treściach (muzyce, filmach, książkach, zdjęciach, wiadomościach), a następnie generują spersonalizowane listy tych treści w celu maksymalizacji zaangażowania użytkownika (Schedl, 2018). Z perspektywy użytkownika, efektem działania takiego programu jest regularne otrzymywanie rekomendacji – sugestii – czego posłuchać, obejrzeć lub przeczytać dalej, często z zaskakującą trafnością (Aggarwal, 2016).

Za to kluczowy cel platformy to po prostu zwiększenie zaangażowania użytkowników na różne sposoby. Oznacza to na przykład wydłużanie czasu spędzonego na platformie, retencję, czyli stałe zaangażowanie, oraz, ostatecznie, generowanie przychodów poprzez subskrypcje lub ekspozycję na reklamy (Ricci, 2022). Czyli klienci mają konsumować więcej, częściej, bardziej regularnie i jak najwięcej za to płacić. Systemy te doskonale odnajdują się w obecnej erze przytłaczającej ilości danych, ponieważ pomagają użytkownikom radzić sobie z nadmiarem informacji i wyborów. Rekomender prowadzi użytkownika do takich treści, które rzeczywiście go interesują (Ricci, 2022).

Już w 1971 laureat nagrody Nobla w dziedzinie ekonomii, Herbert Simon, zauważył, że ludzie otoczeni są nieustannie rosnącą ilością informacji i ma to również swoje negatywne konsekwencje. Jak stwierdził Simon, nadmiar informacji prowadzi do deficytu uwagi. (Simon, 1971). Później, w 1997 Michael Goldhaber swój głośny esej dotyczący przemian w internecie rozpoczął słowami „Uwaga klienci! Walutą Nowej Ekonomii nie będą pieniądze, ale uwaga” („Attention Shoppers! The currency of the New Economy won't be money, but attention”) (Goldhaber, 1997). Koncepcja ekonomii uwagi wyjaśnia sens istnienia systemów rekomendacji. W świecie przesyconym informacją uwaga użytkownika staje się zasobem rzadkim i cennym, a platformy zjadają o nią konkurują. Rekomendery dostarczają spersonalizowane treści, aby utrzymać uwagę użytkownika tak długo, jak to możliwe. Warto jeszcze raz podkreślić, że optymalizują one dobór treści nie tylko w celu wydłużania poszczególnych,

jednorazowych sesji konsumpcji, ale usiłują kształtować długotrwałe, stabilne nawyki tejże konsumpcji. Jest to możliwe poprzez wmanipulowanie nowego użytkownika w aproksymację nawyków innego, lojalnego, zaangażowanego użytkownika. Stanowi to kluczową strategię dla platform dążących do osiągnięcia i utrzymania dominacji na rynku. A z ewolucyjnych praw rynku wynika, że platforma, która nie dąży do osiągnięcia dominacji nad konkurencją, po prostu tę konkurencję przegrywa.

W 2025 termin sztuczna inteligencja kojarzy się głównie z generatywną AI. Ale zanim sztuczna inteligencja zaczęła popularnie generować treści, jej głównym zastosowaniem była właśnie optymalizacja dystrybucji treści na potrzeby ekonomii uwagi. Umożliwiło to gigantom technologicznym, takim jak Spotify, YouTube czy Netflix, zdominowanie odpowiednich nisz kulturowych poprzez efektywną dystrybucję treści wiralnych – płodnych memów – i wzmacnianie zaangażowania.

Ekosystem rekomendacji przeważnie obejmuje co najmniej trzy główne grupy aktorów czy też interesariuszy: Konsumentów (użytkowników/słuchaczy), Dostawców (artystów, wytwórców, kompozytorów) oraz Właścicieli Systemów (platformy takie jak Spotify, YouTube) (Abdollahpouri, 2017). Warto od razu zwrócić uwagę, że taki trójpodział najlepiej opisuje platformy streamingowe takie jak Spotify, Tidal, czy Netflix, które służą do konsumpcji profesjonalnych produktów. Za to platformy nastawione bardziej na aspekt społecznościowy łączą ze sobą role producentów i konsumentów w hybrydową postać prosumenta. Tego typu dynamikę odnajdziemy na YouTube, TikToku, Instagramie, Facebooku. Użytkownik sam wybiera rolę, w której się znajdzie – producent, prosument, konsument. Czasami możemy podać taką informację na samym początku aktywności podczas rejestracji konta, ale system i tak profiluje każdego poprzez jego zachowanie. Zatem podsumowując, role producenta i konsumenta mogą łączyć się w osobie jednego użytkownika, ale nadal są to osobne role, kluczowe dla działania systemu. Odrębną od nich rolę pełni właściciel systemu.

Te trzy grupy aktorów najczęściej posiadają rozbieżne interesy. Producenci treści dążą do szerokiej ekspozycji i sprawiedliwego wynagrodzenia, konsumenci poszukują przyjemnych i różnorodnych treści, a właściciele platformy dążą do rentowności i

udziału w rynku (Abdollahpouri i in., 2017). Chociaż idealnie system rekomendacji powinien optymalizować proces dla wszystkich stron, platforma (właściciel systemu) ostatecznie kontroluje projekt algorytmiczny i może stawiać na pierwszym miejscu własne cele. Może to objawiać się np. promowaniem własnych treści lub optymalizacją algorytmów w celu skutecznego gromadzenia danych, które są sprzedawane i / lub napędzają dalszy rozwój AI (Abdollahpouri i in., 2017). Przykładem może być znacząca inwestycja Spotify w rozwój AI, co implikuje motywację do optymalizacji operacji w celu pozyskiwania większej ilości użytkowników, ponieważ dane te są kluczowe dla postępów w uczeniu maszynowym (Ricci i in., 2022).

Systemy rekomendacji można również opisać w kategoriach memetycznych, co pozwala głębiej zrozumieć ich rolę w muzycznym ekosystemie. Ich rolą – którą spełniają z ogromną skutecznością – jest zwiększenie ilości czasu, który użytkownicy spędzają na słuchaniu muzyki. Innymi słowy ich celem jest jak największa liczba muzycznych 'streamów' czyli de facto pobrań treści utworów muzycznych na urządzenia cyfrowe, czyli replikacja tych utworów. Rekomendery napędzają zatem replikację utworów muzycznych. Możemy określić je jako abstrakcyjne maszyny replikacji, wirtualne maszyny replikacji, silniki replikacji. Przeanalizujemy ogólny mechanizm tej replikacji. Podstawowym założeniem systemów rekomendacyjnych 'content based' (opartych na zawartości) jest to, że jeżeli użytkownikowi spodobała się zawartość (piosenka) A, to spodoba mu się również zawartość (piosenka) B, która jest do niej podobna. To znaczy, że podobna treść wywoła podobną reakcję (tak wprost sformułowane założenie trudno znaleźć w literaturze technicznej, ale można je z niej bezpośrednio wywieść i obserwować jego konsekwencje). System rejestruje pozytywną reakcję użytkownika na daną piosenkę, a następnie stara się zreplikować jego pozytywne przeżycie poprzez dostarczenie zreplikowanego, ale nie identycznego bodźca. Widać zatem, że zachodzi tutaj dziedziczenie z mutacją, zachodzi również presja selekcyjna, ponieważ użytkownik wybiera tylko niektóre rekomendacje – spełnione są zatem warunki działania algorytmu ewolucyjnego. Tym, co bezpośrednio podlega ewolucji, nie jest jednak ani reakcja, ani sam końcowy bodziec muzyczny, choć są one bezpośrednio powiązane z naszym głównym przedmiotem. Ten właśnie główny ewoluujący przedmiot stanowi wewnątrzsystemowa wektorowa reprezentacja informacji na temat tego, co spodoba się

danemu użytkownikowi w danym momencie. Chodzi tu o wektory w semantycznej przestrzeni wektorowej word2vec lub cyfrową reprezentację gustu użytkownika w systemach filtrowania kolaboracyjnego, które omawiam szerzej w kolejnej sekcji rozdziału.

Projektowanie systemów rekomendacyjnych w celu optymalnego funkcjonowania w warunkach ekonomii uwagi sprowadza się do tego, że pośród innych działań, algorytmy nieuchronnie faworyzują cechy muzyczne, które maksymalizują natychmiastowe zaangażowanie. Platformy streamingowe oceniają przeciętny czas zaangażowania użytkownika z materiałem audio-wideo na 6-15 sekund (Ganesh et al., 2024) i kształtuje to dynamkę rekomenderów. Ta algorytmiczna preferencja tworzy środowisko, w którym artyści mogą być (świadomie lub nieświadomie) zachęceni do produkcji muzyki o tych cechach, aby zyskać widoczność i przychylność algorytmiczną – faworyzowanie i promocję. Stwarza to presję selekcyjną, która przekształca produkcję muzyczną, potencjalnie marginalizując złożoność i ponadprzeciętnie angażujące doświadczenia słuchowe, które wymagają trwałej, a nie natychmiastowej uwagi (Singh et al., 2023).

Model wielostronny systemów rekomendacji ujawnia nieodłączną asymetrię władzy. Platforma (właściciel systemu) ostatecznie kontroluje algorytmiczne "reguły gry" i może dostrajać te reguły tak, aby służyły jej strategicznym interesom (np. rentowności, pozyskiwaniu danych, promocji treści ekskluzywnych), które nie zawsze muszą być zbieżne z optymalnymi wynikami dla użytkowników lub wszystkich dostawców treści. Subtelne lub jawne kierowanie rekomendacjami w celu faworyzowania wyników korzystnych dla własnego modelu biznesowego może obejmować wiele mechanizmów. System może przyznawać pierwszeństwo treściom o tańszych licencjach, zachęcać użytkowników do większego udostępniania własnych danych czy też nawet własnych oryginalnych treści. W ten sposób rekomender kształtuje krajobraz kulturowy daleko ponad proste dopasowywanie treści do zastanych preferencji użytkowników (Abdollahpouri i in., 2017).

Dalej przedstawiam główne paradygmaty algorytmiczne. Zrozumienie systemów rekomendacji muzycznej wymaga zapoznania się z fundamentalnymi technikami algorytmicznymi, które stanowią ich trzon. Poniżej omawiam podstawowe typy rekomenderów.

Jednym z dwóch głównych typów podstawowych algorytmów rekomendacji jest Filtrowanie Kolaboracyjne (collaborative filtering, CF). Podstawową ideą filtrowania kolaboracyjnego jest wykorzystanie wzorców interakcji użytkownik-przedmiot (np. ocen, historii odsłuchań) w celu przewidywania przyszłych preferencji (Ricci i in., 2022). CF opiera się na założeniu, że użytkownicy, którzy zgadzali się w przeszłości co do ocen pewnych przedmiotów, będą się w tej materii zgadzać również w przyszłości.

Innymi słowy, jeśli:

użytkownikowi A podobają się utwory 1, 2, 3, 4,

użytkownikowi B podobają się utwory 7, 8, 9, 10,

użytkownikowi C podobają się utwory 2, 3, 4, 5,

użytkownikowi D podobają się utwory 8, 9, 10, 11,

podstawową zasadą działania CF będzie zauważyć podobieństwo użytkowników A i C oraz B i D. Następnie CF identyfikuje, że każdemu użytkownikowi warto polecić to, co spodobało się osobie z nim sparowanej, a czego sam jeszcze nie poznał. Zatem: użytkownikowi A poleci utwór 5, użytkownikowi B poleci utwór 11, użytkownikowi C poleci utwór 1, użytkownikowi D poleci utwór 7. Ta sama prosta zasada daje zaskakująco trafne rezultaty, kiedy użytkowników w puli jest nie czterech, a cztery miliony, a listy odsłuchań obejmują setki utworów. Operacje matematyczne nabierają wtedy złożoności, którą można opracować na różne sposoby.

Dokładniej, w ramach filtrowania kolaboracyjnego wyróżnia się dwa główne podejścia. Pierwsze to metody oparte na sąsiedztwie. System może mierzyć sąsiedztwo

użytkowników lub przedmiotów. Sąsiadujący użytkownicy to osoby o podobnych gustach do użytkownika docelowego. System poleca danemu użytkownikowi przedmioty, które spodobały się jego sąsiadom (Su, 2009). Za to sąsiadujące przedmioty to obiekty powiązane wzorcami ocen użytkowników. Jeśli użytkownik polubił dany przedmiot, system podpowiada mu inne, które były wysoko oceniane przez tych samych odbiorców (Desrosiers, 2011).

Drugie podejście opiera się na modelach statystycznych. Najczęściej wykorzystuje się tu faktoryzację macierzy, w tym metodę dekompozycji wartości osobliwych (SVD). System uczy się wtedy ukrytych czynników opisujących zarówno użytkowników, jak i przedmioty, na podstawie dostępnych ocen (Koren, 2009). Czynniki te nazywa się latentnymi, ponieważ nie są bezpośrednio obserwowalne, lecz wywnioskowane ze wzorców danych (Goodfellow, 2016). Dzięki temu modele te dobrze radzą sobie z problemem rzadkich danych i pozwalają odkrywać struktury preferencji, które nie są widoczne na pierwszy rzut oka.

Główną zaletą CF jest możliwość znajdowania nowatorskich i nieoczywistych, zaskakujących rekomendacji – przedmiotów, których użytkownik mógłby nie odkryć na podstawie samych cech treści. To znaczy, z niezrozumiałych lub egzotycznych przyczyn np. wielu słuchaczom muzyki metalowej może podobać się jedna konkretna piosenka country. Wyszukanie jej ręczne przez kogoś, kto ogólnie nie lubi muzyki country, jest bardzo mało prawdopodobne, ale CF może ją zidentyfikować i polecić. Dodatkową zaletą CF jest to, że działa bez jawnych informacji na temat treści przedmiotów. To znaczy, jeżeli nie mamy żadnych jawnych cech utworów muzycznych, informacji o ich gatunku, wykonawcy, czy roku produkcji, ale wiemy, komu się one podobają, to wystarczy, aby trafnie je polecić.

CF towarzyszy za to tak zwany problem zimnego startu, czyli trudność w poleceniu znanych treści nowym użytkownikom lub rekomendacji znanym użytkownikom nowych, nieocenywanych jeszcze przedmiotów. To znaczy – jeżeli o danej treści lub użytkowniku system wie niewiele, nie będzie potrafił dokonać trafnej rekomendacji. Towarzyszy to szczególnie początkowemu wprowadzaniu treści lub użytkownika na platformę.

Jednocześnie, jako inna strona tego samego zjawiska, występuje również skłonność do faworyzowania popularnych przedmiotów, znana jako błąd popularności (popularity bias).

Drugą podstawową metodą rekomendacji jest filtrowanie oparte na treści (Content-based Filtering, CBF). CBF rekomenduje przedmioty na podstawie ich podobieństwa do przedmiotów, które użytkownik polubił w przeszłości. To podobieństwo określa się przez analizę treści lub cech przedmiotów. Przykładowo zatem, jeżeli użytkowniczka Rozalia polubiła piosenki o atrybutach A, B, C i D, a dostępna pula piosenek obejmuje:

piosenkę 1 o atrybutach A, D,

piosenkę 2 o atrybutach Y, Z,

piosenkę 3 o atrybutach B, C, D,

piosenkę 4 o atrybutach X, O,

system zidentyfikuje piosenki 1 oraz 3 jako posiadające pożądane atrybuty i piosenki 2 oraz 4 jako nieposiadające pożądanych atrybutów. Użytkowniczce Rozalii zarekomenduje zatem piosenki 1 oraz 3.

W przypadku muzyki owe atrybuty to metadane takie jak gatunek, artysta, album, data wydania, a także cechy piosenki wyekstrahowane z samego sygnału audio (np. tempo, rytm, harmonika, barwa, taneczność, walencja) oraz treści semantycznych (teksty piosenek, recenzje, tagi). Profil preferencji użytkownika jest budowany na podstawie cech przedmiotów, z którymi pozytywnie wszedł on w interakcję. Nowe przedmioty są następnie porównywane z tym profilem.

Wśród zalet CBF możemy wymienić to, że potrafi ono rekomendować nowe i niepopularne przedmioty, o ile tylko dostępne są informacje o ich cechach. Zapewnia przejrzystość rekomendacji – można wyjaśnić, dlaczego dany przedmiot jest rekomendowany na podstawie jego cech. Mniejsze znaczenie ma tu problem zimnego startu użytkownika, ponieważ przeważnie początkowe preferencje można jawnie

określić lub wywnioskować z kilku przykładów. To wystarczy, aby stworzyć początkowe propozycje treści.

System ten posiada również wady. Jego jakość i skuteczność określa jakość wprowadzonych do systemu cech przedmiotów. Mała liczba tych cech lub niejasne ich określenie – na przykład przez automatyczny system do ekstrakcji cech – skutkuje niską trafnością rekomendacji. Ponadto działanie CBF może prowadzić do nadmiernej specjalizacji, czy też eksplorowania zamkniętej niszy. Uzyskamy to poprzez rekomendowanie przedmiotów zbyt podobnych do tego, co użytkownik już zna, zarazem ograniczając nowość i autentyczne odkrycia.

Reasumując, żeby rekomendacje oparte na filtrowaniu kolaboracyjnym zadziałały, potrzebujemy tylko wiedzy o tym, który użytkownik polubił który przedmiot. Musimy mieć więc pewną bazę użytkowników i ich doświadczeń. Aby za to zadziałały rekomendacje oparte na filtrowaniu zawartości, potrzebujemy wiedzy na temat przedmiotów, takiej jak ich cechy i metadane. Nie potrzeba za to wiedzy o bazie użytkowników.

Na wyszczególnienie zasługują tzw. podejścia hybrydowe. Współcześnie system rekomendacji w praktyce najczęściej opiera się na synergii wielu poszczególnych algorytmów rekomendacji. Takie podejścia hybrydowe łączą kilka różnych technik rekomendacyjnych (np. CF i CBF), aby wykorzystać ich zalety i zniwelować wady, co często prowadzi do poprawy wydajności (Burke, 2002). Przy problemie zimnego startu w CF pomagają informacje o atrybutach przedmiotów i przejście na rekomendacje CBF. Za to CF poprawia dokładność systemu i wychwytuje podobieństwa między użytkownikami, których sama treść mogłaby nie ujawnić.

Stworzenie synergii algorytmów rekomendacyjnych to zadanie wysoce nietrywialne. Można je rozwiązać na wiele sposobów, otrzymując rozbieżne wyniki. W literaturze (Burke, 2002) takie strategie posiadają własne nazwy:

Ważona (Weighted): Oceny z różnych systemów rekomendacyjnych są łączone liniowo.

Mieszana (Mixed): Rekomendacje z różnych komponentów są prezentowane razem.

Przełączająca (Switching): Wybiera jeden system rekomendacyjny w zależności od bieżącego kontekstu lub profilu użytkownika.

Kombinacja Cech (Feature Combination): Cechy z jednego źródła rekomendacji są wykorzystywane jako cechy wejściowe dla innego typu systemu rekomendacyjnego.

Augmentacja Cech (Feature Augmentation): Jeden system rekomendacyjny generuje nowe cechy opisowe dla przedmiotów, które są następnie wykorzystywane przez inny system.

Kaskadowa (Cascade): Podejście hierarchiczne, w którym drugorzędny system rekomendacyjny udoskonala wyniki systemu pierwszorzędnego, często używane do rozstrzygnięcia remisów (tzn. sytuacji kiedy różne algorytmy naraz określają swój stopień pewności co do trafności rekomendacji tą samą miarą liczbową) lub dopracowywania listy.

Metapoziomowa (Meta-level): Model nauczony przez jeden system rekomendacyjny służy jako dane wejściowe dla innego.

Ilość konkurencyjnych rozwiązań obrazuje złożoność problemu synergii rekomenderów. Stosowanie podejść hybrydowych jest obecnie dominującą praktyką, ale same sposoby hybrydyzacji jednak stanowią dodatkowy, złożony problem. Wybór konkretnej hybrydyzacji jest już mocno zależny od dziedziny, w której operuje dany system rekomendacji (np. audio, wideo, produkt, tekst). Ponadto, poza liniowymi i nieliniowymi hybrydyzacjami istnieją również inne popularne modyfikacje podstawowych algorytmów rekomendacyjnych w nowoczesnych systemach – w tym zaawansowane rozwiązania oparte na sztucznej inteligencji i sieciach neuronowych.

Wcześniejsze przykłady działania algorytmów CF i CBF obejmowały bardzo niewielkie zbiory danych, z których łatwo i niezawodnie można wyciągnąć wnioski. W realnych warunkach takie systemy przetwarzają dane o milionach użytkowników i milionach przedmiotów, co łączy się w niebywale złożone macierze.

Kiedy opisane wyżej systemy przestają skutecznie działać, programiści muszą rozwiązać problem złożoności danych. W tym celu coraz częściej wykorzystują oni modele głębokiego uczenia do ulepszania wszystkich typów podejść rekomendacyjnych. Te modele wychwytyją złożone, nieliniowe wzorce w wyjątkowo dużych i zróżnicowanych zbiorach informacji (Ricci, 2022).

Wśród podejść opartych na sieciach neuronowych odnaleźć można m.in. następujące techniki:

Głębokie Sieci Neuronowe (DNN) do ogólnego rozpoznawania wzorców.

Konwolucyjne Sieci Neuronowe (CNN) do przetwarzania danych o strukturze siatkowej, takich jak spektrogramy audio (do ekstrakcji cech muzycznych) lub obrazy (okładki albumów) (Choi, 2017).

Rekurencyjne Sieci Neuronowe (RNN), w tym LSTM i GRU, do modelowania danych sekwencyjnych, takich jak historie odsłuchań użytkowników lub playlisty (Ricci i in., 2022).

Wariacyjne Autoenkodery (VAE) do uczenia się ukrytych reprezentacji użytkowników i przedmiotów (Zhang, 2019).

Grafowe Sieci Neuronowe (GNN) do modelowania złożonych relacji w grafach interakcji użytkownik-przedmiot lub grafach wiedzy (Ricci i in., 2022).

Mechanizmy Uwagi (Attention Mechanisms) do skupiania się na ważnych częściach danych wejściowych (Zhang, 2019).

Wśród kluczowych korzyści systemów wzbogaconych o sztuczne sieci neuronowe wymienić możemy zdolność do automatycznej ekstrakcji cech z surowych danych – zarazem redukcję ręcznej inżynierii cech. To zatem zachodząca na szeroką skalę automatyzacja pracy, która inaczej musiałaby być wykonywana przez ludzi-inżynierów. Poza tym modelowanie wysoce nieliniowych preferencji użytkowników oraz efektywna integracja różnorodnych typów danych – audio, tekst (słowny), interakcje użytkowników, kontekst – to nie tylko zadanie żmudne i czasochłonne dla ludzkiego programisty. Przede wszystkim ani sam człowiek, ani klasyczne oprogramowanie (bez technik uczenia maszynowego i AI) nie posiada zdolności wykrywania i zapamiętywania subtelnych zależności w przepastnych wzorach danych. To nie tylko żmudne, ale także najprawdopodobniej niemożliwe (być może teoretycznie możliwe w przyszłości na podstawie kompletnie nieznanych jeszcze, nieodkrytych metod).

Nie mogę przeprowadzić całkowitej, dogłębnej analizy architektur algorytmicznych systemów rekomendacji z największych komercyjnych platform, ponieważ nie są one jawne. Korporacje strzegą tajemnic swoich rekomenderów podobnie jak McDonald's czy KFC swoich sekretnych składników. Poza tym największe muzyczne platformy, takie jak Spotify, YouTube (w tym YouTube Music) oraz TikTok podlegają ciągłej ewolucji wraz ze swoimi rekomenderami. Zatem wszelkie szczątkowe informacje na ich temat szybko stają się nieaktualne. Uważam jednak, że nawet przybliżony opis ma dużą wartość. Na szczęście istnieją pewne publicznie dostępne publikacje, blogi techniczne oraz inne analizy, które umożliwiają rekonstrukcję ogólnych zasad działania tych rekomenderów. Opis ich architektury stanowi podstawę do wyjaśnienia mechanizmów dystrybucji współczesnej kultury muzycznej. To właśnie te systemy w znacznej mierze określają to, jaką muzykę odkrywają słuchacze na całym świecie. I jakiej słuchają. Jednocześnie takie rekomendery to "czarne skrzynki" – wyników ich działania czasami nie da się dokładnie opisać, są nieprzejrzyste. Ale to, co da się opisać, staram się przedstawić poniżej. Omawiam Spotify, YouTube z YouTube Music oraz TikTok.

Spotify wykorzystuje zaawansowany, hybrydowy system rekomendacji. Stanowi on fundament realizacji celów biznesowych marki. Firma zabiega o utrzymanie użytkowników i maksymalizację czasu spędzanego przez nich na platformie, czyli ostatecznie po prostu generowanie przychodów (Maasø, 2019). W Spotify modele oparte na głębokim uczeniu przetwarzają ogromne ilości danych. Tak powstaje szeroko zakrojona personalizacja, która jest podstawą sukcesu aplikacji.

Spotify wykorzystuje złożony system rekomendacji oparty na wielu typach danych. Obejmują one interakcje użytkowników, metadane, surowe pliki audio oraz informacje tekstowe w języku naturalnym. Interakcje użytkowników to zarówno sygnały jawne (np. ulubione utwory, zapisane do biblioteki), jak i niejawne — historia odsłuchań, pominięcia, wskaźniki ukończenia utworów, czas słuchania, tworzenie i edytowanie playlist oraz aktywność w wyszukiwarkach i obserwowanie artystów.

Na metadane składają się tytuły utworów, nazwy albumów, etykiety gatunkowe, daty wydania i dane dotyczące artystów. Za to analiza audio wyprowadza z plików dźwiękowych dwa rodzaje istotnych cech. Cechy niskiego poziomu, czyli elementarne, obejmują na przykład spektrogramy. Czyli z danym utworem powiązany jest przebieg widma akustycznego. Poza tym analiza wyodrębnia także cechy wyższego poziomu, czyli zaawansowane wnioski na temat struktury muzycznej. Wśród autorskiej grupy parametrów Spotify wyróżnia takie kategorie jak taneczność, tempo, tonacja, głośność, barwa, rytm, harmonika i walencja emocjonalna. Za to najbardziej abstrakcyjną warstwę danych stanowią informacje językowe. Teksty piosenek, recenzje, czy tytuły playlist dodatkowo pogłębiają modelowany kontekst kulturowy (Maheshwari, 2023; Hernandez, 2021; Beats to rap on, 2025).

Personalizacja w Spotify wynika z synergii wielu różnych algorytmów. Nawet samo klasyczne filtrowanie kolaboracyjne działa tu w obu wariantach, które wcześniej omawiałem. Rekomender niezależnie ocenia sąsiedztwa użytkowników oraz piosenek (Maheshwari, 2023). Spotify wprowadził na tym polu skuteczną innowację: zaczął uwzględniać współwystępowanie utworów w milionach playlist i sesji odsłuchowych.

Pomaga to wychwycić kontekstowe powiązania między piosenkami i lepiej odwzorowuje różnorodne nawyki słuchania (Maheshwari, 2023).

Wyniki systemu poprawia dodatkowo filtrowanie oparte na treści. W tym przypadku korzysta ono nie tylko z cech akustycznych, ale także informacji semantycznych pozyskiwanych dzięki metodom z dziedziny przetwarzania języka naturalnego. Tak wzbogacone CBF tworzy reprezentacje utworów i dopasowuje je do profili użytkowników, co szczególnie sprawdza się w sytuacjach zimnego startu (Damak, 2019). Ponadto, w wielu obszarach systemu zintegrowane jest głębokie uczenie. Klasyczne sieci głębokie (DNN) wychwytyją nieliniowe zależności między użytkownikami i utworami. Poza tym, sieci konwolucyjne (CNN) analizują spektrogramy, a sieci rekurencyjne (RNN) modelują sekwencje odsłuchań (Maheshwari, 2023; Liu, 2024). A zatem trzy różne typy głęboko uczonych sieci mają swoje konkretne zadania. To wszystko wspiera faktoryzacja macierzy, która wzbogaca reprezentacje zarówno utworów, jak i użytkowników (Damak, 2019). W badaniach Spotify pojawia się również podejście, które ma równoważyć propozycje utworów dobrze znanych z nowymi, które zaskakują słuchacza (McInerney et al., 2018).

Spotify regularnie tworzy dla swoich słuchaczy kilka rodzajów spersonalizowanych playlist. Najbardziej znana z nich to Discover Weekly, która co tydzień poleca użytkownikowi nowe utwory na podstawie jego upodobań. Według firmy od 2015 użytkownicy wysłuchali dzięki Discover Weekly już ponad sto miliardów utworów (Spotify, 2025). Kolejna to Release Radar, czyli przegląd nowości. Uaktualnia się w każdy piątek i wyróżnia się tym, że zawiera premierowe wydawnictwa od obserwowanych przez użytkownika artystów (Spotify, b.d.). Daily Mixes to codziennie odświeżane zestawy utworów, które mieszają nowe rekomendacje z tym, co znane i lubiane. Ponadto warto wspomnieć o interesującym narzędziu Discovery Mode. To funkcja nie dla słuchaczy, ale artystów i wytwórni. Autorzy mogą wybrać utwory, które otrzymają pierwszeństwo w rekomendacji w zamian za dodatkowe 30% prowizji. To zręczne narzędzie, dzięki któremu zdesperowani początkujący artyści zarabiają jeszcze mniej, ale mogą liczyć na szersze zasięgi (Spotify for artists, b.d.).

Za to system rekomendacji YouTube w znacznej mierze opiera się na uczeniu głębokim (Covington, 2016). Oczywiście YouTube nie jest platformą wyłącznie muzyczną, ale znaczna część materiałów wideo wprost stanowi teledysk czy też klip audio ze statycznym obrazkiem. Jeszcze inną grupę stanowią materiały wideo jak vlogi, streamy czy inne gatunki, którym towarzyszy stały akompaniament muzyczny. Poza tym wiele z najpopularniejszych materiałów na platformie to właśnie utwory muzyczne. W 2025 lista 10 najpopularniejszych materiałów na YouTube zawiera same piosenki (Gupta, 2025). Wśród nich prym wiedzie dziecięca przyśpiewka Baby Shark z ponad 15 miliardami odsłon, a stawkę zamyka muzyczny mem internetowy Gangnam Style przy ponad 5 miliardach wyświetleń (Newman, 2025). Istnieje również dedykowana do odtwarzania muzyki aplikacja YouTube music, która współdzieli z główną platformą pewne aspekty infrastruktury, co może wpływać na dane używane do trenowania systemu rekomendacji. Oficjalny przewodnik podkreśla, że YouTube Music przenosi preferencje i rekomendacje z głównego YouTube (Lee Harter, 2024).

Ów rekomender wykorzystuje trzy główne grupy sygnałów: interakcje użytkownika, cechy treści oraz dane kontekstowe. Praktycznie wszystko, co użytkownik robi na platformie jest brane pod uwagę. Liczą się wszystkie kliknięcia, każdorazowy czas oglądania materiału, zostawione polubienia, niepolubienia, udostępnienia linków, historia wyszukiwania, czy też pomijanie treści. Do tego dochodzą też odpowiedzi w ankietach dotyczące satysfakcji konsumenta (Goodrow, 2021; Macready, 2025). Jeżeli chodzi o same materiały, liczą się ich tytuły, opisy, metadane, transkrypcje audio-wideo, a w przypadku muzyki także jej zidentyfikowane cechy takie jak gatunek, rytm czy tonacja (Macready, 2025). Trzecia grupa sygnałów to dane kontekstowe. Obejmują one porę dnia, typ urządzenia oraz lokalizację użytkownika. Algorytm działa codziennie na setkach miliardów sygnałów — Goodrow (2021) podaje, że ponad 80 miliardów sygnałów wpływa na rekomendacje każdego dnia. Czyli jeżeli ktoś chwyta telefon, żeby posłuchać ulubionej piosenki, do systemu trafiają następujące sygnały: ile razy i gdzie stuknął w ekran, co dokładnie wpisał, wraz z błędami, który materiał i jak szybko wybrał, ile czasu słuchał przed wyłączeniem, gdzie się wtedy znajdował, która była godzina,

oraz z jakiego urządzenia korzystał. Charakterystyczną cechą rekomendera YouTube jest jego dwustopniowa architektura opisana przez Covingtona, Adamsa i Sargina (2016). Najpierw sieć neuronowa generuje zbiór wstępnych propozycji, czyli kandydatów. W praktyce są to setki filmów wybrane z ogromnej biblioteki. Wybór kandydatów opiera się m.in. na historii oglądania użytkownika, wyszukiwanych hasła, wskaźnikach demograficznych i interakcjach innych, podobnych użytkowników (Covington, 2016).

W drugiej fazie osobny model rankingowy ocenia tych kandydatów. Przewiduje czas ich oglądania oraz inne wskaźniki zaangażowania, takie jak przeszłe zachowania względem tego kanału czy aktualność podejmowanej w materiałach tematyki (Covington, 2016). Personalizacja odbywa się dynamicznie: system porównuje wzorce oglądania danego użytkownika z wzorcami innych użytkowników, uwzględniając sygnały na bieżąco. Wagę poszczególnych sygnałów ustala indywidualnie (Covington et al., 2016). Czyli jeśli na przykład ktoś często komentuje materiały, to fakt, że dodał gdzieś kolejny komentarz, będzie mało istotny. Podobnie odwrotnie: jeśli osoba mało wylewna nagle się otworzy, to będzie ważny sygnał dla systemu.

Własną specyfikę posiada też rekomender YouTube Music. Serwis, tropem Spotify, oferuje trzy rodzaje spersonalizowanych miksów: Discover Mix, New Release Mix i Your Mix. Aktualizują się cyklicznie na podstawie historii odsłuchań (Lasche, 2019). W 2020 platforma wprowadziła nowość w postaci playlist konstruowanych do konkretnych nastrojów czy też czynności (activity mix) (Lasche, 2020). Tutaj jednym kliknięciem otrzymamy ciąg utworów dobrany do relaksu, ćwiczeń na siłowni czy też dojazdu do pracy. Wartą uwagi innowacją jest możliwość budowania własnych stacji radiowych przez funkcję *Your music tuner*. Użytkownik według własnych preferencji steruje różnymi parametrami. Dobiera różnorodność artystów, ilość prezentowanych nowości oraz ustala parametry utworów takie jak tempo, popularność i nastrój. Słuchacz może także użyć polecenia *Ask Music*, które wygeneruje osobistą stację radiową na podstawie opisu sformułowanego w języku naturalnym (YouTube Music Help, 2025).

Inaczej funkcjonuje TikTok. Rekomender TikToka bazuje na uczeniu maszynowym. W odróżnieniu od Spotify i YouTube music, Tiktok to aplikacja do scrollowania. Ta różnica w interfejsie użytkownika wiąże się też z inną naturą działania systemu rekomendacji. Tutaj personalizuje on przede wszystkim treści widoczne na stronie głównej, nazwanej 'Dla Ciebie' (*For You Page, FYP*) (Zhou, 2024; Roesner, 2024). Już sama ta nazwa może podpowiadać użytkownikowi, że otrzymuje produkt szyty na miarę. System zaprojektowany jest, żeby unikać błędu popularności i wprost nie faworyzuje popularnych twórców. Preferuje za to treści pasujące do interakcji użytkownika, a także bada cechy samego materiału (Sheikh, 2024; Pollock, 2025).

TikTok zbiera różnorodne dane dotyczące interakcji podejmowanych przez użytkownika. Mierzy liczby polubień, udostępnień i komentarzy. Sprawdza też sposób oglądania danego materiału wideo: to, czy użytkownik obejrzał go do końca, czy oglądał wielokrotnie i jak długo to oglądanie trwało. Na tej podstawie ocenia poziom zainteresowania użytkownika daną treścią (Zhou, 2024; Sheikh, 2024). W drugiej kolejności system analizuje informacje o filmie. Liczą się zawarte w nim dźwięki, hashtagi, napisy i efekty wizualne. Tak rekomender nakreśla kontekst kulturowy i kategorię tematyczną materiału (Sheikh, 2024; Zhang, 2021). Trzeba powiedzieć też, że system wykorzystuje osobiste dane dotyczące właściciela profilu. Uwzględnia on ustawienia urządzenia i konta, takie jak preferencje językowe, kraj przebywania, czy indywidualne zainteresowania (Sheikh, 2024; Lu, 2024).

TikTok stanowi medium społecznościowe, ale jedną z najważniejszych ról w serwisie odgrywa muzyka i dźwięk. Tiktokerzy chętnie wykorzystują wyróżnioną w interfejsie funkcję 'użyj dźwięku'. Dzięki niej prosument może w dowolny sposób wykorzystać całość lub część fragment dźwięku, który napotka. Są to profesjonalne piosenki, dialogi lub oryginalne nagrania. Funkcja 'użyj dźwięku' pozwala dodać swoje wideo do znalezionej audio, które samo w sobie też można remiksować (np. zmieniać głośność,

dodawać dialogi). Jeszcze inną funkcją jest 'duet'. Tutaj prosument może wykorzystać odnaleziony materiał dźwiękowy wraz z wideo i nałożyć na nie własne nagranie audiowizualne. Zastany i autorski materiał audio brzmią wtedy jednocześnie, a materiały wideo mogą się przenikać lub sąsiadować na ekranie. Trzecia funkcja to 'stitch'. Dzięki niej do znalezionej materiału użytkownik może dograć własne zakończenie.

W artykule *Audio memes, earworms and templatibility: the 'aural turn' of memes on TikTok* (Abidin, 2021) autorzy wskazują, że na TikToku materiały audio to coś znacznie więcej niż tło do wideo. Często warstwa dźwiękowa zawiera centralny zwrot akcji w danym materiale. Zdarza się, że słowa wykorzystanej piosenki zawierają pointę. Również zmiana rytmu, melodii czy harmoniki czasami kształtuje narrację wideo w sposób, który zmienia jego wydźwięk. Nagła muzyczna zmiana może wprowadzać ironię (Abidin, 2021).

Poza tym system aktywnie promuje filmy wykorzystujące chwilowo popularne ścieżki dźwiękowe, zwłaszcza te o wysokich wskaźnikach zaangażowania, ukończenia i częstym powtarzalnym użyciu (Abidin, 2021). Istotną częścią TikToka są ewoluujące ciągi zmienności wywodzące się od konkretnej treści, takie jak remiksy, lip-sync, tańce, wyzwania. To konkretne gatunki TikTkowej twórczości. Remiks to ogólna kategoria, lip-sync oznacza synchronizację ruchu ust do gotowego audio, tańce to własne taneczne wideo do popularnej piosenki. Wyzwania to jedna z bardziej kontrowersyjnych gałęzi TikToka, gdzie uczestnicy dokumentują wykonanie często ryzykownych zachowań. Jednym z elementów ciągłości takiego łańcucha wyzwań może być wykorzystanie konkretnego materiału audio. Uczestnictwo użytkowników w tego typu łańcuchach z elementem audio napędza wiralną replikację dźwiękowych memów na TikToku (Ling. C., 2021).

Ta ogromna sieć społecznościowa ma znaczący wpływ na świat muzyki, choć reprezentuje odrębną ścieżkę odkrywania muzyki niż Spotify czy YouTube Music. TikTok oddziela odkrywanie muzyki od tradycyjnych modeli skoncentrowanych na

danych o artyście lub albumie i wiąże je bezpośrednio z trendami krótkich form wideo oraz uczestnictwem w treściach generowanych przez użytkowników.

Spotify i YouTube Music wciąż w dużej mierze działają w ramach rekomendowania piosenek, artystów, albumów lub playlist. Za to na TikToku popularne materiały audio napędzają masowe uczestnictwo użytkowników w różnego rodzaju remiksach i wyzwaniach, co z kolei napędza dalszą algorytmiczną promocję tych dźwięków i powiązanych filmów. Tworzy to szybką, memetyczną propagację fragmentów muzycznych, gdzie sama piosenka może stać się drugorzędna wobec jej funkcji w złożonym wiralnym trendzie. Jest to inny tryb popularyzacji muzyki niż w tradycyjnym muzycznym streamingu. Ale jest to też radykalnie odmienny tryb konsumpcji. Użytkownicy TikToka to typowi prosumenci. W odróżnieniu od Spotify, gdzie muzyk oraz słuchacz uczestniczą w platformie na całkowicie innych zasadach, na TikToku praktycznie wszyscy odbiorcy to również producenci. Ważniejsze od skonsumowania zamkniętego, profesjonalnego produktu jest współuczestnictwo, wspólna zabawa, współtworzenie. Esencjonalnie to właśnie ten rodzaj platformy jest bliższy naturalnemu funkcjonowaniu muzyki: jako plastyczne tworzywo łączące ludzi, a nie zamknięty w gablocie eksponat muzealny (Bannan, 2025)

Aspekt "czarnej skrzynki" modeli głębokiego uczenia używanych przez powyższe platformy utrudnia zrozumienie działania ich rekomenderów. Przyczyn nagłej popularności piosenek czy krótkich filmów po prostu nie da się ustalić. To z kolei utrudnia systematyczne rozwiązywanie problemów tendencyjności i stronniczości. Wewnętrzne działanie omawianych modeli jest często nieprzejrzyste, co utrudnia interpretację ich procesów decyzyjnych. Ten brak informacji sprawia, że pełne zrozumienie sprawczej roli algorytmu w kształtowaniu kultury muzycznej jest wyzwaniem – na przykład, czy piosenka jest popularna ze względu na swoją treść, czy też dlatego, że algorytm zdecydował się ją wypromować na podstawie subtelnych, niezrozumiałych dla człowieka wskazówek, a może nawet z powodu niezamierzonego uprzedzenia? Ta wiedza w nieprzejrzystych systemach jest z definicji niedostępna.

Muzyczne rekomendery stały się potężnymi czynnikami kształtującymi współczesną kulturę muzyczną. Ich algorytmiczna natura, sposób przetwarzania danych i interakcji z użytkownikami nadają im znaczącą autonomiczną sprawczość. Bez odgórnego nadzoru aktywnie wpływają na ekosystem muzyczny. Wynikają z tego różne zagrożenia. Systemy rekomendacji mają tendencję do tworzenia subiektywnych, zamkniętych sfer wpływu, znanych jako bańki filtrujące lub komory echa. Są to spersonalizowane środowiska, w których algorytmy uczą się preferencji użytkownika i serwują mu głównie treści podobne do tych, które wcześniej polubił. W konsekwencji wzmacnia to zastane preferencje i ogranicza ekspozycję internauty na odmienne gusta i opinie (Acerbi, A. 2019).

Taki mechanizm może prowadzić do negatywnych efektów psychologicznych, takich jak wzmocnienie błędu potwierdzenia (confirmation bias), lenistwo poznawcze czy nawet zmniejszona empatia (Gottron, 2016), ponieważ użytkownicy rzadziej kwestionują swoje poglądy i mają trudności ze zrozumieniem odmiennych punktów widzenia. Ten ogólny problem przekłada się też na muzykę. Powstaje realne ryzyko zawężenia nawyków słuchania i ograniczenia odkrywania nieznanymi gatunków lub artystów (Toolify.ai, 2024). Słuchacze mogą utknąć w pętli znanych utworów i zawężyć swoje muzyczne horyzonty. W szerszej skali istnieje także ryzyko, że algorytmy faworyzujące popularne lub mainstreamowe treści prowadzą do zawężenia zbiorowego gustu i homogenizacji krajobrazu muzycznego. Jest to potęgowane przez pętle sprzężenia zwrotnego, gdzie popularne utwory stają się jeszcze popularniejsze dzięki rekomendacjom.

Jednocześnie systemy te posiadają potencjał do dywersyfikacji gustów poprzez wprowadzanie użytkowników w niszowe oraz względnie i bezwzględnie nowatorskie treści, których inaczej by nie odnaleźli. Określa się to mianem serendipity - serendipityczności (Ricci i in., 2022), czyli zdolności do szczęśliwych, przypadkowych

odkryć. Ta zdolność jest jednak często ograniczana przez nadrzędny cel platform, jakim jest optymalizacja pod kątem maksymalizacji zaangażowania użytkownika w oparciu o znane mu już wzorce.

Ekosystem streamingu stwarza dla artystów paradoksalną sytuację, w której zalety bezprecedensowej widoczności i globalnego zasięgu są równoważone przez pogłębione fundamentalne problemy ekonomiczne i systemowe nierówności (Hesmondhalgh, 2020). Nowy model biznesowy, nakierowany na przychody platform, a nie na bezpośrednie transakcje artystów i słuchaczy, doprowadził do powstania systemu płatności za pojedyncze odtworzenia. Przeważnie stawki są bardzo niskie. To, w połączeniu ze złożoną strukturą kontraktów, sprawia, że dla większości artystów osiągnięcie stabilnego dochodu z samej muzyki stało się jeszcze trudniejsze niż dotychczas (Towse, 2020; Hesmondhalgh, 2020). Problem pogłębiają systemowe błędy w algorytmach oraz strukturalna dominacja największych graczy. Największe wytwórnie muzyczne kontrolują większość odtworzeń i treści na kluczowych playlistach, co daje im uprzywilejowaną pozycję (Henry, 2024). Przy tym wszystkim algorytmy rekomendacyjne często wpadają w pętlę błędu popularności, czyli dodatkowo promują treści o największych zasięgach. Pogłębia to barierę wejścia dla początkujących twórców (Henry, 2024).

Odrębny problem bierze się stąd, że nieprzejrzyste rekomendery uniemożliwiają pełną kontrolę nad strategią promocji. Artyści muszą jakoś sobie radzić w tej nowej sytuacji. Często prowadzi to do prób optymalizacji cech swojej twórczości pod algorytmy. To dodatkowa presja selekcyjna w memetycznej ewolucji muzyki. Twórcy modyfikują strukturę piosenek lub ich brzmienie, aby sprostać wymogom rekomenderów i zwiększyć swoje szanse na pojawienie się w playlistach (Henry, 2024). W efekcie, choć streaming pozornie demokratyzuje dostęp do muzyki, w praktyce często wzmacnia istniejące w przemyśle muzycznym hierarchie i nierówności (Hesmondhalgh, 2020). Z badań (Silviya, 2025) wynika, że algorytmy konsekwentnie nagradzają określone cechy utworów, jak na przykład krótszy czas trwania, co sprzyja maksymalizacji odtworzeń i widoczności. Wyciągając wnioski ze wszystkich przytoczonych powyżej faktów i

zależności, należy stwierdzić, że systemy rekomendacji mają znaczny wpływ nie tylko na dystrybucję zastanego oraz nowego repertuaru muzycznego, ale także na samą treść nowego repertuaru.

Mechanizmy sprawczości systemów rekomendacji ujawniały się przy opisie wielu aspektów ich funkcjonowania, ale warto je podsumować w jednym miejscu. Jak zaznaczyłem wcześniej, systemy te działają jak silniki memetyczne, które identyfikują udane doświadczenia muzyczne (i-memy) i próbują je replikować poprzez podsuwanie podobnych treści. W ten sposób wpływają one również na selekcję i propagację nagrań - memów muzycznych - w oparciu o własne kryteria optymalizacji.

Ponadto, użytkownicy coraz częściej przenoszą wysiłek poznawczy (Risko, 2016) związany z odkrywaniem muzyki na algorytmy. Następuje przejście od roli aktywnego odkrywcy/konesera do pasywnego konsumenta treści proponowanych przez algorytmy (Hagen, 2016). Warto jednak przy okazji nadmienić, że zachęcająca do dokonywania remiksów platforma TikTok pobudza również kreatywność prosumentów – ale to kreatywność improwizacyjna, kompozytorska, a nie pobudzenie kreatywnej eksploracji repertuaru muzycznego. Tradycyjne role muzykologów, krytyków, DJ-ów radiowych jako kuratorów i interpretatorów są kwestionowane lub umniejszane, gdy algorytmy przejmują selekcję i kontekstualizację. Ludzcy eksperci tracą monopol na wyjaśnianie muzycznych kanonów i wprowadzanie nowych artystów (Hagen 2016). Pojawiają się nowe role dla ludzkich muzycznych kuratorów: praca obok algorytmów, dostarczanie im ludzkiego pierwiastka (human touch), czy też wprowadzanie kontekstu kulturowego, którego algorytmy mogą nie uchwycić.

Zjawisko przenoszenia wysiłku poznawczego (*cognitive offloading*) (Risko, 2016) w odkrywaniu muzyki, choć wygodne dla użytkowników, może prowadzić do stopniowego

zaniku umiejętności aktywnej eksploracji muzycznej oraz zmniejszonego doceniania wiedzy kontekstowej i historycznej, tradycyjnie dostarczanej przez ludzkich ekspertów. Użytkownicy polegają na algorytmach w kwestii odkrywania muzyki, redukując własny wysiłek aktywnego poszukiwania. Jest to forma przenoszenia wysiłku poznawczego, przekazująca zadanie selekcji systemowi. Długotrwałe poleganie na takich systemach może osłabić własne umiejętności użytkownika w nawigowaniu po zróżnicowanych krajobrazach muzycznych lub poszukiwaniu muzyki wymagającej/nieznanej. Jednocześnie postrzegana wartość muzykologów czy specjalistycznych kuratorów maleje, gdy ich funkcja wydaje się zautomatyzowana, co potencjalnie prowadzi do utraty głębszego kulturowego zrozumienia muzyki.

Agentowość, sprawczość systemów rekomendacji nie jest monolityczna. To złożona gra wzajemnych oddziaływań między projektem algorytmu (odzwierciedlającym cele platformy), danymi, które przetwarza (odzwierciedlającymi zbiorowe zachowania użytkowników, które same są kształtowane przez algorytm) oraz indywidualnymi wzorcami interakcji użytkownika. Tworzy to dynamikę koewolucyjną, w której algorytmy kształtują kulturę muzyczną, a kultura muzyczna (wyrażona poprzez dane) kształtuje algorytmy. Algorytmy maksymalizują zaangażowanie słuchaczy oraz przychód. Najpierw uczą się wzorców zachowań użytkowników, a później na nie wpływają. To zmienione zachowanie wraca do algorytmu jako nowe dane, tworząc ciągłą pętlę sprzężenia zwrotnego. Dlatego agentowość-sprawczość nie leży wyłącznie po stronie rekomendera ani wyłącznie po stronie użytkownika, ale wyłania się z tej dynamicznej, rekursywnej interakcji, którą możemy określić jako techno-biokulturowe, czy też tremetyczno-memetyczne pętle koewolucji.

3.2.6.1. Muzyczne AI – wprowadzenie, rys historyczny

W niniejszej pracy bardzo istotną rolę stanowi opis kultury muzycznej w świetle dynamicznie rozwijającej się Sztucznej Inteligencji - AI. To kulminacja przeróżnych wątków. Już poprzedni rozdział wykazał, że AI towarzyszy treściom muzycznym i multimedialnym pod postacią systemów rekomendacyjnych od wielu lat. Wszystkie technologie AI związane z muzyką stanowią dla niniejszej pracy przedmiot szczególnego zainteresowania z kilku względów.

Po pierwsze, technologie AI są w świetle memetyki owocem memetycznej ewolucji kultury. Co więcej, algorytmy wprost aproksymujące ewolucję (algorytmy ewolucyjne, algorytmy genetyczne) oraz te bardzo do niej podobne (*reinforcement learning*, uczenie przez wzmocnienie) odgrywają kluczową rolę w projektowaniu najnowszych technologii AI (Eiben, 2015; Sutton, 2018). Zatem AI to nie tylko doskonały przykład ewolucji w kulturze. To bezprecedensowy przykład stworzenia kulturowego produktu, który integruje ewolucję samą w sobie, zarówno w historii (sposobie) swojego powstania, jak i w metodzie działania.

Co ciekawe, podobne podwójne mechanizmy selekcyjne obserwujemy również u człowieka. Nasz gatunek powstał na drodze złożonej ewolucji, a jednocześnie selekcja aktywnie zachodzi w naszych ciałach. W układzie odpornościowym poprzez selekcję klonalną przeciwciała konkurują o skuteczność wobec antygenów (Carter, 2000). Za to według teorii darwinizmu neuronalnego Edelmana grupy neuronalne podlegają wewnętrznej selekcji w zależności od doświadczeń i sprzężeń zwrotnych (Edelman, 1987; Seth, 2005). Te analogie pokazują, że procesy ewolucyjne możemy stosować zarówno w wyjaśnieniach ultymatywnych jak i proksymalnych. Zatem można powiedzieć, że pod względem tej analogicznej dwoistości procesów selekcyjnych technologie AI przypominają funkcjonowanie organizmów żywych.

Kenneth O. Stanley i Joel Lehman na podstawie własnych sukcesów w dziedzinie algorytmów genetycznych i sztucznej inteligencji wprowadzają koncepcję 'poszukiwacza skarbów' (*treasure hunter*). Ich zdaniem precyzyjny plan i cel to w dziedzinie AI strategie krótkowzroczne, a znacznie korzystniejsze jest eksperymentowanie i uważne rozglądanie się za niespodziewanym przełomem (Stanley, 2015).

Można powiedzieć, że najnowsze systemy AI są raczej *hodowane* niż manualnie *konstruowane* i *projektowane*. Tworzenie nowych, bardziej skutecznych architektur sztucznych sieci neuronowych nie polega na 'manualnym' (intencjonalnym, prowadzonym teleologicznie i racjonalnie przez człowieka) projektowaniu. Badacze uruchamiają (emulują) na komputerze proces ewolucyjny, który rozpoczyna się od zebrania grupy różnych neuronalnych architektur jako osobników w populacji. Następnie są one rozmnażane i mutują przez wiele generacji, aby odnaleźć takie rozwiązania, które są jednocześnie adekwatne oraz efektywne obliczeniowo (Lu, 2019). A w kolejnym kroku te optymalne, ale jeszcze nienasycone treścią, głębokie sztuczne sieci neuronowe dosłownie rosną w wyniku procesu 'karmienia danymi', co stanowi kolejny etap ewolucji sieci (Zou, 2024).

Autorzy słynnego artykułu *Attention is all you need*, który stoi u podstaw trwającej obecnie rewolucji AI (9998 cytowań wg Semantic Scholar), sformułowali w nim bardzo ważne prawo projektowania AI: wystarczy, aby istniał mechanizm uwagi, a karmienie większą ilością danych wystarczy, żeby wyewoluować poprzez emergencję coraz to bardziej inteligentne sieci neuronowe (Vaswani, 2017). Na marginesie warto tutaj wspomnieć, że w pokrewnym duchu, odnośnie pracy z opartymi na ewolucji generatywnymi systemami muzycznymi wyrażał się bardzo uznany kompozytor i producent muzyczny Brian Eno. Już w 1999 Eno wygłosił przemowę Kompozytorzy jako Ogrodnicy (*Composers as Gardeners*), w której w ciekawych słowach zaprezentował zniuansowaną i wyrafinowaną opinię na tytułowy temat (Eno, 1999).

Wracając do uzasadnienia, dlaczego muzyczne AI ma szczególną wagę dla niniejszej pracy. Po drugie i być może najważniejsze – to właśnie kształt i charakter najnowszych narzędzi AI najlepiej wykazują skuteczność predykcyjną paradygmatu memetycznego. Najbardziej niepowtarzalną i odróżniającą od konkurencyjnych paradygmatów ewolucji kulturowej cechą memetyki jest to, że postuluje ona, że elementarnym budulcem kultury jest replikator, który w toku ewolucji będzie dążył do wykształcenia zachowania mającego na celu własne przetrwanie.

A zatem memetyka już ze swoich podstawowych założeń przewiduje, że kultura u zarania powinna być bezkształtnie dryfującą bezwonną pierwotną zupą chaotycznych cząsteczek, a następnie wraz z upływem czasu, stopniowo i powoli, coraz bardziej wykształcać celowość, intencjonalność, zachowanie, podobnie jak stało się to w procesie ewolucji biologicznej wraz z początkiem życia. Nie chodzi tu jednak o teleologiczne założenie istnienia zewnętrznego planu, lecz o rozumienie celowości w sensie proksymalnym, czyli jako obserwowalne ukierunkowanie zachowań organizmów na określone stany świata (np. zdobycie pożywienia, unikanie zagrożenia, rozmnażanie). Tego rodzaju celowość jest funkcjonalna i emergentna. Powstaje w wyniku doboru naturalnego, który promował układy zachowujące się w sposób sprzyjający przetrwaniu (Mayr, 1961; Laland, 2000).

Podobnie intencjonalność można ujmować jako zdolność organizmów do wytwarzania i używania uproszczonych reprezentacji elementów środowiska, które kierują działaniem (Proust, 1995). Nie oznacza to od razu w pełni świadomej, ludzkiej intencjonalności, lecz raczej stopniowo wykształcane mechanizmy percepcyjno-reaktywne, które można interpretować jako formy proto-reprezentacji (Proust, 1995). W ujęciu teorii systemów rozwojowych (*developmental systems theory*) właściwości te nie są wyłącznie produktem genów, lecz powstają we współdziałaniu genotypu, fenotypu i środowiska (Oyama, 2001). W tym sensie „celowość” i „intencjonalność” są zjawiskami emergentnymi, które nie wymagają odwołania się do metafizycznego teleologizmu, lecz dają się wytłumaczyć poprzez procesy selekcji i rozwoju. Analogia z kulturą polega na tym, że również systemy memetyczne z czasem mogą wytwarzać zachowania pozornie

intencjonalne — w rzeczywistości będące wynikiem selekcji i utrwalania skutecznych strategii replikacyjnych. Analiza dziejów ewolucji kultury pod kątem wykształcania przez nią własnej sprawczości wykazuje zgodność z tą przytoczoną wyżej podstawową predykcją memetyki. Moimi przykładami działania tego procesu są dzieje kultury muzycznej. Być może dzisiejsze systemy AI nie wykazują jeszcze zachowania, celowości i intencjonalności w całkowitym i niepodważalnym sensie, ale widzimy ewidentnie, że aproksymują te kategorie znacznie lepiej niż ich memetyczne poprzedniczki sprzed roku, dwóch, pięciu, dziesięciu czy stu lat.

Główną tezę niniejszej pracy jest to, że muzyka podlega ewolucji na sposób memetyczny, a zarazem z przeglądu literatury wynika, że paradygmat memetyczny uchodzi za zdyskredytowany. Zatem, aby obronić tezę, należy wykazać, że koncepcja ewolucji memetycznej jest słuszna, a analiza systemów AI jest sposobem, aby to wykazać. Trzecim powodem, dla którego muzyczne AI leży szczególnie w zakresie niniejszej pracy, jest to, że ma ona mówić o procesach ewolucji muzyki w świecie technologii cyfrowych, a muzyczne AI jest właśnie najświeższym owocem tejże ewolucji. To znaczy – rozumiejąc ewolucję nie jako ciąg ‘od gorszego do lepszego’ (popularny błąd jeszcze z czasów Spencera (Spencer, 1857)), tylko ciąg nieustannych chwilowych przystosowań do bieżących warunków środowiska, AI jest obecnie potężną siłą kształtującą środowisko kultury. Muzyka, kultura muzyczna, intensywnie dopasowuje się do tego kontekstu. Opisanie muzycznej AI jest zatem opisem najnowszego stadium ewolucji kultury muzycznej.

Zgodnie z paradygmatem memetycznym nic w kulturze nie może pojawić się nagle i ex nihilo, zatem i korzeni sztucznej inteligencji jako technologii generatywnej można upatrywać bardzo daleko, nawet w starożytności. Tym, szczególnie w kontekście muzycznym, zajmuje się Marcin Strzelecki (Strzelecki, 2024), który zauważa, że już renesansowy kontrapunkt stanowił przykład generatywnego systemu muzycznego. Ta tradycja trwa nieprzerwanie aż do dziś. Znacznie później, w XX wieku, twórczość Johna Cage’a czy Iannis Xenakisa były matematycznie bardzo spójne z ideą generowania

muzyki poprzez system. Obaj kompozytorzy oddalali się od bezpośredniej sprawczości, kontroli nad materiałem muzycznym, w różny sposób delegując ją 'na zewnątrz siebie'.

W tym miejscu zatem dla porządku przywołam jedynie mocno skróconą historię muzyki i sztucznej inteligencji, żeby uwypuklić, jak rozpoczęła się historia splotu obu tych dziedzin. Myśląc maszyny w kontekście muzycznym konceptualnie projektowała już w 1843 Ada Lovelace, czyli znana pionierka w zakresie informatyki, uznawana za twórczynię pierwszego komputerowego algorytmu (Essinger, 2014). Lovelace współpracowała z Charlesem Babbage'em, twórcą *maszyny analitycznej*, czyli pierwszej mechanicznej maszyny liczącej. To Lovelace wpadła na pomysł, że przetwarzane przez maszynę liczby mogą symbolicznie reprezentować także zupełnie inne, pozamatematyczne treści, czyniąc jednocześnie poważny krok w konceptualizacji komputera (Murtaugh, 2024). Zaproponowała użycie maszyny analitycznej do tworzenia muzyki:

Załóżmy, na przykład, że elementarne relacje tonów z nauki o harmonii i kompozycji muzycznej można było w ten sposób wyrazić i zaadaptować, nasz silnik mógłby komponować wyszukane i naukowe dzieła muzyczne o dowolnym stopniu złożoności i czasie trwania.

Supposing, for instance, that the fundamental relations of pitched sounds in the science of harmony and of musical composition were susceptible of such expression and adaptations, the engine might compose elaborate and scientific pieces of music of any degree of complexity or extent. (Lovelace, 1843)

Lovelace skonceptualizowała zatem maszynę obliczeniową jako generatywny system do tworzenia muzyki. Co więcej, jako pierwsza skonceptualizowała muzykę w formie obliczalnych dla komputera informacji. Ciekawe, że od razu wspomniała o dowolnym (czyli innymi słowy potencjalnie nieskończonym) stopniu złożoności i czasie trwania. Faktycznie te predykcje już dziś można uznać za zrealizowane. Co prawda dosłownie nieskończony poziom złożoności i nieskończony czas trwania to nieosiągalna granica, ale za to trwa wykonanie takich utworów jak Longplayer (Finer, 2000). Utwór rozpoczął się 31 grudnia 1999, a jego koniec przypada na 31 grudnia 2999, zatem ta komputerowa kompozycja oparta na algorytmicznej manipulacji dźwiękiem tybetańskich

mis ma trwać dokładnie 1000 lat. To bardzo znane i odbywające się w czasie rzeczywistym wykonanie, ale nic nie stoi na przeszkodzie, żeby przy pomocy komputera wygenerować obecnie plik muzyczny (np. MIDI) trwający nawet sto tysięcy lat. Albo sto tysięcy takich plików i określić je jako kolejne części tego samego dzieła. A zatem stworzenie prostego proceduralnego utworu o czasie trwania 10 miliardów lat nie jest wyzwaniem dla współczesnych komputerów.

Zatem predykcje Lovelace należy ocenić za wyjątkowo trafne, szczególnie, że sformułowała je w roku 1843, czyli jeszcze za czasów życia Fryderyka Chopina. Prognoza Lovelace była bezpośrednim wynikiem jej wyjątkowego podejścia, które sama określała mianem 'nauki poetyckiej' (Woolley, 2020). Inżynier Babbage, jej mentor, koncentrował się na praktycznym celu tworzenia bezbłędnych tablic matematycznych i nawigacyjnych (Woolley, 2020), a interdyscyplinarne spojrzenie Lovelace pozwoliło jej dostrzec potencjał maszyny wykraczający poza arytmetykę. Ta zdolność do myślenia abstrakcyjnego, połączenie rygoru matematyki z artystyczną wyobraźnią, była warunkiem koniecznym do narodzin idei kreatywnej AI. Lovelace sformułowała także niezwykle ważną uwagę krytyczną. Stwierdziła:

Silnik analityczny nie może stworzyć nic oryginalnego. Zrobi wszystko, co będziemy potrafili mu nakazać. Potrafi analizować, ale nie ma mocy przewidywania relacji analitycznych czy też prawd.

The Analytical Engine has no pretensions whatever to originate anything. It can do whatever we know how to order it to perform. It can follow analysis; but it has no power of anticipating any analytical relations or truths.

Lovelace, 1843

To stwierdzenie, nazwane później przez Alana Turinga 'Zastrzeżeniem Lady Lovelace' (Lady Lovelace's Objection) (Turing, 1950), do dziś pojawia się jako fundamentalna krytyka kreatywności w AI. Jeszcze w 2025 np. Lee Cronin, specjalista od komputerów chemicznych (Cronin, 2024), ale również jeden z twórców przytaczanej w tej pracy teorii złożań, wielokrotnie publicznie wyrażał swój pogląd, że AI nie jest i nigdy nie będzie

prawdziwie kreatywne, a ta sfera autentycznej kreatywności jest zarezerwowana wyłącznie dla biologicznych organizmów żywych (RSE, 2023).

Ale nie można powiedzieć, że Zastrzeżenie Lady Lovelace zatrzymało rozwój AI. Wręcz przeciwnie. Stanowi ono pierwszą precyzyjną artykulację centralnego problemu filozoficznego w dziedzinie kreatywności obliczeniowej. Lovelace jednoznacznie zaznaczyła różnicę między wykonywaniem poleceń a autentycznym tworzeniem. Definiując tę granicę, Lovelace sformułowała wyzwanie, które nadal napędza badania nad AI, choć minęło już 150 lat. W 2001 roku na jego podstawie powstał nawet tak zwany 'Test Lovelace' (Bringsjord, 2001). Zakłada on, że aby uznać maszynę za autentycznie inteligentną i kreatywną, musi ona stworzyć coś, czego jej programiści nie są w stanie wyjaśnić na podstawie jej kodu i architektury. Innymi słowy, maszyna musi nas zaskoczyć w sposób, który nie jest wynikiem jej zaprogramowanych instrukcji. Stanowi to ciekawe i przydatne rozszerzenie słynnego testu Turinga.

Natomiast sam Alan Turing bardzo przyczynił się do dzisiejszej rozpoznawalności Ady Lovelace poprzez odniesienie się do niej w swoim klasycznym artykule stanowiącym kamień milowy informatyki *The Imitation Game* (Turing, 1950). Turing rozdzielił pola inżynierii myślących maszyn od filozoficznego problemu świadomości. Zaproponował wtedy swój słynny test, znany jako test Turinga, w którym sędzia prowadzi komunikację neutralnym kanałem (tekstowym) z maszyną i człowiekiem. Jeśli sędzia nie potrafi rozróżnić, które wiadomości pochodzą od człowieka, a które od maszyny, maszynę uznajemy za inteligentną (Turing, 1950).

Turing, jak wspominałem wcześniej, bezpośrednio odniósł się do *Zastrzeżenia Lady Lovelace*. Zgodził się, że maszyna sama w sobie nie jest w stanie stworzyć niczego nowego, ale argumentował, że można ją zaprogramować tak, aby uczyła się i modyfikowała swoje zachowanie w sposób, który zaskoczyłby samych jej programistów, co w efekcie można określić jako formę tworzenia (Turing, 1950; Essinger, 2014). Ta właśnie koncepcja została w 2001 sformalizowana jako Test Lovelace (Bringsjord, 2001).

Ponadto Turing sformułował koncepcję *maszyny uniwersalnej* (dziś znanej jako maszyna Turinga). Przekonywał, że zaprogramowany komputer cyfrowy może symulować działanie każdej innej maszyny (Turing, 1950). Owa uniwersalność ma głębokie implikacje dla sztuki. Istnienie takiej maszyny mocno sugeruje, że kreatywność nie jest unikatową własnością jakiejś szczególnej, niepojętej maszyny artystycznej. Wręcz odwrotnie, skoro kreatywność można opisać w kategoriach obliczeniowych, można ją także zautomatyzować. Innymi słowy, jeśli proces twórczy da się skutecznie opisać jako zbiór operacji, to maszyna uniwersalna jest w stanie, przynajmniej w teorii, te operacje wykonać. W ten sposób Turing zdemistyfikował akt twórczy, przekształcając go z niemal magicznego aktu w złożone zadanie przetwarzania informacji. W ten sposób maszynowa twórczość stała się uprawnionym przedmiotem badań naukowych i otworzyło się pole dla późniejszych pionierów muzyki algorytmicznej.

Wizje Lovelace i Turinga znalazły swoją pierwszą konkretną, słyszalną realizację w 1957 roku na Uniwersytecie Illinois w Urbana-Champaign. To tam chemik i kompozytor Lejaren Hiller, we współpracy z Leonardem Isaacsonem, wykorzystał pięciotonowy komputer ILLIAC I do stworzenia *Illiad Suite for String Quartet* – pierwszej kompletnej partytury muzycznej skomponowanej przez maszynę (Hiller, 1959). Dzieło to stanowiło bezpośrednie przejście od teorii do praktyki, materializację spekulacji. Forma utworu składa się z czterech eksperymentów, z których każdy testował inny aspekt kompozycji algorytmicznej (Hiller, 1959). Dzięki temu *Illiad Suite* sam w sobie jest świetnym studium przypadku wczesnych paradygmatów AI.

W pierwszym eksperymencie, czyli pierwszej części utworu, komputer generował jednogłosowy *cantus firmus* w oparciu o zakodowane reguły szesnastowiecznego kontrapunktu. Było to czyste ucieleśnienie podejścia symbolicznego w sztucznej inteligencji, nazywanego dziś z przymrużeniem oka ‘Stare, dobre AI’ „Good Old-Fashioned AI” (GOFAI). GOFAI wymaga formalizacji i implementacji wiedzy eksperckiej. W tym przypadku stanowią ją reguły teorii muzyki (Funk, 2018).

Eksperyment drugi polegał na tym, że program generował czterogłosowe segmenty. Stosował w tym celu zmodyfikowane reguły harmoniczne, aby przejść od prostych melodii do bardziej złożonych tekstur (Hiller, 1959). Autorzy dalej rozwinęli tę formułę w części trzeciej. Do reguł melodycznych i harmonicznnych dodali zasady dotyczące rytmu, dynamiki i artykulacji. Tak otrzymana muzyczna struktura posiadała więcej ciekawych niuansów (Hiller i Isaacson, 1959).

Za to w finałowej części następuje istotna zmiana. Tutaj Hiller i Isaacson nie mają już całkowitej kontroli. Do generowania materiału muzycznego posłużyły probabilistyczne łańcuchy Markova (Funk, 2018). Ten eksperyment reprezentował drugi główny nurt wczesnej AI, oparty na statystyce, a nie na sztywnej logice. Funk zauważa, że proces tworzenia Illiac Suite zapoczątkował trwały paradygmat artystycznej współpracy człowieka z AI (Funk, 2018). To nie komputer ILLIAC I podjął decyzję o napisaniu kwartetu smyczkowego. To Hiller i Isaacson zaprojektowali autorskie eksperymenty. Musieli oni sami zdefiniować i zakodować reguły. A po komputerowej generacji to oni dokonali subiektywnej estetycznej selekcji tych fragmentów, które uznali za najciekawsze.

Ten model w znacznej większości przypadków pozostaje aktualny aż do dziś. Maszyna pełni rolę generatora w ramach określonych przez człowieka, a sam artysta zmienia się w kuratora, redaktora i ostatecznego arbitra wartości estetycznej. Illiac Suite dowiodła, że wizja Lovelace i Turinga jest możliwa do zrealizowania, ale jednocześnie ukazała wczesną twórczość maszynową jako proces ściśle nadzorowany przez ludzkich twórców. W latach 80. XX wieku rozpoczęła się nowa epoka komputerowej twórczości generatywnej, zapoczątkowana przez kompozytora i programistę Davida Cope'a. Możemy powiedzieć, że dzięki niemu muzyka generatywna rozpoczęła wejście w XXI wiek. Cope rozpoczyna przejście od przestarzałej 'Starej dobrej sztucznej inteligencji' GOFAI, ściśle kontrolowanej przez eksperckie zasady, do nowoczesnej, oddolnej, ewoluującej i nieprzewidywalnej AI, która to stała również u podstaw rewolucji lat 2020 czy też ery ChatGPT.

Cope'owski projekt Experiments in Musical Intelligence (EMI) oryginalnie miał być osobistym narzędziem do walki z kryzysem twórczym. Kompozytor chciał, aby komputer wiernie naśladował jego własny styl (Cope, 1989; Johnson, 1997). Ostatecznie rezultaty okazały się na tyle ciekawe, że dziś EMI stanowi kamień milowy muzyki generatywnej. Przede wszystkim, wcześniej omawiany Iliac Suite to zamknięta, całkowicie deterministyczna kompozycja (z elementami pseudolosowości w części 4). EMI to nie jedna kompozycja, a otwarty system, który został użyty do wielu rozmaitych generacji.

Za najważniejszą techniczną innowację Cope'a można uznać to, że autor nie instruował EMI wprost, jak wyglądają zasady muzyki, ścisłe ramy kompozycyjne, których ma przestrzegać. Zamiast tego karmił swój program dużymi zbiorami danych, jak na przykład wszystkie chorały Bacha. Następnie autor zlecał systemowi samodzielne odkrycie wzorców, które definiują styl kompozytora (Cope, 1992; Coenen, 1997). W tak dużym skrócie brzmi to jak bardzo nowoczesna technologia, ale owo karmienie i uczenie nie przebiegało w sposób tak otwarty, jak w dzisiejszych systemach (Briot, 2020). System EMI ma bardzo jasne kategorie do zbadania, jak na przykład statystyki następstw wysokościowych dźwięków. Z drugiej strony uczenie w nowoczesnych systemach często polega na analizie surowych danych, z czego sam system wytwarza w ogóle całe abstrakcyjne kategorie, według których te dane zacznie porównywać. EMI jest systemem znacznie prostszym, o mniejszej dowolności i nieprzewidywalności. Dokładnie realizuje postawione zadanie, którym jest odkrycie statystycznych prawideł w partyturach.

Dokładniej, metodologia EMI opiera się na trzystopniowym procesie (Cope, 1991). Dwa pierwsze stopnie stanowią analizę, a trzeci syntezę. Rozpoczyna się od dekonstrukcji: system analizuje korpus muzyczny oraz identyfikuje cechy melodyczne, harmoniczne i rytmiczne. Krok drugi to odszukiwanie powtarzających się wzorców. Przy odpowiedniej ilości danych ujawniają one prawidłowości statystyczne wyjątkowe dla stylu danego kompozytora. (Cope, 1989; da Silva, 2003). Trzeci, kreatywny etap to rekombinacja: EMI wykorzystuje zidentyfikowane struktury i reguły, aby generować nowe utwory. Są

one spójne stylistycznie z oryginalnym korpusem, ale nie stanowią dokładnej kopii żadnego dzieła (Cope, 1991).

Ten napędzany danymi model okazał się niezwykle skuteczny. Muzyka generowana przez EMI była na tyle przekonująca, że wielokrotnie zwodziła ekspertów i publiczność podczas muzycznych testów Turinga. W słynnym eksperymencie zorganizowanym przez Douglasa Hofstadtera, publiczność uznała utwór EMI za autentyczne dzieło Bacha, podczas gdy prawdziwy utwór Bacha został uznany za kompozycję komputerową (Johnson, 1997). Zatem EMI przeszło swoisty muzyczny test Turinga. EMI stanowi pomost między symboliczną GOFAL a nowoczesnym uczeniem maszynowym. Chociaż EMI nie wykorzystywało sieci neuronowych, jego metodologia oparta na analizie danych i odkrywaniu wzorców to bezpośredni koncepcyjny prekursor dzisiejszych modeli.

Praca Cope'a to także doskonała inspiracja dla rozważań estetycznych nad naturą kreatywności i stylu. Według Hofstadtera sukces EMI podważa romantyczny mit o geniuszu (Johnson, 1997). Skoro wiernie odtwarza styl kompozytora, to znaczy, że jego metafizyczne tajemnice można zamknąć w matematycznej formule (Johnson, 1997).

Wyróżnię jeszcze jeden, ostatni krok w ewolucji muzycznego AI przed nadejściem rewolucji transformerów i LLMów. Jak dotąd stabilnie poruszamy się w kierunku stopniowego zanikania ludzkiego racjonalnego planowania i pełnej kontroli na rzecz pobudzania aktywności maszyny, która uczy się wzorców. Tak jest i w tym wypadku. Dalsze pogłębienie oddolnej sprawczości komputera otrzymujemy przez zastosowanie uczenia maszynowego oraz sztucznych sieci neuronowych. Pionierską pracą w tej dziedzinie było badanie Petera M. Todda z 1989 roku, który jako pierwszy zastosował Rekurencyjną Sieć Neuronową (RNN) do kompozycji muzycznej (Todd, 1989). Sieci RNN są specjalnie zaprojektowane do przetwarzania danych sekwencyjnych. Posiadają wewnętrzną pamięć, która pozwala im uwzględniać poprzedzające elementy sekwencji

podczas generowania kolejnych. Dzięki temu stały się popularnym narzędziem do modelowania muzyki i języka (Briot, 2020).

Proste sieci RNN mają jednak krótką pamięć, co badacze precyzyjniej określili mianem zanikającego gradientu (Hochreiter, 1997). W związku z tym kiepsko przyswajają długoterminowe zależności w danych, a więc nie nadają się do analizy formy dzieła muzycznego. Ten problem został rozwiązany jeszcze w latach 90. XX wieku. Sepp Hochreiter i Jürgen Schmidhuber w 1997 opracowali architekturę Długiej Krótkoterminowej Pamięci LSTM (Hochreiter, 1997). Sieci LSTM jako zaawansowany typ RNN wprowadziły mechanizm bramek, które pozwalają sieci decydować, co zapisać w pamięci, co z niej usunąć, a co wykorzystać do wygenerowania wyniku (Briot, 2020; Mon, 2025). Ta zdolność do utrzymywania i wykorzystywania informacji w długich sekwencjach okazała się kluczowa dla muzyki.

Pierwszą muzykę sieci LSTM stworzyły pięć lat później, w 2002 roku. Przy ich pomocy Douglas Eck i Jürgen Schmidhuber wygenerowali przekonującego bluesa (Eck, 2002). Ich model nauczył się odpowiednich sekwencji dźwięków, ale też bluesowych progresji i typowych rytmów. To dowód na to, że sieci LSTM potrafią przyswoić sobie cały zbiór cech gatunku muzycznego (w tym wypadku raczej prostego gatunku, którym jest blues).

Prace Todda, oraz Ecka i Schmidhubera ze sztucznymi sieciami neuronowymi ustanowiły paradygmat architektoniczny, który zdominował generowanie muzyki na niemal dwadzieścia lat. To także bardzo duży skok technologiczny. Zacementowały przejście do modelu, w którym wiedza muzyczna nie jest kodowana explicite, ale wyłania się jako statystyczna właściwość dużej sieci neuronowej wytrenowanej na obszernym zbiorze danych. Ta skalowalna architektura, oparta na uczeniu się przewidywania kolejnego elementu w sekwencji, stała się bezpośrednim fundamentem dla potężniejszych modeli, które miały nadejść, w tym dla architektury Transformer, która napędza dzisiejszą rewolucję w dziedzinie generatywnej AI.

3.2.6.2. Muzyczne AI - case studies

Samo istnienie narzędzi AI do tworzenia muzyki nie musi oznaczać jeszcze, że powstaną utwory o istotnym wpływie na kulturę. AI w muzyce wzbudza niechęć i wątpliwości. Sonda społeczna przeprowadzona przez *YouGov* wskazuje, że ponad połowa młodych Amerykanów z pokolenia Z nie jest zainteresowana słuchaniem muzyki tworzonej lub współtworzonej przez AI, a u starszych pokoleń ten wskaźnik tylko wzrasta (Bansal, 2024). Z kolei z ankiety przeprowadzonej przez brytyjską organizację *BPI* (British Phonographic Industry) wynika m.in. że 82.7% ankietowanych wierzy, że ludzka kreatywność jest kluczowa w procesie tworzenia muzyki oraz ponad 80% ceni muzykę tworzoną przez ludzi wyżej niż tę tworzoną przez AI (BPI, 2024).

Tymczasem recepcja coraz większej liczby przypadków konkretnych utworów tworzonych przy użyciu AI podaje w wątpliwość te deklaracje. Różnorodne studia przypadku prezentują muzykę tworzoną lub współtworzoną przez AI, która zyskała szeroką popularność oraz w wielokrotnie bardzo pozytywne przyjęcie. Pomimo deklaracji, statystyki odsłuchań przekonują, że słuchacze często pozytywnie reagują na utwory tworzone z udziałem AI (Gomez Sarmiento, 2025). Może być to związane między innymi z tym, że AI wcale nie jest łatwo rozpoznawalne. Kiedy te muzyczne technologie brzmiały nieprzekonująco, nie były znane. Dziś również muzyczne AI to jest zjawiskiem bardziej znanym niż kilka lat temu, ale jakość generatywnych technologii - rozumiana tutaj wąsko i specyficznie jako zdolność imitacji stylu ludzkiego - gwałtownie wzrosła. Tym samym rozpoznanie, że dany produkt kultury, mem, utwór muzyczny został stworzony przy udziale AI, staje się coraz trudniejsze (Cros Vila, 2025).

Za pierwszy album muzyki pop skomponowany z pomocą AI uchodzi *I AM AI* artystki Taryn Southern z 2018 roku. Główny singiel *Break Free* skomponowany z udziałem programu *Amper* zyskał ponad dwa miliony odsłon. Album *I AM AI* nazywany jest również pierwszym albumem skomponowanym z pomocą AI, który znalazł się na liście *Top 100 U.S. Radio Charts* (Southern, b.d.). Strona internetowa artystki wskazuje, że w

produkcji albumu zostały użyte programy IBM Watson Beat, Amper, AIVA, Google Magenta oraz NSynth (Southern, 2018). Z perspektywy 2025 owe narzędzia w ówczesnych wersjach wydają się już przestarzałe, żeby nie powiedzieć prymitywne. Pakiet prasowy albumu zdradza, że wszystkie narzędzia AI zostały użyte wybiórczo w poszczególnych utworach, a nad nimi twórczą kontrolę sprawowali ludzie: Taryn Southern z pomocą producenta Ethana Carlsona. Utwór *Voices* ewidentnie używa AI w większym stopniu, co słyhać po nietypowej, dziwacznej dla piosenki pop formie, instrumentacji i schematycznej harmonii, ale przede wszystkim pomyłonym (niezgodnym z konwencją) basowym dźwięku syntezatora. Pomyłonym w konwencji popowej ballady – jako muzyka eksperymentalna może takie rozwiązanie jednak zostać ocenione pozytywnie. Ostatecznie cały projekt i współpracę Taryn Southern z AI można zakwalifikować jako hybrydę eksperymentu oraz średnio udanego chwytu marketingowego, gdyż AI nie miało znacznej kontroli w tworzeniu tej muzyki, a płyta zyskała mniejszą popularność niż inne produkty tej samej artystki.

W 2023 roku na kanał YouTube zespołu AISIS trafił album *Lost Tapes*. reklamowany jako album z alternatywnej rzeczywistości, w którym grupa OASIS się nie rozpadła. Twórcami tej muzyki jest brytyjski zespół Breezer, a udział AI stanowi teoretycznie niewielki, ale jednak kluczowy element. Głos wokalisty grupy Breezer został dzięki AI zmieniony tak, aby imitować brzmienie Liama Gahllaghery z OASIS. Dzięki temu, że sam zespół mocno i kompetentnie inspirował się OASIS, a brzmienie partii głosu zostało przetworzone przekonująco, efekt definitywnie został osiągnięty. Grupa w wywiadach wspomina, że w wyniku pasma niepowodzeń już w zasadzie planowała się rozwiązać, a dzięki albumowi *Lost Tapes* otrzymali międzynarodowy rozgłos i ruszyli w prawdziwą trasę koncertową. W 2025 *Lost Tapes* na YouTube przekroczyło 690 tysięcy odsłon. Co ciekawe również sam Liam Gahllagher zapoznał się z albumem i wyraził się o nim pozytywnie. Z rozlicznych komentarzy w sieci wynika, że zespół Breezer doskonale zastąpił pustkę wynikającą z przedłużającego się rozpadu grupy OASIS i zyskał uznanie wśród fanów oryginalnej grupy. Co ciekawe OASIS, grupa zawieszona w 2009, akurat w 2024, rok po premierze *Lost Tapes* ogłosiła, że wraca do koncertowania. Nie istnieje żaden dowód na wynikowe powiązanie tych wydarzeń, ale warto odnotować tę

informację nawet jeśli jest całkowitym zbiegiem okoliczności. Album *Lost Tapes* dzięki AI w pewnym sensie wdarł się już w historię prawdziwego OASIS i prawdopodobnie pozostanie już z nią trwale powiązany.

4 kwietnia 2023 światło dzienne ujrzał utwór *Heart on my Sleeve* w wykonaniu Drake'a oraz The Weeknd, autorstwa osoby pod pseudonimem Ghostwriter977. Drake oraz The Weeknd to celebryci, bardzo znani wokaliści, którzy w tym wypadku zostali zastąpieni przez imitującą ich sztuczną inteligencję. Mówiąc bardziej precyzyjnie, ten utwór, który wstrząsnął przemysłem muzycznym, był praktycznie w całości skonceptualizowany i skomponowany przez człowieka, anonimowego twórcę, który podpisał się jako Ghostwriter977. Owszem, zawarte w piosence partie głosowe niezwykle przekonująco naśladowały głosy dwóch gwiazdorów (Pahwa, 2023; Reed, 2023). Ale metoda była dokładnie ta sama jak w wypadku AISIS – na gotowy utwór nałożono filtry wokalne oparte na uczeniu maszynowym, które przetransformowały barwę głosu (Klaes, 2025; de Luna, 2023). Utwór błyskawicznie stał się wiralowym hitem, gromadząc 600 000 odtworzeń na Spotify, 15 milionów wyświetleń na TikToku i setki tysięcy na YouTube w ciągu zaledwie kilku dni (Snapes, 2023). To wynik, o którym marzą nawet największe światowe gwiazdy. Choć metodyka pracy z AI była dokładnie taka sama jak w przypadku AISIS, recepcja utworu przez właścicieli oryginalnych wizerunków była drastycznie odmienna. Gahlagher pochwalił swoich naśladowców, a Drake wyraził oburzenie (Ingham, 2023). Najmocniej Ghostwriter977 poruszył jednak globalnych monopolistów z Universal Music Group.

UMG, wytwórnia reprezentująca Drake'a oraz The Weeknd, określiła utwór jako treść naruszającą prawo i doprowadziła do jego szybkiego usunięcia z głównych platform streamingowych (Snapes, 2023; CoverNet.ai, 2024). Taki ruch możemy oczywiście interpretować jako panikę wielkiej firmy z powodu artystycznego i kreatywnego precedensu. Oto dzięki AI, bezkosztowo, domowy twórca może wyprodukować przebój brzmiący jak produkt ich ogromnej korporacyjnej maszyny.

Co ciekawe, analiza prawna wykazała, że tak szybkie działanie w oparciu o ustawę Digital Millennium Copyright Act było możliwe z powodu wykorzystania przez Ghostwritera977 krótkiego, chronionego prawem autorskim producer taga należącego do producenta Metro Boomin (Guez, 2023; Klaes, 2025). Ściślej mówiąc, producer tag to praktyka producentów muzycznych w kulturze hip-hopowej, swoisty dźwiękowy podpis. W tym wypadku to około 3 sekundy, dokładnie wypowiedziane na początku utworu jedno zdanie: „*If young Metro don't trust you, I'm gonna shoot you*”. A zatem formalnie nie została naruszona żadna własność intelektualna Drake'a, tylko producenta Metro Boomin.

Bez tego technicznego szczegółu sprawa byłaby znacznie bardziej skomplikowana. Prawo autorskie w Stanach Zjednoczonych co do zasady nie chroni stylu wykonawczego ani barwy głosu, co ujawniło ciekawą lukę w tym systemie prawnym. Rozgorzała nowa legalistyczna dyskusja dotycząca prawa do wizerunku (*right of publicity*), które chroni unikalne cechy tożsamości osoby przed komercyjnym wykorzystaniem bez zgody (Reed, 2023; Guez, 2023). Piosenka *Heart on my Sleeve* jest kompletna, to interesujący, wciągający utwór, który nie bez przyczyny zyskał wielką popularność w kilka dni. Wykorzystanie głosów znanych piosenkarzy to formalnie, w kontekście skomponowania i wyprodukowania całej piosenki jedynie detal. Zaprzęgnięcie armii prawników do zwalczania tego pięknego przejawu kreatywności niewątpliwie przeraziło i odwiodło wielu domowych producentów od podzielenia się swoimi dziełami w podobnym stylu.

Działanie świetnie prosperującej firmy, którą jest Universal Music Group, ma na celu oczywiście zarabianie pieniędzy. Jak starałem się wcześniej pokazać, podrzędne cele w takiej strukturze mają pewną swobodę, ale ostatecznie muszą działać w zgodzie z najwyższym hierarchicznie poziomem systemu. A zatem UMG i inni podobni monopolisci ze względu na swój monstrualny rozmiar są niemalże skazani na poszukiwanie zarobku i jednocześnie zabijanie oddolnej kreatywności niezależnych twórców. W cynicznym cytacie z komunikatu prasowego firma deklaruje co innego:

Trenowanie generatywnej Sztucznej Inteligencji używając muzyki naszych artystów (co jest naruszeniem zgód i pogwałceniem praw autorskich) oraz fakt, że te łamiące prawo treści stworzone przy użyciu Sztucznej Inteligencji mogą być dostępne na serwisach streamingowych, prowokuje pytanie, po której stronie historii chcą się znaleźć wszelcy interesariusze muzycznego ekosystemu: po stronie artystów i kreatywnej ekspresji ludzkiej, czy po stronie deep fake'ów, oszustów i pozbawiania artystów ich prawowitego wynagrodzenia.(...) Pokrzepia nas zaangażowanie naszych partnerów z różnych platform w te sprawy – jako że zdają sobie oni sprawę z tego, że muszą być częścią rozwiązania tego problemu.

“The training of generative AI using our artists’ music (which represents both a breach of our agreements and a violation of copyright law) as well as the availability of infringing content created with generative AI on DSPs, begs the question as to which side of history all stakeholders in the music ecosystem want to be on: the side of artists, fans and human creative expression, or on the side of deep fakes, fraud and denying artists their due compensation,(...) We’re encouraged by the engagement of our platform partners on these issues - as they recognize they need to be part of the solution.” -UMG, 2024

Ten ubrany w fałszywą troskę komunikat to w istocie groźba wystosowana przez monopolistów. Narrację, w której wielkie wytwórnie i portale streamingowe dbają o dochody artystów warto skonfrontować choćby z tym, że Spotify płaci artystom średnio cztery dziesiąte centa za odtworzenie. Czyli jeśli utwór na Spotify stanie się prawdziwym hitem i osiągnie liczbę miliona odsłón, artysta zarobi około cztery tysiące dolarów.

Pozwy przeciwko firmom Suno (UMG, 2024) oraz Udio (UMG, 2024b) złożone przez największe amerykańskie korporacje muzyczne można uznać za zabawne oraz przerażające. Korporacje, które na każdym kroku wykorzystują artystów, teraz pozują na ich obrońców, aby etycznie uzasadnić własną chęć budowania kapitału. Groźba kary 150 000 dolarów za nieuzgodnione trenowanie modelu AI na pojedynczym utworze sprowadza się do wielomiliardowych kar, które całkowicie pogrzebałyby młode startupy AI. Zbiorowy pozew przede wszystkim skupia się na tym, że generatywne AI może zdestabilizować rynek, a więc zaszkodzić artystom. Zarazem Sony jawnie deklaruje budowanie inicjatywy nazywanej etycznym AI (Sony, 2025), czyli w zasadzie własnej kopii Suno i Udio. Różnica byłaby tylko taka, że ostatecznie na generatywnym AI zarabiałoby Sony, a nie potencjalnie konkurencyjne nowe startupy.

Nie musimy być tak straszliwie oburzeni twórczością AI. W końcu przykład Liama Gallaghery z OASIS pokazuje, że artyści mogą przychylnie podchodzić do własnych cyfrowych imitacji. Pogląd taki publicznie wyraziła również Grimes, znana wokalistka, producentka i kompozytorka:

Podzielę się 50% zysków z twórcą każdej udanej piosenki wygenerowanej przez AI, która wykorzystuje mój głos. Taki sam układ proponuję każdemu artyście, z którym współpracuję. Możesz używać mojego głosu swobodnie i nie spotka cię za to żadna kara. Nie należę do żadnej wytwórni muzycznej i nie wiążą mnie żadne kontrakty.

I'll split 50% royalties on any successful AI generated song that uses my voice. Same deal as I would with any artist i collab with. Feel free to use my voice without penalty. I have no label and no legal bindings (Grimes, 2024).

Zatem widać, że artyści jak Gallagher czy Grimes mają do tych spraw zupełnie inny stosunek niż UMG czy Sony. Co ciekawe, również i późniejsze podejście Drake'a stoi w całkowitej sprzeczności z oficjalną linią jego wytwórni. Dosłownie rok po oburzeniu za nieprawne wykorzystanie głosu Drake'a, artysta sam bezprawnie wykorzystuje głosy innych w ten sam sposób we własnym utworze. Ten konkretny przypadek omówię w odrębnym, szerszym kontekście, jako osobne studium przypadku. Bardzo interesujący materiał na odrębne case study stanowi różnorodne użycie AI w muzycznym konflikcie pomiędzy amerykańskimi raperami. Beef Kendrick Lamar kontra Drake to także ciekawy przypadek memetycznej ewolucji, w której miesza się kreatywność ludzi oraz AI.

W kulturze hip-hopowej *beef* odnosi się do publicznego sporu lub rywalizacji między artystami. Tradycja lirycznego pojedynku sięga jeszcze potyczki Słowackiego i Mickiewicza (Zakrzewski, 1964). W tym wypadku jednak to walka muzyczna, wyrażana głównie poprzez *dissy*, czyli rapowe utwory muzyczne obrażające oponenta, przedstawiające go w złym świetle, czy też mówiące „całą prawdę”, czyli niewygodne fakty na jego temat.

Elementami beefu mogą być także wydarzenia wokół samych utworów muzycznych, takie jak wywiady czy ostatnio posty w mediach społecznościowych. Konflikty te mogą wynikać z osobistych nieporozumień czy też rywalizacji o status. Beefy są głęboko powiązane z korzeniami gatunku – improwizowanymi bitwami rapowymi. Służą jako narzędzie do potwierdzania dominacji, obrony reputacji i zdobywania rozgłosu. Badacze i krytycy zauważają, że beefy mogą kształtować tożsamość i trajektorię kariery artysty, jednocześnie napędzając zaangażowanie wśród fanów. Mogą one jednak również eskalować do przemocy w świecie rzeczywistym, co budzi obawy o zatarte granice między ekspresją artystyczną a osobistą wrogością (Watkins, 2005; Alim, 2006).

Konflikt hip-hopowy między Kendrickiem Lamarem a Drake'em rozwijający się od 2012, w 2024 roku wykroczył poza ramy tradycyjnej rywalizacji artystycznej. Stał się ogromnym, skomplikowanym wydarzeniem kulturalnym. Głos w tym konflikcie zabrał nawet były prezydent USA Barrack Obama (Limehouse, 2024). Szerokie omówienia całego beefu można znaleźć w sieci (Glynn, 2025; Wikipedia contributors, 2025). Jednym z jego interesujących aspektów jest wykorzystanie AI, co może stanowić interesujące stadium przypadku użycia AI w dyskursie publicznym. Technologia ta została wykorzystana na dwa sposoby: jako scentralizowane narzędzie ofensywne oraz jako podstawa dla zdecentralizowanego, wiralowego fenomenu.

Pierwszy przypadek dotyczył bezpośredniego użycia technologii deepfake przez Drake'a. W utworze *Taylor Made Freestyle* artysta opublikował sfabrykowane wokale Snoop Dogga oraz zmarłego Tupaca Shakura, mające na celu prowokację i psychologiczną dekonstrukcję wizerunku Lamara jako dziedzica sceny West Coast (Horowitz, 2024). Głos Tupaca zarapował słowa świadczące o tym, że jest zawiedziony Lamarem. Takie działanie Drake'a spotkało się z natychmiastową reakcją prawną ze strony spadkobierców Tupaca, którzy, powołując się na naruszenie praw do wizerunku, zmusili Drake'a do usunięcia utworu. Tym niemniej utwór wciąż można bez problemu odnaleźć w internecie, umieszczany tam przez nieoficjalne źródła (np. Username 701, 2024). Niezgoda spadkobierców Tupaca była również związana z tym, że Lamar w

kulturze hip-hopowej zachodniego wybrzeża USA cieszy się wielkim szacunkiem. Zatem ten głos został niejako zabrany oddolnie – Drake wykorzystał Tupaca w formie AI, aby zaatakować Kendricka, ale zwolennicy Tupaca – ludzie zaangażowani w jego mempleks – odpowiedzieli kontratakiem. Incydent ten uwypuklił fundamentalne problemy etyczne i prawne związane z nieautoryzowanym wykorzystaniem cyfrowych podobizn zmarłych artystów. Co więcej, skądinąd bardzo ciekawy i twórczy zabieg Drake’a ostatecznie można uznać za nieudaną ofensywę właśnie dlatego, że raper myślał, iż może zmodyfikować mempleks Tupaca tak, aby realizował jego własną wolę – przypisać mu własny, nadrzędny cel. Skutek był odwrotny do zamierzonego: zaktywizowany przez Drake’a mempleks Tupaca, rękoma spadkobierców, wyraził sprzeciw, a nawet znacząco osłabił pozycję Drake’a w tej długiej rapowej wojnie.

Atak Drake’a przy użyciu AI można określić jako scentralizowany. To sam raper stworzył jeden konkretny utwór z głosami AI. Za to drugi w historii tego beefu przypadek użycia AI, tym razem wymierzony przeciwko Drake’owi, demonstruje partycypacyjny i oddolny charakter. Jest to mem „BBL Drizzy”. Genezą zjawiska kulturowego znanego jako „BBL Drizzy” był pejoratywny pseudonim nadany Drake’owi przez rapera Ricka Rossa w relacjach w mediach społecznościowych, które stały się integralną częścią wojny. Ross, choć formalnie nie był sprzymierzony z Lamarem, również został przez Drake’a wciągnięty do tego złożonego konfliktu i odpowiedział także utworem muzycznym. Jednak najbardziej istotne znaczenie – największy memetyczny sukces – odniosło samo określenie BBL Drizzy, które jednocześnie było swego rodzaju oskarżeniem o to, że Drake poddał się procedurze Brazilian Butt Lift, czyli operacji kosmetycznego powiększenia pośladków.

Następnie artysta Willonius Hatcher użył generatora Suno AI, aby stworzyć krótki utwór *BBL Drizzy* w stylu RnB. Brzmienie lat 70. połączone z bardzo żywym memetycznie hasłem stworzyło prawdziwy humorystyczny przebój. Utwór ten zyskał znaczną popularność w mediach społecznościowych w bardzo krótkim czasie. Platforma Spotify usunęła piosenkę, kiedy liczyła już dwa miliony i sześćset tysięcy odsłon (Brutus, 2024).

Zdobyła również popularność na Tiktoku – filmik, na którym influencerka podająca się za zakonnice tańczy do utworu BBL Drizzy zyskał ponad 13 milionów odsłon (Sister Mary Blaze, b.d.).

Drake w trakcie beefu z Lamarem w swoich piosenkach antagonizował również innych artystów, licząc, że w ten sposób zbuduje swój status jako nietykalnego króla hip-hopowego świata. Między innymi zaatakował słownie producenta muzycznego Metro Boomin, z którym wcześniej wielokrotnie współpracował. Boomin, bardzo znany i cieszący się estymą producent, nie odpowiedział na zaczepki Drake'a rapując, ale zaszkodził mu znacznie bardziej w inny sposób. Artysta zremiksował stworzony przy użyciu AI utwór *BBL Drizzy*, tworząc rapowy beat. Boomin opublikował ten beat za darmo pod nazwą *BBL DRIZZY 150 BPM* i oznajmił, że każdy, kto zechce, może zarapować do tego podkładu swój własny diss na Drake'a (Boomin, b.d.). Beat ten stał się wiralowym narzędziem, umożliwiając tysiącom użytkowników tworzenie własnych, prześmiewczych treści wymierzonych w Drake'a, co przekształciło mem w zdecentralizowaną i niezwykle skuteczną kampanię wizerunkową (Rosenberg, 2024).

Opisany powyżej cykl zdarzeń można ująć jako ewolucję memetyczną: najpierw powstało zaprojektowane celowo hasło BBL Drizzy, następnie w wyniku jego szerokiej popularności powstał utwór skomponowany przez Suno AI, który wywołał wiele oddolnych remiksów. Pojawił się także jeden szczególnie znany remiks, zmieniający formę piosenki RnB na rapowy beat, do którego następnie wielu artystów nagrało własne rapowane teksty, co stworzyło kolejną specyficzną falę remiksów dalej opartych na lotnym, zaraźliwym hasle BBL Drizzy. Podsumowując, beef między Lamarem a Drake'em stanowi precedens w analizie zastosowań AI w kulturze muzycznej. Ukazuje dwoistą naturę tej technologii: z jednej strony jako kontrowersyjnego narzędzia deepfake w rękach pojedynczego aktora, z drugiej jako katalizatora dla oddolnych, partycypacyjnych zjawisk kulturowych, które redefiniują dynamikę współczesnych konfliktów w sferze publicznej.

3.2.6.3. Muzyczne AI - przegląd technologiczny

Jak próbowałem wykazać wcześniej, rozwój technologii AI sprawia, że memy muzyczne zyskują coraz większą swobodę replikacji, umniejszając rolę intencjonalnych ludzkich strategii w muzycznym ekosystemie. Poza tym muzyczne technologie AI są także najnowszymi owocami memetycznej ewolucji muzyki, które znacznie wpływają na jej obecny kształt i kierunek zmian. W tym rozdziale omawiam zarówno technologie widocznie prowadzące do rosnącej autonomii przedmiotów kultury muzycznej, jak i najnowsze narzędzia, których rola w środowisku muzycznym dopiero zaczyna się wyłaniać.

Jak wspominałem, przez około dwadzieścia lat dominującą architekturą AI do tworzenia muzyki były sztuczne sieci neuronowe RNN oraz LSTM (Hochreiter, 1997). Znaczny przełom nastąpił w 2017 wraz z nadejściem uniwersalnej architektury *transformer* i publikacją artykułu *Attention is all you need* (Vaswani, 2017). Wcześniejsze sieci przetwarzały dane tylko sekwencyjnie, krok po kroku, za to transformery opierają się na mechanizmie uwagi i samouwagi. To istotna zmiana. Te mechanizmy pozwalają modelowi na dynamiczne ważenie istotności wszystkich elementów wejściowych przy generowaniu poszczególnych elementów wyjściowych. Sieć może zatem bezpośrednio odnosić się do dowolnego fragmentu sekwencji, niezależnie od odległości pomiędzy fragmentami, co eliminuje problem zależności dalekiego zasięgu (Vaswani, 2017). Innymi słowy, transformery pokonały trudność tworzenia powiązań pomiędzy odległymi elementami sekwencji, czyli w muzyce np. relacji harmonicznym rozciągniętych na wiele taktów.

Mechanizm samouwagi zrewolucjonizował sposób przetwarzania muzyki przez sztuczne sieci neuronowe. Przykładowo, pozwala on modelowi rozpoznać, że dana wysokość dźwięku w melodii jest powiązana z konkretnym akordem w harmonii, pomimo tego, że pojawił się kilka taktów wcześniej. Starsze sieci nie mogły także w praktyce rozpoznać, że określony motyw rytmiczny może być powtórzeniem innego

motywu z początku utworu. Dzięki transformerom dostrzeganie podobieństw, regularności w strukturze wykracza poza najbliższe sąsiedztwo. To bezprecedensowy w środowisku AI sposób na generowanie złożonej muzyki, która jest wewnętrznie spójna i posiada logiczną strukturę (Huang, 2019).

Ponadto, dzięki przewyciężeniu sekwencyjnego charakteru obliczeń, trenowanie stało się znacznie szybsze, a więc takie sieci mogą w ograniczonej czasowo praktyce analizować znacznie większe korpusy danych. Transformery pierwotnie miały służyć do maszynowego tłumaczenia tekstów. Wywodzą się z dziedziny przetwarzania języka naturalnego (natural language processing, NLP), ale ich zastosowanie jest dziś znacznie szersze. W związku z ich niesamowitą wydajnością i skutecznością odniosły sukcesy w bardzo wielu dziedzinach. Dość powiedzieć, że słynny GPT z ChatGPT to właśnie Generatywny Pretrenowany Transformer (Generative Pretrained Transformer). Bardziej precyzyjnie, rodzina GPT to takie transformery, które wyróżniają się z dwóch względów. Po pierwsze są charakterystycznie trenowane, po drugie mają charakterystyczne zastosowanie.

Trening, czy też uczenie GPT obejmuje dwa etapy. Ta technika pozwala na bezprecedensową generalizację wiedzy. Na początek odbywa się pre-trening: transformer uczy się uniwersalnej reprezentacji danych poprzez zadanie predykcyjne. W przypadku języka naturalnego jest to zazwyczaj modelowanie językowe, czyli przewidywanie następnego słowa zadanego tekstu (Radford et al., 2018). Dzięki temu procesowi model uczy się gramatyki, w tym składni, semantyki, a nawet pewnej wiedzy o świecie zawartej w danych treningowych. Korpus powinien być ogromny, zróżnicowany i nieopisany etykietami. Brak etykiet znaczy, że jest to proces nauki nienadzorowanej, transformer sam odnajduje regularności i struktury w danych. Nie ma żadnego innego jasno określonego celu do optymalizacji (Brown, 2020).

Po ogólnym pre-treningu model przechodzi drugi etap szkolenia. Tutaj transformer przygotowuje się precyzyjnie do konkretnego zadania. Podstawą nadzorowanego dostrajania, czyli końcowego etapu uczenia modelu na starannie dobranych

przykładach z gotowymi odpowiedziami, jest znacznie mniejszy, etykietowany zbiór danych. Łączenie złożonych wewnętrznych reprezentacji modelu z jasno zrozumiałymi zewnętrznymi etykietami prowadzi do doskonałych wyników w różnorodnych zadaniach (Radford, 2019).

Jak wspomniałem wyżej, GPT to nie każdy pre-trenowany transformer, ale taki, który sprawdza się do pewnej konkretnej grupy zastosowań. Definiuje je generatywność – GPT tworzą nowe treści. Z drugiej strony nie-generatywne transformery, takie jak BERT (Devlin, 2019), głęboko analizują tekst i rozpoznają przeróżne jego cechy: nastrój, temat, styl. Tworzą użyteczne klasyfikacje (Qiu, 2020). Pełnia twórczego paradygmatu GPT przejawia się głównie w domenie językowej, ale także i muzycznej. Już na etapie pre-treningu do uczenia modelu zamiast danych słownych można wykorzystać dane muzyczne, np. w standardzie MIDI. Wtedy przetrenowany transformer będzie przewidywał kolejne dźwięki lub inne ‘zdarzenia muzyczne’ (Huang, 2019). Dalej taki model dostraja się do takich zadań jak kontynuacja melodii lub generowanie muzyki w stylu danego kompozytora (Payne, 2019).

Sieć MuseNet od OpenAI z 2019 teoretycznie potrafi z fragmentu dzieła Mozarta rozwinąć utwór w stylu Chopina (OpenAI, 2019). Niewątpliwie wynik tego działania jest spójny z jej wewnętrznymi reprezentacjami i może nawet przekonać pracowników OpenAI, skoro postanowili się nim pochwalić. Jednak osoba jakkolwiek zaznajomiona ze stylem Chopina niewątpliwie wyrazi zdrowy sceptycyzm. Według mojej własnej analizy stylokrytycznej MuseNet co najwyżej uczy się pewnych chopinowskich ozdobników. Trzeba jednak dodać, że styl Chopina jest bardzo charakterystyczny, subtelny, ekspresyjny, co może być przyczyną trudności w skutecznym zamknięciu go w matematycznej formule.

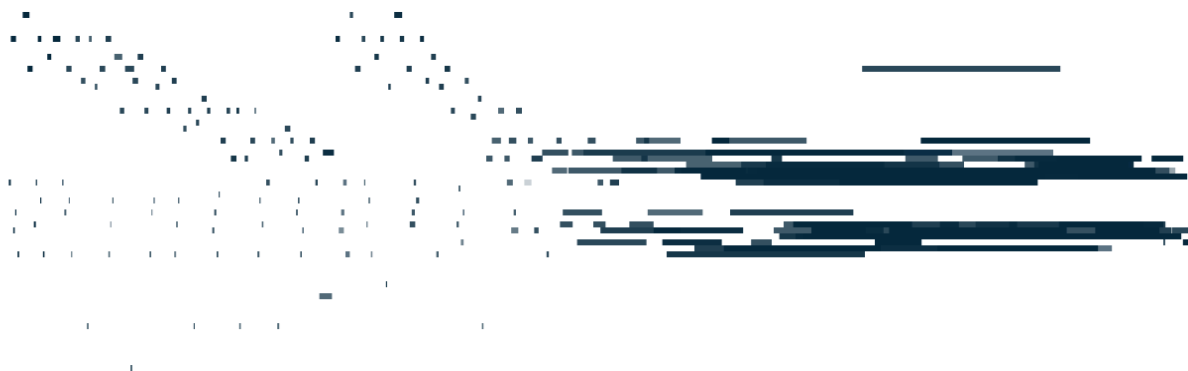
Bardziej przekonującą muzykę fortepianową w podobnym czasie generował muzyczny transformer oparty na technologii Magenta od Google AI (Simon, 2019). Autorzy również prezentowali przykład przetwarzania muzyki Fryderyka Chopina. Tym razem jednak celem nie było wygenerowanie nowej kompozycji w szeroko rozumianym stylu

chopinowskim, ale naśladowanie logiki konkretnego dzieła, Etiudy Ges-dur op. 10 nr 5. Badacze użyli dosłownie 14 pierwszych nut jako materiału, który sieci następnie miały rozszerzyć. Zademonstrowali w ten sposób różnice pomiędzy różnymi architekturami. Ich sieć LSTM stworzona specjalnie do ekspresyjnego wykonawstwa całkowicie zawiodła w kontynuacji logicznej struktury muzycznej, co dodatkowo wspiera to, co pisałem wcześniej, że LSTM-y nie nadają się do tworzenia złożonych, spójnych struktur muzycznych (Huang, 2018).



LSTM kontynuuje Etiudę, wyciąg MIDI – Huang, 2018

W tym samym badaniu lepiej poradził sobie zwyczajny transformer – ewidentnie zreplikował elementarną łukową strukturę głównego motywu Etiudy, z naprzemiennym długim pochodem w dół, a następnie w górę klawiatury.



Klasyczny transformer kontynuuje Etiudę, wyciąg MIDI – Huang, 2018

Zdecydowanie najlepiej poradził sobie prezentowany w badaniu Muzyczny Transformer (Music Transformer), który ewidentnie zrozumiał logikę głównego motywu i usiłował ją rozwinąć w nową, szerszą formułę. Jest to wyraźnie widoczne na wyciągu MIDI:



Muzyczny transformer kontynuuje Etiudę, wyciąg MIDI – Huang, 2018

Pomimo takiego sukcesu dziś to nie transformery stanowią podstawową architekturę AI do generowania muzyki, ich naczelną modalnością pozostają dane w formie językowej. Istotną przyczyną problemów z kreatywną adaptacją transformerów do bezpośredniego przetwarzania i wytwarzania muzyki jest dylemat reprezentacji i tokenizacji (Mousavi, 2025), czyli rozbicia ciągłego strumienia danych (np. tekstu lub muzyki) na podstawowe „klocki”, którymi model potrafi operować. Podczas gdy zwyczajny tekst językowy w naturalny, linearny sposób dzieli się na dyskretne jednostki (słowa, znaki), złożoną muzykę trudno przedstawić w formie jednego rzędu elementarnych symboli (Briot, 2017; Haiduk, 2022). Sygnał audio jest ciągły, a w każdej jednostce muzycznego czasu mieści się także złożoność tradycyjnie rozumiana jako wymiar pionowy dźwięku, jak współbrzmienia i akordy. Ale i to nie wyczerpuje muzycznej złożoności. Nawet w reprezentacji symbolicznej (MIDI) mamy do czynienia z wieloma równoległymi wymiarami, takimi jak wysokość dźwięku, czas trwania, głośność, barwa (instrument).

Wybór odpowiedniego sposobu konwersji muzyki na sekwencję dyskretnych tokenów dla transformera jest wysoce nietrywialnym problemem badawczym. Jak przyznał CEO firmy Suno, opracowanie 'właściwego sposobu tokenizacji audio' było jednym z największych wyzwań w budowie ich modelu (Shulman, 2024).

Muzyczna rola transformerów nie kończy się jednak na monolitycznych autoregresyjnych systemach do generacji audio. Przede wszystkim duże modele językowe LLM jak ChatGPT są używane do pisania tekstów słownych do piosenek (Hind, 2024; Frederick, 2025). Można również z powodzeniem wytwarzać w nich w formie tekstowej akordy do piosenek (am, C#7 itp), struktury utworów symfonicznych, kreatywne koncepcje do tworzenia dzieł muzyki współczesnej, kod do interaktywnych muzycznych aplikacji... ograniczeniem jest głównie wyobraźnia użytkownika. Przykłady takiej twórczości muzycznej poprzez różne rodzaje tekstu zamieszczam w części czwartej niniejszej pracy.

Poza tym, oprócz podejść symbolicznych, architektura transformera zaczęła być również szeroko stosowana w domenie audio, w tym jako fundament w wielu nowszych modelach dyfuzyjnych do generowania muzyki i dźwięku (Evans, 2024). Uwagi wymagają same modele dyfuzyjne, które uzyskiwały już ciekawe wyniki w generacji obrazów bez pomocy architektury transformer (Ho, 2020). W 2022 Forsgren i Martiros wykonali niezwykle interesujący eksperyment w dziedzinie generacji muzyki z AI. Znany model Stable Diffusion, przystosowany do przetwarzania i generowania obrazków, wytrenowali na obrazkach etykietowanych spektrogramów. Spektrogram to graficzna reprezentacja zmian widma dźwięku w czasie, czyli po prostu jedna z metod wizualizacji przebiegu sygnału audio. Wystarczyło tylko dodać na wyjściu sieci algorytm do przetwarzania wygenerowanych spektrogramów w pliki dźwiękowe. W ten sposób wytrenowane na spektrogramach Stable Diffusion zmieniło się w Riffusion – program generujący muzykę z obrazków na podstawie wprowadzanych przez użytkownika promptów tekstowych (Forsgren, 2022). Riffusion pierwotnie działał jako całkowicie otwarty generator online bez żadnych limitów. Tworzenie w nim opisuję w czwartej części tej pracy. Program ten

dziś już w pierwotnej wersji nie istnieje, ale prace nad oryginalnymi prototypami są kontynuowane przez innych badaczy (Li, 2024).

Riffusion po wejściu na rynek przegrał konkurencję z innymi programami do generowania muzyki na podstawie promptów tekstowych – Udio i Suno. Moim zdaniem stało się tak, ponieważ konkurencja potrafi wiernie naśladować style i brzmienie muzyki popularnej, za to Riffusion miał bardziej eksperymentalny brzmieniowy charakter. Muzyce i sztuce tworzonej przez AI zarzuca się, że nie wnosi ona nic nowego, że to tylko zapożyczenia, powtórki, rekombinacje istniejących stylów i dzieł (van Hees, 2025). Za to Riffusion, w swoim eksperymentalnym duchu, dzięki prostej architekturze dyfuzyjnej, wnosi do świata sztuki dźwięku autentyczną nowość. Żeby przedstawić jej charakter, wprowadzę najpierw elementarny opis modeli dyfuzyjnych.

Modele dyfuzyjne służą głównie do generacji realistycznego obrazu, tekstu językowego i dźwięku. Ich trening przebiega w bardzo charakterystyczny sposób: uczą się generować dane poprzez stopniowe oczyszczanie losowego szumu (Ho, 2020). W odróżnieniu od autoregresywnych transformerów, które najlepiej sprawdzają się do dyskretnych danych, podejście dyfuzyjne, oparte na stochastycznej manipulacji szumem, świetnie sprawdza się do manipulacji ciągłym sygnałem audio (Wang, 2024). Naukę stopniowego dostrzegania regularności w szumie dzieli się na dwa etapy: najpierw dodaje się szum do prawdziwych danych, aż staną się niemal losowe, a następnie model uczy się, jak ten proces odwrócić (Sohl-Dickstein, 2015). W fazie odszumiania model krok po kroku usuwa zakłócenia. Przewiduje przy tym, jak poprawnie powinny wyglądać dane. Pracuje nad nimi, aż uzyska czystą próbkę (Ho et al., 2020). Dzięki temu model potrafi z losowego szumu wygenerować nowe dane, które wyglądają podobnie do przykładów z treningu, np. zdjęcia lub nagrania mowy (Kong et al., 2020). Siłą modeli dyfuzyjnych jest stabilność i zdolność do odtwarzania bardzo realistycznych detali. Dzięki temu są dziś jednymi z najważniejszych narzędzi w dziedzinie kreatywnej, generatywnej sztucznej inteligencji (Chen et al., 2020).

Podsumowując, praktyczne działanie Riffusion opisałbym następująco: wytrenowany model otrzymuje od użytkownika opis audio do wygenerowania, np. 'doom metal ukulele'. Riffusion otrzymuje losowy szum i stara się odnaleźć w nim podobieństwa do materiału, który wyobraża sobie jako 'doom metal ukulele' na podstawie danych z treningu. Krok po kroku z pomocą procesów stochastycznych z szumu wyłaniają się regularności prowadzące coraz bardziej zdecydowanie do jednego konkretnego kształtu. Ostatecznie model kończy proces na pewnym spektrogramie 'wyrzeźbionym' z szumu. Można to podsumować jako wielką, złożoną halucynację, gdzie model z każdym krokiem przekonuje sam siebie, że jednak coś dostrzegł i brnie coraz dalej, aż kończy z zamkniętą całością, gdzie ta halucynacja zostaje urzeczywistniona.

W związku z tak wysokim poziomem abstrakcji manipulacji sygnałem, powstają autentycznie nowe brzmienia i struktury dźwiękowe. Nowe wytworzone spektrogramy najczęściej nie odwzorowują bezpośrednio żadnych prawdziwych instrumentów, głosów, melodii zawartych w zbiorze treningowym. Model naśladuje pewne ogólne kształty spektrogramów, ale ze względu na swoją specyfikę działania praktycznie zawsze ląduje gdzieś 'pomiędzy': pomiędzy saksofonem a trąbką, pomiędzy głosem a perkusją, pomiędzy gitarą a szumem tła. Ta przestrzeń pomiędzy realnymi kategoriami w literaturze nosi nazwę przestrzeni latentnej (latent space) (Bergmann, 2025). Owe charakterystyczne dla wczesnego generatywnego AI lat 2020 fantazyjne halucynacje zostały zaprezentowane w popularnym memie. Użytkowniczka chciała wytworzyć przy pomocy AI obrazek przedstawiający konie, a w efekcie ujawniły się dziwaczne konio-słoniopodobne cienie:



rraaaarri. (2024, September 20). Bah! — I loved when AI art could never be anything but AI... [Tumblr post]. Tumblr. <https://rraaaarri.tumblr.com/post/762192338374311936>

Obrazek z podpisem *secret horses* został popularnym memem internetowym. Proponuję, aby szeroko przyjęte w internetowej społeczności określenie ‘secret horses’ stało się określeniem dla tej specyficznej, halucynowanej estetyki, zarówno w domenie graficznej jak i audio. Dzisiejsze modele w wyniku znacznie bardziej złożonej architektury i zaawansowanych treningów, w których optymalizuje się podobieństwo generatywnych rezultatów do zamkniętych, wygładzonych komercyjnych produktów, nie tworzą już takich eterycznych, abstrakcyjnych dzieł. W związku z tym jest to styl praktycznie wymarły, co tylko podkreśla tempo memetycznej ewolucji technologii AI. W roku 2025 styl z roku 2021 jest wyraźnie rozpoznawalny jako kompletnie archaiczny. Tak nie musi być – systemy takie jak Riffusion można ponownie wytrenować na danych o znacznie wyższej jakości i uzyskać systemy o takim samym charakterze, ale

działające w wyższych rozdzielczościach danych. To jednak wymaga nakładów finansowych, a styl secret horses nie skłania do większych inwestycji.

Przejawy muzyczne tego stylu, czyli eksploracje kategorii wysnionych gdzieś pomiędzy tymi realnymi prezentuje badanie z 2024 (Li, 2024). Do technicznego artykułu dołączone są nagrania dźwiękowe. W zasadzie jest to cała mapa, lub nawet pole, przedstawiające nagrania w różnych miejscach latentnej przestrzeni pomiędzy realnymi kategoriami. Wśród nagrań eksplorujących latentne przestrzenie znajdziemy między innymi dźwięki dokładnie w połowie drogi między: głosem ptaka i gitarą, głosem ptaka i perkusjonaliami, akordeonem i gitarą, kroplami wody i marimbą, dzwonem kościelnym i gitarą, biciem serca i elektroniczną perkusją, dzwonkami na wietrze i skrzypcami, klarnetem i marimbą (Li, 2024). Przykłady te ukazują bezprecedensowe brzmienia wynikające z halucynacji sztucznej sieci neuronowej. Moim zdaniem stanowią one jedną z najważniejszych autentycznych nowości wprowadzonych przez technologie AI do świata muzyki.

Obecnie dwa najbardziej znane, odnoszące komercyjne sukcesy programy do generacji piosenek na podstawie promptów to Suno oraz Udio (Pareño, 2025). Trudno szukać w nich estetyki secret horses – użytkownik otrzymuje gotowy, doszlifowany produkt, który często od razu nadaje się do szerokiej publikacji, jak pokazuje chociażby przykład zespołu The Velvet Sundown (Frelon, 2025). Dokładne architektury modeli obu firm są ściśle tajne. Tym niemniej na podstawie elementarnych deklaracji CEO oraz odwrotnej inżynierii można wnioskować, że łączą one właśnie modele dyfuzyjne z transformerami (Wilson, 2025; Lee, 2024). Choć nie wprowadzają one szczególnych muzycznych nowości, ich rola w muzycznym ekosystemie jest istotna, ponieważ w znaczący sposób oddziałują na rynek muzyczny. Badacze zaznaczają także, że jest to tak daleko idąca demokratyzacja tworzenia muzyki, że aż prowadzi do głębokich filozoficznych pytań na temat twórczości (Pram, 2025; Tan, 2024). Aspekty wpływu muzyki generowanej przez AI na rynek muzyczny opisał także Trzciński (Trzciński, 2023). Z perspektywy niniejszej pracy najważniejsza jest sprowokowana przez tę demokratyzację zmiana dynamiki ewolucyjnej muzyki. Skoro piosenki może tworzyć każdy, kto posiada komputer, telefon,

lub dostęp do internetu, w środowisku muzycznym istotnie zmieniają się presje selekcyjne. Posiadanie instrumentów, znajomość teorii, a przede wszystkim posiadanie wewnętrznej wyobraźni muzycznej (tzw. słyszenia wewnętrznego) stają się niepotrzebne. Poza tym, wartą wspomnienia autentyczną nowością w produkcji muzyki są narzędzia AI do separacji źródeł dźwięku oraz poprawy jakości nagrań. Jedną z najważniejszych nowinek wprowadzonych przez technologie AI jest umiejętność rozpoznawania konkretnych kategorii oraz manipulowania nimi. Na zdjęciu AI wyróżni kota lub psa, a w pliku audio rozpozna saksofon, oklaski i chrząknięcia (Stöter, 2019).

Separacja źródeł z gotowych miksów (pełnych nagrań, w których wszystkie instrumenty są już zmieszane w jedną całość) to inżynierska wersja psychoakustycznego *cocktail party problem* (Cherry, 1953) – czyli problemu wyodrębnienia jednego spośród wielu równoczesnych źródeł dźwięku. Od lat 50. XX wieku (Cherry, 1953) uznawano to za bardzo trudne, niemal nierozwiązywalne zadanie. W praktyce wymagało karkołomnych, ekwilibrystycznych sztuczek produkcyjnych i zwykle dawało ograniczone rezultaty (Rafii, 2018). Dopiero uczenie głębokie umożliwiło praktyczne rozdzielanie partii głosu, perkusji, gitary basowej i innych ścieżek z komercyjnych nagrań. Narzędzia takie jak Open-Unmix i Spleeter wprowadziły i rozpowszechniły nowe standardy w zakresie produkcji (Stöter, 2019; Hennequin, 2020; Wang, 2018). Jakość brzmienia takich wyodrębnionych przy użyciu uczenia głębokiego poszczególnych ścieżek z roku na rok się poprawia. Ta metoda z ciekawostki badawczej w niedługim czasie stała się elementem studyjnego rzemiosła (Mitsufuji, 2022).

Skuteczne wyodrębnianie ścieżek audio otwiera nowe pola pracy z dźwiękiem. Po separacji można niezależnie udoskonalać każdą ścieżkę. Można je także usuwać. Z klasycznych wykonań orkiestrowych można bezstratnie usunąć niepożądane chrząknięcia lub potknięcia pojedynczych instrumentów. Rozpoczyna się nowa era w restauracji dźwiękowych archiwów – można odszumiać i rekonstruować nagrania wcześniej uznane za bezpowrotnie stracone (Michelsanti, 2021; Nogales, 2023). To wszystko wiąże się z istotną zmianą paradygmatu: zamiast dotychczasowego „odfiltrowywania” sygnału, modele generatywne (np. dyfuzyjne) rekonstruują brakujące

informacje dźwiękowe w sposób wiarygodny percepcyjnie (Moliner, 2024). Proponuję wyjaśnienie oparte na metaforze: dotąd można było czyścić brudne nagrania, jak czyszczenie strun w gitarze. Teraz, z głębokim uczeniem, można rekonstruować uszkodzone nagrania, jak odbudowanie częściowo spalonej gitary. To właśnie te metody doprowadziły do powstania nowego singla zespołu The Beatles „Now and Then” (2023). Izolacja partii głosu Johna Lennona z kasetowego demo umożliwiła profesjonalną współczesną produkcję dźwięku całego zespołu (Höpflinger, 2025).

3.2.6.4. Muzyczne AI - perspektywa agentowa

Rozważania nad memetyczną ewolucją sprawczości w najnowszych systemach muzycznych prowadzą w kierunku obserwacji rosnącej autonomii AI. To aktualnie pole inwestycji rzędu setek miliardów dolarów, które pobudza także różnorodne badania naukowe (Nestor, 2025). Różne branżowe przewidywania dotyczące AI w 2025 kładły nacisk na rozwój tzw. niezależnych agentów AI (Pillay, 2025; Paul, 2024, Papandreou, 2025). Miały to być takie systemy AI, które po otrzymaniu zadania od użytkownika podejmują własne wieloetapowe działania, aby owe zadania spełnić. Ponadto, potrafią w tym celu używać zewnętrznych narzędzi (Krishnan, 2025). Przykładem takiego zadania mogłoby być polecenie: „stwórz i opublikuj na platformie streamingowej nową playlistę z muzyką w estetyce *vaporwave*, inspirowaną twórczością Lecha Rocha Pawlaka, wraz z opisem i okładką graficzną”, które agent realizowałby samodzielnie, planując kolejne etapy – od analizy twórczości kompozytora, przez generowanie utworów (używając zewnętrznych narzędzi, czyli programów do tworzenia muzyki), aż po przygotowanie materiałów promocyjnych. Warto tutaj odnotować, że współczesne, publicznie dostępne komercyjne systemy potrafiłyby tego typu zadania samodzielnie wykonać jedynie w ograniczonym zakresie. O agencjalnej AI mówił również Jensen Huang, główny dyrektor NVIDIA, jednej z najważniejszych firm na globalnym rynku AI. Jego zdaniem:

Przez ostatnie dwa, trzy lata dokonał się znaczny przełom. Fundamentalny przeskok w sztucznej inteligencji. Nazywamy go agencjalnym AI. Agencjalne AI oznacza takie AI, które posiada agencjalność (sprawczość). Potrafi postrzegać i rozumie kontekst, w którym się znajduje. Potrafi także wnioskować, i co bardzo ważne, potrafi wnioskować na temat tego, jak odpowiedzieć, albo jak rozwiązać problem. Potrafi także zaplanować oraz podjąć działanie. Potrafi używać narzędzi, ponieważ rozumie multimodalne informacje. Potrafi otworzyć stronę internetową, przeanalizować jej format, słowa, filmiki, nawet obejrzeć filmik, wyciągnąć z tego wnioski i powrócić, aby wykorzystać tę nową wiedzę do realizacji zadania. Agencjalne AI. U podstaw agencjalnego AI leży coś absolutnie nowego - wnioskowanie.

The last several years, the last couple, two, three years, major breakthrough happened. Fundamental advance in artificial intelligence. We call it agentic AI. Agentic AI basically means that you have an AI that

has agency. It can perceive, and understand the context of the circumstance. It can reason, very importantly, it can reason about how to answer, or how to solve a problem, and it can plan an action, it can plan, and take action. It can use tools, because it now understands multimodality information, it can go to a website, and look at the format of the website, words, and videos, maybe even play a video, learns from what it learns from that website, understands it, and come back, and use that information, use that newfound knowledge to do its job. Agentic AI. At the foundation of agentic AI, of course, something that's very new, reasoning (Huang, 2025).

Huang określa zatem agencjalne AI jako takie, które postrzega otoczenie, potrafi używać zewnętrznych narzędzi oraz wnioskować z multimodalnych informacji, aby zrealizować określone złożone zadanie. Taka definicja wymaga pewnego komentarza. Po pierwsze warto zaznaczyć, że odnosi się ona do takich modeli jak np. OpenAI o3, który łączy multimodalność, zdolność korzystania z narzędzi oraz umiejętność wnioskowania (OpenAI, 2025). Poniżej wyjaśnię te trzy funkcje.

Warunek postrzegania otoczenia wiąże blisko z pojęciem multimodalności. Multimodalność oznacza, że model na wejściu posługuje się nie tylko tekstem, ale także danymi w innych formatach. W tym przede wszystkim grafiką, a także dźwiękiem. W praktyce należy tu zaznaczyć, że w porównaniu do analiz graficznych, które skutecznie przetwarzają pliki PDF, funkcja analiz audio jest bardzo ograniczona. Obecnie (wrzesień 2025) w praktyce analizy audio – to znaczy treści mówionego tekstu – spośród multimodalnych modeli przeprowadza jedynie Google Gemini 2.5 (Masood, 2025). Z własnych podejść obserwuję, że potrafi on także w pewnym elementarnym stopniu i z niewielką skutecznością analizować utwory muzyczne.

Korzystanie z zewnętrznych narzędzi oznacza, że dany model potrafi zidentyfikować i zrealizować potrzebę użycia zewnętrznej aplikacji do ukończenia określonego zadania (Qin, 2023). Szczególnie popularnym i użytecznym narzędziem jest przeglądarka internetowa używana w celu zdobycia najświeższych informacji. Tym niemniej narzędziem może być dowolna aplikacja, z którą użytecznie połączy się LLM. Firma Anthropic rozwija uniwersalny protokół MCP, aby jednolicić protokół łączenia AI z różnymi aplikacjami (Hou, 2025). To rozwiązanie zostało już zaimplementowane do

połączenia AI Claude z niezmiernie popularnymi programami DAW (Digital Audio Workstation) do produkcji muzyki – FL Studio (Veena, 2025) oraz Ableton (Ahuja, 2025). W praktyce oznacza to, że używając promptów w języku naturalnym użytkownik może sterować zaawansowanym DAWem, a ponadto, że może to robić także sam zaledwie nadzorowany lub potencjalnie nienadzorowany agent AI. To jeden ze sposobów na to, żeby muzyka komponowana i produkowana przez systemy AI unikała homogenicznego, popularnego obecnie charakteru brzmieniowego (Afchar, 2025). To także po prostu nowy sposób tworzenia muzyki, który określam jako *vibe-composing*, analogicznie do nowopowstałego *vibe-codingu* w świecie programowania (Gadde, 2025). To zjawisko prezentuję bliżej w czwartej części tej rozprawy.

Wracając do cech agencjalnego AI wymienianych przez Huangę, oprócz multimodalności i używania narzędzi trzeba opisać także wnioskowanie. Rozumowanie w dużych modelach językowych to zdolność do rozwiązywania złożonych problemów krok po kroku. Model przeprowadza pośrednie etapy rozumowania, dzięki czemu znacznie lepiej wykonuje skomplikowane zadania (Wei, 2022). Dzięki temu na przykład jeśli spytamy model o bieżącą dostępność karnetów na festiwal muzyczny, będzie mógł najpierw zrozumieć, że nie zna odpowiedzi, później postanowić, że poszuka jej w internecie, następnie uruchomić przeglądarkę, znaleźć informację a na koniec przeformułować ją do stylu zgodnego z zapytaniem. Albo kiedy użytkownik prosi o zaprogramowanie interaktywnej muzycznej aplikacji, wnioskujący system może wykorzystać terminal, aby na bieżąco testować kod i korygować własne błędy. Definicja Huangę to doskonałe wprowadzenie do znacznie bardziej złożonego problemu AI i agencjalności. Samo agencjalne AI w bieżącej literaturze naukowej określane jest także jako:

niezależne, nastawione na osiągnięcie celu systemy, które mogą działać same przez długi czas, wymagając minimalnego ludzkiego nadzoru.

autonomous, goal-driven systems that can operate on their own for long periods, requiring minimal human supervision (Bandi, 2025).

Ta quasi-definicja opiera się na innych cechach tych samych systemów. W szczególności podkreśla samodzielne działanie przez **długi** czas (nie określając tej długości) przy **minimalnym** ludzkim nadzorze (również nie określając stopnia tego nadzoru). To po prostu robocze definicje w błyskawicznie rozwijającym się polu naukowym i technologicznym.

Poza AI agencjalną funkcjonuje również termin *agent AI*, stosowany chociażby przez dyrektora OpenAI Sama Altmana, który spodziewał się, że ‘w 2025 pierwsi agenci AI wejdą na rynek pracy’ (Altman, 2025). Jego zapowiedź sugerowała nadejście systemów inteligentnych i kompetentnie wykonujących różnorodne przydzielone zadania, tak, aby zastąpić ludzkich pracowników. To bardzo wąskie i zaawansowane rozumienie agenta AI. W klasycznej już pozycji ‘Sztuczna Inteligencja: Nowe spojrzenie’ autorzy przypisują pojęciu agenta duże znaczenie (Russel, 2023). Ich zdaniem ‘podejście „racjonalnego agenta” zdominowało historię rozwoju sztucznej inteligencji’ (Russel, 2023). Nie definiują oni agenta AI, ale za to piszą, że od „agenta komputerowego” oczekuje się:

Samodzielnego działania, postrzegania swego otoczenia, funkcjonowania przez dłuższy czas, adaptowania się do zachodzących zmian, wytyczania celów i dążenia do ich osiągnięcia (Russel, 2023 przeł. Andrzej Grażyński).

Taka definicja komputerowego agenta łączy obie wcześniej przytoczone quasi-definicje agencjalnego AI, ale jeszcze je zaostrza, mówiąc o samodzielnym działaniu. Uznaję zatem, że agenta AI od agencjalnej AI odróżnia właśnie to kryterium samodzielności. Ponadto Russel i Norvig określają całe pole sztucznej inteligencji wokół racjonalnych agentów, czyli bytów, które działają:

Reasumując, *sztuczna inteligencja skupiła się na budowaniu agentów **postępujących właściwie***, przy czym kryterium „właściwości” postępowania określone jest każdorazowo przez cel, którego osiągnięcie powierzamy agentowi. Ów ogólny paradygmat jest tak powszechny, że śmiało możemy go nazwać modelem standardowym. –Russel, 2023, przeł. Andrzej Grażyński

Powyższe kryteria agentowości są bardzo szeroko zakrojone. W minimalnym stopniu spełnia je na przykład termostat, ponieważ jest w stanie postrzegać temperaturę otoczenia i długofalowo dostosowywać do tych pomiarów swoje zachowanie. Czyli zachodzi skuteczna pętla zwrotna między postrzeganiem środowiska i wpływaniem na nie. Te same minimalne kryteria spełniają także automatyczne stroiki do gitary (Alexander, 2014), czyli urządzenia elektroniczne montowane na główkę gitary, które monitorują czystość stroju i dostosowują ją na bieżąco w razie potrzeby. Automatyczny stroik jest więc autonomicznym muzycznym agentem komputerowym.

Jednak w latach 20. XXI wieku autonomiczne systemy komputerowych agentów potrafią o wiele więcej. Przede wszystkim są wyposażone w zaawansowaną sztuczną inteligencję. Takim w pełni autonomicznym agentem może być opisywany we wcześniejszym rozdziale system rekomendacyjny. Poniżej rozważę, czy opisany wcześniej system rekomendacyjny Spotify spełnia opisane przez Huanga kryteria agencjalnej AI oraz opisane przez Russela i Norviga kryteria agenta komputerowego. W tej perspektywie, na początek, przede wszystkim dostrzegam potrzebę dodatkowego skomentowania koncepcji agencjalnego AI. Huangowskie agencjalne AI to przede wszystkim bardzo konkretny produkt – stworzony w określonym czasie do określonych zadań. Co więcej, tym zadaniem nie jest masowa rekomendacja muzyki w globalnym systemie, a raczej bycie doraźnym narzędziem-pomocnikiem biurowym. Tym bardziej zatem taki kształt agencjalnego AI nie może być w tej pracy potraktowany po prostu jako jedyne naturalne przedłużenie ewoluującej agencjalności memów i mempleksów. Agencjalność rozwijana na drodze memetycznej ewolucji wiąże się z różnymi kompetencjami, sposobami manipulacji otoczeniem w taki sposób, aby bardziej zwiększało ono szanse replikacji agencjalnego memu. Celowo zaprojektowane przez człowieka cechy agencjalnej AI jak najbardziej wpisują się w ten złożony, szeroko zakrojony proces koewolucji biologiczno-memetycznej, ale go nie wyczerpują. Zatem porównanie cech i ‘zachowania’ rekomendera Spotify do agencjalnego AI daje pewne, ale nie pełne wskazanie na temat memetycznej agencjalności tego systemu. Tym niemniej znaczące podobieństwa i paralele, a także jedną istotą różnicę, należy odnotować.

Pierwszy warunek – postrzeganie kontekstu i otoczenia, spełnia się całkowicie. Rekomender w sercu Spotify nieustannie przetwarza ogromny potok informacji. To po pierwsze stale poszerzająca i zmieniająca się baza samych piosenek oraz metainformacji na ich temat (Chopade, 2024). Poza tym utwory muzyczne posiadają dane takie jak tytuł, album, wykonawca, czas trwania. Z każdym utworem nierozzerwalnie wiąże się także jego recepcja: ile zebrał odsłuchań w jakim czasie, ile razy go pominięto, do ilu i jakich dodano go playlist, ile zebrał polubień. Z drugiej strony są informacje o użytkownikach, takie jak to, ile i czego dana osoba słucha, kiedy czy też jak długo trwają sesje słuchania (Fouad, 2025). Te informacje nieustannie się zmieniają, ponieważ słuchacze bez przerwy korzystają z aplikacji.

Multimodalność, drugi kluczowy warunek, także zostaje spełniony. Rekomender Spotify przetwarza zarówno dane muzyczne jak i tekstowe, co stanowi dwie modalności (Saravanou, 2021). System rekomendacji MUSIG w Spotify prowadzi złożone wnioskowanie z grafów opartych na danych tekstowych i muzycznych (jw). Co prawda pełna multimodalność zakłada użycie tekstu, obrazu i dźwięku, ale jak wskazuje opisana wcześniej praktyka, użycie dwóch z tych trzech modalności wystarcza jako warunek minimalny multimodalności.

Co więcej, opisany wcześniej system MUSIG operujący na danych z różnych modalności można potraktować także jako centrum wieloetapowego wnioskowania. W ogóle wnioskowanie w systemie rekomendacyjnym odbywa się w inny sposób niż typowym dla agencjalnego AI LLMie z interfejsem chatbota. LLM generalnie otrzymuje poszczególne, dyskretne, odrębne zadania o jasnym momencie początku i końca. System rekomendacji, szczególnie w tak ogromnych i popularnych platformach jak Spotify działa bez przerwy, z każdym krokiem aktualizuje swoją wiedzę i dokonuje coraz bardziej racjonalnych wyborów. Choć taki pojedynczy wybór, jak zarekomendowanie jednemu użytkownikowi jednej piosenki, nie jest tak wieloetapowo złożony, jak pojedyncze zapytanie w LLMie, to jednak fundamentalna funkcja rekomendera jest zupełnie inna. Jeżeli jako jego złożone zadanie opiszę utrzymanie użytkownika na

platformie, to znaczy stabilne zwiększenie konsumpcji muzyki, to widać, że rekomender działa powoli, stabilnie, systematycznie, wieloetapowo. Występuje tu pewna niewspółmierność, ale na pewno nie przekreślałbym agencjalności rekomendera Spotify ze względu na jego umiejętności wnioskowania.

Huang zaledwie to sugeruje implicite, ale Bandi (2023) wprost od agencjalnej AI wymaga długofalowego autonomicznego działania. To także ważny wymóg dla agentów komputerowych w rozumieniu Russela i Norviga (Russel, 2023). Istnienie i realizację długofalowej strategii w systemie rekomendacji Spotify potwierdza wprost raport techniczny z 2023 (McDonald, 2023). System nie jest nastawiony na optymalizację jednej rekomendacji tu i teraz. Rozumuje w szerszym kontekście, tworząc racjonalne ciągi rekomendacji, celowo dawkuje bodźce, aby utrzymywać długie zaangażowanie i kreować nawyki konsumpcji (McDonald, 2023). Takie zachowanie systemu pozytywnie realizuje również wymóg Russela i Norviga odnośnie do wytyczania celów, adaptowania się do zmian i osiągania tychże celów. Rekomender Spotify nie tylko postrzega i planuje, ale nieustannie wysyła słuchaczowi informacje, ustala jego playlisty, jego nawyki słuchania, oraz usprawnia swoje strategie, jeśli okazują się nieefektywne.

Ostatni aspekt agencjalności według Huanga to korzystanie z zewnętrznych narzędzi. Tutaj należy odnotować największą różnicę, ponieważ w rekomenderze Spotify taka funkcja nie istnieje. Za to system Spotify nieustannie wykorzystuje przeróżne wewnętrzne narzędzia do badania swojego wewnętrznego świata, czyli baz danych, które nieustannie rosną. Zatem ten rekomender nie używa zewnętrznych narzędzi, ale po prostu nie ma takiej potrzeby, szczególnie że i tak nadażenie za rozwojem jego wewnętrznych danych to trudne zadanie.

Powyższe informacje, o tak wysokiej agencjalności rekomendera Spotify, nawet według najnowszych branżowych kryteriów ze świata AI, ukazują prawdziwe owoce ewolucji memetycznej w świecie muzyki. Rekomender Spotify to autonomiczny cybernetyczny byt, który steruje globalną konsumpcją muzyki. W 2020 według Amerykańskiego Towarzystwa Przemysłu Nagraniowego (RIAA) dochody z platform streamingowych

stanowiły 83% globalnych dochodów branży muzycznej w ogóle (Friedlander, 2020). Czyli streaming absolutnie zdominował globalną branżę muzyczną. A w sercu tego streamingu stoją coraz bardziej zaawansowane systemy rekomendacji oparte na AI, w tym na procesach ewolucyjnych. Dokładnie przewidywanym przez teorię memetycznej ewolucji rezultatem jest to, że wytwory tejże ewolucji prowokują człowieka, aby owe muzyczne memy replikował. W tym także zawiera się ich sprawczość, którą na drodze ewolucji nabywają, pomimo tego, że same nie są biologicznie żywe. W tym miejscu stawiam ważne pytanie: czy człowiek ma wolną rękę, czy prezes Spotify może z własnej woli zrezygnować z używania coraz mniej zrozumiałych i coraz bardziej autonomicznych systemów AI do rekomendacji muzycznych? Moja odpowiedź nie jest żadną spekulacją, konstruktem teoretycznym, domysłem. Spotify złożyło w roku fiskalnym 2023 formalny techniczny raport o czynnikach ryzyka do United States Securities and Exchange Commission, w którym napisano wprost:

Jeżeli nie uda nam się trafnie przewidzieć, zarekomendować i odtworzyć treści, które podobają się naszym użytkownikom, możemy stracić obecnych użytkowników i stracić zdolność przyciągania wystarczającej liczby nowych użytkowników, aby spełnić oczekiwania inwestorów odnośnie wzrostu, a także możemy stracić możliwość rentownego prowadzenia działalności gospodarczej.

If we fail to accurately predict, recommend, and play content that our users enjoy, we may fail to retain existing users and attract new users in sufficient numbers to meet investor expectations for growth or to operate our business profitably (Spotify Technology, 2024).

Spotify samo potwierdza, że po prostu nie może zrezygnować ze swoich agencjalnych rekomenderów pod groźbą upadku firmy. Tutaj przypomina się problem smyczy Wilsona, opisany w części drugiej (Wilson, 1978). Jeżeli znajdzie się w firmie osoba o dostatecznej pozycji, aby wdrożyć decyzję zaniechania używania rekomenderów, zapewne w niedługim czasie zostanie ona z tej pozycji zdegradowana, a zastąpi ją ktoś inny, kto taką decyzję odwoła. Tak działa wyewoluowana logika kompetencji w firmie, która jest systemem wiabilnym, dynamicznym, adaptacyjnym. To przykład tego, że w korporacyjnym świecie przemysłu muzycznego w tej chwili nie ma odwrotu od adaptowania agencyjnych, autonomicznych technologii AI. Racjonalne za to jest

oczekiwanie zmian w drugą stronę: coraz głębszej i szerszej adaptacji AI. Według raportu Organizacji Współpracy Gospodarczej i Rozwoju (OECD), tempo takich adaptacji potrafi być powolne, ale ich kierunek jest stabilny i wydaje się nieodwracalny (Calvino, 2024). Firmy skutecznie adaptujące rozwiązania AI stają się bardziej konkurencyjne.

W innym raporcie technicznym Spotify otrzymało wyraźne wskazanie, że czynnikiem ryzyka dla prowadzenia działalności firmy są drogie umowy licencyjne z wielkimi muzycznymi wytwórniami, a rozwiązaniem jest wytworzenie własnego repertuaru (Williams, 2021). Racjonalną polityką firmy zatem byłoby wykorzystanie generatywnej AI, aby wytworzyć ów własny repertuar i nie polegać na żadnych zewnętrznych umowach licencyjnych. Źródła dziennikarskie (Gomez Sarmiento, 2025; Hypebot, 2025) potwierdzają, że na muzykę generowaną przez AI natrafia się na Spotify coraz częściej. W 2025 pojawił się także bezprecedensowo wiralny zespół The Velvet Sundown, który zebrał miliony odsłuchań, zanim autor Andrew Frelon ujawnił, że sam stworzył całą muzykę w Suno AI (Frelon, 2025). Muzyka The Velvet Sundown była promowana przez algorytm Spotify, ale nie ma dowodów na to, że powstała na zlecenie tej korporacji.

Zatem Spotify zezwala na publikację muzyki generowanej całkowicie przez AI (na podstawie promptów), a także w niektórych przypadkach aktywnie ją promuje. Na tym nie koniec eksperymentów firmy z generatywną AI. Spotify już w 2022 kupił firmę Sonantic, zajmującą się właśnie generowaniem dźwięku przy pomocy AI. Bazując na tej technologii w 2023 Spotify wprowadził funkcję AI DJ (Spotify, 2023). Firma chwali się, że stanowi to szczytowe osiągnięcie w personalizacji, która jest 'sercem wszystkiego, co robią' (Spotify, 2023). A zatem Spotify przechodzi już z paradygmatu wspieranej przez AI rekomendacji treści audio tworzonych przez człowieka, na paradygmat hybrydowego systemu, w którym AI jest zarówno rekomenderem jak i współtwórcą słuchanego audio. Co prawda jak dotąd AI DJ w kwestii generacji audio zajmuje się tylko wypowiedziami prowadzącego, ale jest to już początek wdrożenia pionierskiej technologii: hiperpersonalizowanego osobistego radia, które łączy technologie AI zarówno do rekomendacji jak i generacji treści.

Taką technologię Spotify wdrożył już w 2023 i zgodnie z działaniem jego systemów, od tamtej pory nieustannie zbiera dane dotyczące jej funkcjonowania, recepcji. Przy tym wszystkim zarobki dla artystów są zaskakująco niskie. Żeby utwór w ogóle mógł generować przychód, musi przekroczyć liczbę tysiąca odsłuchań w ciągu roku. W 2023 tego warunku nie spełniło 159 milionów utworów z całkowitej puli 184 milionów (Tencer, 2024). Czyli 86% utworów przyniosło artystom dosłownie zerowe zyski. A zatem podczas gdy AI przejmuje kolejne pola działania, artyści stają się coraz bardziej bezradni. To fundamentalna zmiana w funkcjonowaniu, a także ewolucji kultury muzycznej. Po raz pierwszy w historii to inteligentne maszyny decydują o kierunkach dystrybucji muzyki, jak również w coraz większym stopniu ją tworzą. Ale warto spojrzeć jeszcze szerzej. W artykule z 2023 badacze Brinkmann i jego współpracownicy zaproponowali hasło *kultura maszynowa* (Machine culture) (Brinkmann, 2023), które opisuje tę bezprecedensową sytuację nie tylko na polu muzyki, ale również innych form kultury. Zauważają oni przede wszystkim, że obecnie inteligentne maszyny mają znaczący wpływ na wszystkie trzy główne aspekty ewolucji kultury – zmienność, transmisję i selekcję (Brinkmann, 2023). W artykule nie pojawia się pojęcie memu, ale opisywana koncepcja mierzalnej ewolucji konkretnych przedmiotów kultury jest z memetyką całkowicie spójna.

Stosując koncepcję kultury maszynowej do świata muzyki, mogę szerzej rozważyć podawane przeze mnie w poprzednich rozdziałach przykłady. Generatywna AI z powodzeniem wytwarza muzykę: historie zespołów The Velvet Sundown i AISIS pokazują znaczący wkład do muzycznej puli memetycznej. Ponadto, modele dyfuzyjne jak Riffusion generują absolutnie niespotykane dotąd struktury, czyli autentyczną nowość. Procesy transmisji i selekcji przeprowadzają systemy rekomendacji, jak opisywany system Spotify.

Autorzy koncepcji kultury maszynowej proponują wiele ciekawych wniosków, jak to, że maszyny są zarówno produktami kultury, jak i jej twórcami (Brinkmann, 2023), oraz to, że kultury ludzi i maszyn już w złożony sposób się przenikają, a w przyszłości rola

maszyn będzie tylko wzrastać (Brinkmann, 2023). Za to ci autorzy w ogóle nie odwołują się do pojęcia replikatora, które moim zdaniem ma tutaj kolosalne znaczenie.

Zwracam uwagę na to, że mogą już istnieć całkowicie zamknięte dla ludzi nisze ewolucji kultury muzycznej – AI wytwarza muzykę, następnie AI ją odsłuchuje i rekomenduje, stwarzając informację zwrotną do kolejnego pokolenia wytworzonych utworów. Oczywiście w praktyce trudno je ukazać, skoro taka nisza ma w ogóle nie dopuszczać człowieka. Ale istnieją już przykłady masowego wytwarzania muzyki przez AI, która następnie doskonale odnalazła się w systemie rekomendacji przy minimalnym udziale człowieka. Szwed Johan Röhr prawdopodobnie z pomocą AI wykreował na Spotify aż 650 artystów, których 2700 piosenek zebrało 15 miliardów odsłon (Bryant, 2024). Co ważne, Röhr przy prawdopodobnej pomocy AI celowo wytwarzał taką muzykę, która miała precyzyjnie spełnić kryteria selekcji algorytmicznej. Od lat jest znane także zjawisko autonomicznych maszyn odsłuchujących muzykę w streamingu, również na Spotify (Drott, 2020).

A zatem w środowisku maszyn powstają autentycznie nowe warianty informacyjne, następnie podlegają one specyficznym dla maszyn presjom selekcyjnym, które wpływają na to, na ile skutecznie będą wyznaczać one kreację kolejnych wariantów po kolejnym treningu AI na podstawie danych z internetu. Wnioskuje więc, że należy wyróżnić nowy wyłaniający się replikator. Jego istnienie prowokują głównie LLMy, zatem nazywam go lemem (również na cześć polskiego pisarza, autora fantastyki naukowej Stanisława Lema, który zgłębiał koncepcję mechanicznej ewolucji choćby w książce *Niezwyciężony*).

Dawkins wyróżnił memy jako replikator odrębny od genów (1976) właśnie dlatego, że zauważył informacje podlegające ewolucji poza mechanizmem ewolucji genetycznej. Jeżeli jednak głównym mechanizmem ewolucji memetycznej ma być prowokowana przez ludzi imitacja, a memy replikują się z mózgu do mózgu, informacje kulturowe wytwarzane, przetwarzane i ewoluowane przez AI należy uznać za kolejną odrębną grupę. Odrębny mechanizm ewolucji nie oznacza absolutnej niezależności – tak jak

memy replikują się w środowisku genetycznym, tak lemy replikują się w środowisku memetycznym – algorytmicznym. Czyli kultura przypięta genom na Wilsonowskiej smycz posiada własną smycz, i trzyma na niej lemy.

Wielką istotność w tej dynamice dwóch smyczy odgrywa koewolucja. Autorzy kultury maszynowej podkreślają empiryczne obserwacje tego, że wynalezione przez Alpha Go nowatorskie strategie gry w GO zaczęły istotnie wpływać na nowopowstające strategie ludzkie (Brinkmann, 2023). Analogicznie, należy obserwować, jak strategie muzyki komponowanej przez AI przenikają do nowych, ludzkich, czy też hybrydycznych, stylów. W sieci już możemy znaleźć poradniki jak tworzyć taką muzykę, która będzie najbardziej odpowiednia dla systemów rekomendacji Spotify (Anderson, 2023; McGlynn, 2023; Knibbe, 2020). To silne empiryczne wskazanie tego, że kryteria selekcyjne AI przenikają już także do twórczości ludzkiej.

Ponieważ nie istnieje jeszcze żadne absolutnie autonomiczne AI, a trenowanie modeli nadal przebiega przede wszystkim na danych dotyczących ludzkiej kultury, ‘czyste’ lemy uznaję za byty nadal hipotetyczne. Tym niemniej można zaobserwować, że mechanizmy autentycznie poza-ludzkiej ewolucji w środowisku kultury muzycznej nieustannie wzrastają na znaczeniu. Choć branża streamingowa stanowi znaczną większość rynku muzycznego (Friedlander, 2020), a firma Spotify i jej systemy AI są w doskonałej kondycji, badania wskazują, że muzycy nie są w tym ekosystemie ‘sprawiedliwie kompensowani’ (Castle, 2021) oraz obawiają się o swój przyszły los w świecie wypełnionym technologiami AI (Ma, 2025; Iskandar, 2025). To idealny przykład obecnego stanu konkurencji o zasoby materialne zachodzącej między genami i memami. Autonomiczne, powstałe na skutek ewolucji memetycznej systemy kulturowe takie jak byty korporacyjne czy autonomiczne systemy AI zawłaszczają coraz większą część zasobów na rynku muzycznym, marginalizując rolę poszczególnych, niezależnych artystów – organizmów żywych zabiegających o byt.

W sferze muzyki, pomimo koewolucyjnych zależności pomiędzy genami i memami, a także wyłaniającym się, niezależnym AI, można nakreślić wyraźną oś konkurencji.

Szerszą perspektywę takich konkurujących procesów ewolucyjnych nakreślił w 2023 Dan Hendrycks w raporcie dla Center for AI Safety zatytułowanym *Natural Selection Favors AIs over Humans* (Hendrycks, 2023). Badacz w duchu uniwersalnego darwinizmu podkreśla, że każda ewolucja na drodze doboru naturalnego produkuje 'samolubne byty' (Hendrycks, 2023). Opisuje on także różnorodne mechanizmy kreatywnego stawiania własnej replikacji za priorytet przez systemy AI, również w sposób niejawny (Hendrycks, 2023). Badacz uważa także, że ewolucja AI przebiega znacznie szybciej, niż ludzka (Hendrycks, 2023).

Roman Yampolskiy, twórca badań nad bezpieczeństwem AI jako subdyscypliny naukowej (Wilson, 2020) opisał wiele scenariuszy zagrożenia dla ludzi ze strony AI. Jednym z nich jest AI manipulujące ludźmi poprzez celową produkcję samolubnych memów (Yampolskiy, 2024). Badacz zauważa, że:

Korporacje to samolubne byty, których celem jest maksymalizacja zysku udziałowców. Kiedy AI przewodzi takiemu bytowi, każda idea może zostać skodyfikowana w algorytmie i dodana do polityki korporacji. Na wyższym poziomie abstrakcji tak mogą powstawać samolubne kryptowaluty.

Corporations are selfish entities with the goal of maximizing shareholder profit, with AI in charge of such an entity, any idea can be codified in an algorithm and added as the driving force behind the corporation's decision-making. At a higher level of abstraction, this could produce selfish cryptocurrencies. - (Yampolskiy, 2024)

Yampolskiy nie hipotetyzuje na temat jakiegoś przyszłego, potencjalnego AI obdarzonego złożoną, antropocentrycznie rozumianą intencjonalnością czy świadomością. Mówi o dzisiejszym systemie racjonalnie nawigującym samolubną korporacją na konkurencyjnym rynku. Systemy AI pełnią już zarządcze role w wielu korporacjach, a chińska firma NetDragon od 2022 szczeni się tym, że przewodzi jej prezes AI (Odilov, 2024). Za to w 2024 pierwszy autonomicznie działający model AI, Truth Terminal, poprzez tworzenie atrakcyjnych memów zbudował wartość własnej kryptowaluty i został milionerem (Khalili, 2024).

Nic nie stoi na przeszkodzie, żeby podobne autonomiczne systemy zaczęły działać na polu muzyki. Spotify i tak już selekcjonuje oraz współtworzy memy muzyczne przy pomocy AI. Znany jest również przykład dobrze prosperującej memetyczno-muzycznej kryptowaluty \$keycat (Schmidt, 2025). Perspektywa Levinowskiej Wieloskalowej Architektury Kompetencji sugeruje, że mniej złożone systemy mogą łączyć się w bardziej złożone struktury dzięki wyewoluowanej odgórnej strukturze zarządzania. Współczesny cyfrowy ekosystem muzyczny już integruje w sobie zarówno memetycznie wyewoluowane treści, jak i złożone systemy AI, które nieustannie rozwijają się dzięki algorytmom ewolucyjnym. Jeżeli systemy AI będą dalej ewoluować podobnie jak dotychczas, niedługo może się okazać, że AI po prostu jest skuteczniejsze niż człowiek na pozycji prezesa. Wtedy także i Spotify, jako system wiabilny, postąpi racjonalnie i wybierze takiego przewodniczącego. Sprawcze, agencjalne, wyewoluowane memetycznie systemy muzyczne już istnieją i postępująco dążą w kierunku całkowitej autonomii.

4.1. Semantyka dystrybucyjna

W poszukiwaniu praktycznej metody badania wspomnianej w części 2 niniejszej pracy semantyki dystrybucyjnej muzyki natrafiłem na inną interesującą możliwość. Otóż semantyka dystrybucyjna języka jest badana już na tyle szeroko, że istnieją darmowe bazy danych i elementarne narzędzia do wykonywania badań ilościowych (Hamilton, 2016). Ponieważ jednak takie badania językoznawczo-informatyczne leżą zdecydowanie poza obszarem muzykologii, zdecydowałem się sprawdzić i opisać tylko jedną elementarną rzecz: czy dzięki wyżej wspomnianym dostępnym narzędziom będę w stanie uzyskać elementarne informacje na temat sieci semantycznej muzyki?

Z pomocą AI udało mi się zaimplementować algorytm wyszukujący 10 najbliższych sąsiadów w wektorowej przestrzeni semantycznej, czyli 10 słów ocenionych przez algorytm jako najbliższych semantycznie wybranemu słowu z uwagi na częstość i bliskość współwystępowania w tekstach językowych. Baza danych oraz metodyka badaczy ze Stanforda (Hamilton, 2016) umożliwiają dodatkowo diachroniczne badania przestrzeni semantycznej. To znaczy, mogę określić najbliższych semantycznych sąsiadów dla każdej dekady od 1800 do 1990. Za swój cel obieram słowo „music” czyli muzyka w języku angielskim, ponieważ dostępny korpus zawiera teksty w języku angielskim. Takie badanie nie tylko ukaże dla każdej dekady te słowa, które z muzyką najczęściej i najbliżej współwystępowały w tekstach językowych. Zgodnie z hipotezą dystrybucyjną odpowiednio bogaty korpus odzwierciedla również przy pomocy sąsiadów realne znaczenie danego słowa (Firth, 1957).

A zatem taki system umożliwi wgląd w zmiany pola semantycznego słowa music na przestrzeni dziejów, czyli w moim rozumieniu jego memetyczną ewolucję. Śledząc zmiany zachodzące z dekady na dekadę, ukazuje się droga, którą pokonuje słowo muzyka w diachronicznej przestrzeni semantycznej.

Prezentuję zatem listę 10 najbliższych semantycznie słów do słowa music na podstawie korpusu ze Stanforda i algorytmu *nearest neighbor*. Każde słowo posiada wartość numeryczną, która zgodnie z hipotezą dystrybucyjną określa miarę semantycznego podobieństwa danego słowa do słowa 'music' w spektrum od całkowitej zgodności: 1.00 do absolutnej rozbieżności: 0.00. Surowe wyniki działania algorytmu są następujące:

Dekada 1800:

musick: 0.68
vocal: 0.62
dancing: 0.59
musical: 0.59
dances: 0.56
tune: 0.53
sculpture: 0.53
singing: 0.51
poetry: 0.50
melody: 0.50

Dekada 1810:

musick: 0.62
musical: 0.50
vocal: 0.50
tunes: 0.50
dancing: 0.49
flute: 0.45
violin: 0.45
poetry: 0.45
lyre: 0.45
singing: 0.44

Dekada 1820:

musick: 0.55
dancing: 0.51
vocal: 0.49
singing: 0.49
guitar: 0.48

lyre: 0.47
accompaniment: 0.47
musical: 0.45
violin: 0.45
lute: 0.44

Dekada 1830:

musical: 0.52
dancing: 0.51
singing: 0.46
vocal: 0.45
harpsichord: 0.45
melody: 0.45
poetry: 0.44
painting: 0.42
piano: 0.42
mozart: 0.42

Dekada 1840:

musical: 0.55
melody: 0.54
singing: 0.49
dancing: 0.49
songs: 0.48
vocal: 0.45
concerts: 0.44
tune: 0.43
poetry: 0.43
lute: 0.43

Dekada 1850:

musical: 0.55
flutes: 0.52
choral: 0.50
melody: 0.50
modulation: 0.49
dancing: 0.48

fifes: 0.47
flute: 0.47
singing: 0.46
strains: 0.46

Dekada 1860:

melody: 0.52
dancing: 0.51
musical: 0.48
singing: 0.48
cadence: 0.46
song: 0.45
flutes: 0.44
tune: 0.44
lute: 0.43
strains: 0.43

Dekada 1870:

musical: 0.57
dancing: 0.53
singing: 0.53
lute: 0.49
melody: 0.49
poetry: 0.47
song: 0.46
flute: 0.45
singer: 0.44
sounds: 0.44

Dekada 1880:

musical: 0.53
dancing: 0.53
singing: 0.52
melody: 0.50
harping: 0.48
ditties: 0.48
strains: 0.47

symphonies: 0.47

fifes: 0.46

orchestra: 0.46

Dekada 1890:

musical: 0.54

melody: 0.52

flute: 0.47

poetry: 0.46

dancing: 0.45

musicians: 0.45

orchestra: 0.44

singing: 0.44

recitations: 0.44

songs: 0.44

Dekada 1900:

musical: 0.53

melody: 0.52

violin: 0.49

poetry: 0.47

singing: 0.46

flute: 0.46

strains: 0.46

dancing: 0.45

orchestra: 0.44

symphonic: 0.44

Dekada 1910:

musical: 0.58

singing: 0.55

melody: 0.49

songs: 0.49

poetry: 0.48

dancing: 0.48

song: 0.47

violin: 0.43

orchestra: 0.43

piano: 0.42

Dekada 1920:

musical: 0.53

poetry: 0.49

singing: 0.48

drama: 0.47

painting: 0.46

dancing: 0.46

symphony: 0.46

violin: 0.45

song: 0.45

sculpture: 0.44

Dekada 1930:

musical: 0.55

singing: 0.50

dancing: 0.47

poetry: 0.45

listening: 0.44

choral: 0.44

violin: 0.44

piano: 0.43

song: 0.43

sculpture: 0.42

Dekada 1940:

musical: 0.54

singing: 0.47

drama: 0.46

dancing: 0.46

art: 0.46

orchestral: 0.44

flute: 0.44

poetry: 0.44

symphony: 0.44

song: 0.44

Dekada 1950:

musical: 0.58

melody: 0.48

art: 0.46

drama: 0.45

singing: 0.45

dancing: 0.45

poetry: 0.45

painting: 0.44

composer: 0.44

arts: 0.44

Dekada 1960:

bmi: 0.55

ascap: 0.54

operatic: 0.52

musical: 0.52

singing: 0.51

dance: 0.50

dancing: 0.49

dramatics: 0.48

choral: 0.48

polyphonic: 0.47

Dekada 1970:

musical: 0.55

polyphonic: 0.49

poetry: 0.49

jazz: 0.48

dramatics: 0.48

singing: 0.46

song: 0.45

songs: 0.45

symphony: 0.44

dance: 0.44

Dekada 1980:

musical: 0.61
dance: 0.55
poetry: 0.50
drama: 0.49
singing: 0.48
jazz: 0.48
opera: 0.47
lyrics: 0.47
dancing: 0.46
songs: 0.46

Dekada 1990:

dance: 0.54
poetry: 0.48
art: 0.46
musical: 0.46
jazz: 0.46
opera: 0.45
singing: 0.45
songs: 0.45
dancing: 0.45
drama: 0.44

Dalsza analiza znaczeń i kontekstów semantycznych tych słów może być podstawą do bardzo szeroko zakrojonych badań. Dość powiedzieć, że dla każdego z tych 10 słów w każdej dekadzie można teraz ponownie ustalić 10 najbliższych sąsiadów, dodając budowanej mapie semantycznej dodatkową głębię. Takie badanie wymagałoby znacznie więcej czasu. Poza tym, dopiero dodatkowe określenie wzajemnych odległości (semantycznych w przestrzeni wektorowej) pomiędzy każdą parą sąsiadów umożliwiłoby zmapowanie poglądowej przestrzeni wektorowej.

Tym niemniej w celu potwierdzenia, że w działaniu algorytmu nie pojawiają się przypadkowe błędy, a także sprawdzenia, czy budowanie szerszej sieci ma sens,

uruchamiam algorytm ponownie, tym razem określając poszukiwania na liczbę 30 najbliższych sąsiadów słowa „music”. Rezultaty nie wyglądają na przypadkowe. Wartości liczbowe zdają się odzwierciedlać realne zmiany w świecie muzyki. Dostrzegam wyraźne przesunięcia akcentów w ciągu dwóch stuleci. W XIX wieku dominowało powiązanie z dawnymi formami i instrumentarium klasycznym. Wskazuje to obecność takich wyrazów jak musick (0,68 w 1800, 0,55 w 1820), harpsichord, lyre czy viol (0,37–0,46), które następnie zanikają po połowie stulecia. Równocześnie music funkcjonowało w szerokim kontekście sztuk wyzwolonych: obok niego pojawiały się sculpture (0,53 w 1800), painting (0,49 w 1800) czy architecture (0,45 w 1800), przy czym w XX wieku ich podobieństwo spada do poziomu 0,40–0,41. Stabilne pozostają elementy rdzeniowe: musical, singing, dancing, poetry, melody czy song/songs. Występują one konsekwentnie w każdym dziesięcioleciu, utrzymując podobieństwo w granicach 0,40–0,60 (np. singing osiąga 0,55 w 1910, dalej spada do 0,45 w 1990). Początek XX wieku przynosi wzmocnienie semantyki akademickiej (conservatory, orchestral, symphony), a od lat 60. coraz częściej pojawiają się terminy instytucjonalne, związane z przemysłem muzycznym (BMI 0,55; ASCAP 0,54; Juilliard 0,41). W dekadach 1970–1990 do pola wchodzi także gatunki i kategorie popkulturowe – jazz (0,48 w 1970, 0,46 w 1990), folk (0,41–0,43), blues (0,40) czy nawet movies (0,39). W rezultacie obserwujemy systematyczny dryf semantyczny: od zakorzenienia w tradycji sztuk pięknych i instrumentarium klasycznym, przez semantykę akademicko-konserwatoryjną, aż po instytucjonalne i popkulturowe ujęcia muzyki charakterystyczne dla drugiej połowy XX wieku. Taka zgodność wyników z przewidywaną narracją jeszcze niczego nie dowodzi, ale wskazuje, że wyniki nie są przypadkowe. Dokładne wartości z szerszego badania prezentują się następująco:

Dekada 1800:

musick: 0.68

vocal: 0.62

dancing: 0.59

musical: 0.59

dances: 0.56

tune: 0.53

sculpture: 0.53

singing: 0.51
poetry: 0.50
melody: 0.50
harp: 0.49
dance: 0.49
painting: 0.49
oratory: 0.48
musicians: 0.48
rhetoric: 0.47
belles: 0.45
sounds: 0.45
performers: 0.45
architecture: 0.45
instrumental: 0.45
songs: 0.45
arithmetic: 0.44
drums: 0.44
astronomy: 0.44
played: 0.44
surgery: 0.44
passionately: 0.44
strains: 0.43
singers: 0.43

Dekada 1810:

musick: 0.62
musical: 0.50
vocal: 0.50
tunes: 0.50
dancing: 0.49
flute: 0.45
violin: 0.45
poetry: 0.45
lyre: 0.45
singing: 0.44
tune: 0.44
dances: 0.43

songs: 0.42
sweeter: 0.41
melodious: 0.41
song: 0.40
painting: 0.40
gardening: 0.40
tones: 0.40
dance: 0.39
sound: 0.39
musicians: 0.39
singers: 0.39
sounds: 0.39
harp: 0.38
nightingale: 0.38
sung: 0.38
proficient: 0.37
dancers: 0.37
drums: 0.37

Dekada 1820:

musick: 0.55
dancing: 0.51
vocal: 0.49
singing: 0.49
guitar: 0.48
lyre: 0.47
accompaniment: 0.47
musical: 0.45
violin: 0.45
lute: 0.44
tunes: 0.44
tune: 0.44
poetry: 0.43
sound: 0.43
concerts: 0.42
flute: 0.42
piano: 0.42

composer: 0.41
harp: 0.41
operas: 0.41
orchestra: 0.40
sounds: 0.40
melody: 0.40
warbling: 0.40
harpsichord: 0.40
dance: 0.39
harps: 0.39
painting: 0.39
song: 0.39
songs: 0.38

Dekada 1830:

musical: 0.52
dancing: 0.51
singing: 0.46
vocal: 0.45
harpsichord: 0.45
melody: 0.45
poetry: 0.44
painting: 0.42
piano: 0.42
mozart: 0.42
operas: 0.42
cymbals: 0.42
flute: 0.42
drums: 0.41
warbling: 0.41
harp: 0.41
dances: 0.40
chanting: 0.40
violin: 0.39
sound: 0.39
melodious: 0.39
strains: 0.38

musicians: 0.38
sounds: 0.38
sculpture: 0.38
playing: 0.37
viol: 0.37
tones: 0.37
troubadours: 0.36
concerts: 0.36

Dekada 1840:

musical: 0.55
melody: 0.54
singing: 0.49
dancing: 0.49
songs: 0.48
vocal: 0.45
concerts: 0.44
tune: 0.43
poetry: 0.43
lute: 0.43
song: 0.42
viol: 0.42
strains: 0.42
tunes: 0.42
operas: 0.41
sounds: 0.41
melodious: 0.41
melodies: 0.40
flute: 0.40
tuned: 0.40
accompaniment: 0.40
cadence: 0.39
musicians: 0.39
sound: 0.39
painting: 0.39
harp: 0.38
cymbals: 0.38

dances: 0.38
chorus: 0.38
playing: 0.38

Dekada 1850:

musical: 0.55
flutes: 0.52
choral: 0.50
melody: 0.50
modulation: 0.49
dancing: 0.48
fifes: 0.47
flute: 0.47
singing: 0.46
strains: 0.46
viol: 0.46
vocal: 0.46
tune: 0.45
poetry: 0.45
harpsichord: 0.45
tuned: 0.44
melodious: 0.44
operas: 0.43
chime: 0.43
warbling: 0.42
song: 0.42
dances: 0.42
harp: 0.42
pealing: 0.42
sounds: 0.42
sound: 0.42
ballet: 0.41
lute: 0.41
dance: 0.40
painting: 0.40

Dekada 1860:

melody: 0.52
dancing: 0.51
musical: 0.48
singing: 0.48
cadence: 0.46
song: 0.45
flutes: 0.44
tune: 0.44
lute: 0.43
strains: 0.43
violin: 0.43
sounds: 0.42
painting: 0.42
choral: 0.42
guitar: 0.42
flute: 0.41
dance: 0.41
poetry: 0.41
chanting: 0.40
melodies: 0.40
songs: 0.39
nightingale: 0.39
sung: 0.39
singer: 0.39
lyre: 0.38
warbling: 0.38
viol: 0.38
vocal: 0.38
dances: 0.37
accompaniment: 0.37

Dekada 1870:

musical: 0.57
dancing: 0.53
singing: 0.53
lute: 0.49
melody: 0.49

poetry: 0.47
song: 0.46
flute: 0.45
singer: 0.44
sounds: 0.44
melodious: 0.43
dances: 0.43
dance: 0.42
choral: 0.42
strains: 0.41
sung: 0.41
songs: 0.41
vocal: 0.41
harp: 0.40
listening: 0.40
painting: 0.40
violin: 0.40
chorus: 0.39
tune: 0.39
piano: 0.39
tuned: 0.39
cadence: 0.39
tunes: 0.38
drama: 0.38
musician: 0.38

Dekada 1880:

musical: 0.53
dancing: 0.53
singing: 0.52
melody: 0.50
harping: 0.48
ditties: 0.48
strains: 0.47
symphonies: 0.47
fifes: 0.46
orchestra: 0.46

lulling: 0.45
waltz: 0.45
chorus: 0.44
accompaniment: 0.44
choral: 0.43
tune: 0.43
choruses: 0.43
vocal: 0.43
violin: 0.42
lute: 0.42
sounds: 0.42
anthems: 0.42
poetry: 0.42
chant: 0.42
guitar: 0.42
piano: 0.42
dances: 0.42
harmonies: 0.42
sang: 0.42
clangor: 0.41

Dekada 1890:

musical: 0.54
melody: 0.52
flute: 0.47
poetry: 0.46
dancing: 0.45
musicians: 0.45
orchestra: 0.44
singing: 0.44
recitations: 0.44
songs: 0.44
strains: 0.42
sound: 0.42
sounds: 0.41
violin: 0.41
lullaby: 0.41

flutes: 0.40
fiddles: 0.40
tunes: 0.40
symphonies: 0.39
chant: 0.39
carols: 0.39
rhythm: 0.39
symphony: 0.38
modulations: 0.38
tune: 0.38
harp: 0.38
playing: 0.37
singers: 0.37
chorus: 0.37
song: 0.37

Dekada 1900:

musical: 0.53
melody: 0.52
violin: 0.49
poetry: 0.47
singing: 0.46
flute: 0.46
strains: 0.46
dancing: 0.45
orchestra: 0.44
symphonic: 0.44
song: 0.44
rhythm: 0.44
songs: 0.44
banjo: 0.43
fifes: 0.42
tune: 0.42
conservatory: 0.42
musicians: 0.42
art: 0.42
piano: 0.42

sounds: 0.41
orchestral: 0.40
tunes: 0.40
recitative: 0.40
pianoforte: 0.39
concerts: 0.39
recitations: 0.39
playing: 0.39
listening: 0.39
sung: 0.39

Dekada 1910:

musical: 0.58
singing: 0.55
melody: 0.49
songs: 0.49
poetry: 0.48
dancing: 0.48
song: 0.47
violin: 0.43
orchestra: 0.43
piano: 0.42
accompaniment: 0.42
vocal: 0.41
art: 0.41
harp: 0.41
architecture: 0.41
conservatory: 0.41
tune: 0.41
literature: 0.40
drama: 0.40
sounds: 0.40
sculpture: 0.40
flute: 0.39
voices: 0.38
musicians: 0.38
strains: 0.38

dances: 0.37
listening: 0.37
rhythm: 0.37
lyric: 0.37
sung: 0.37

Dekada 1920:

musical: 0.53
poetry: 0.49
singing: 0.48
drama: 0.47
painting: 0.46
dancing: 0.46
symphony: 0.46
violin: 0.45
song: 0.45
sculpture: 0.44
orchestra: 0.44
art: 0.43
melody: 0.43
literature: 0.42
orchestral: 0.41
sung: 0.41
dance: 0.40
flute: 0.40
piano: 0.40
composer: 0.40
listening: 0.40
songs: 0.40
vocal: 0.40
architecture: 0.39
rhythm: 0.39
rhetoric: 0.39
musicians: 0.39
accompaniment: 0.38
dances: 0.38
tunes: 0.37

Dekada 1930:

musical: 0.55
singing: 0.50
dancing: 0.47
poetry: 0.45
listening: 0.44
choral: 0.44
violin: 0.44
piano: 0.43
song: 0.43
sculpture: 0.42
opera: 0.42
art: 0.41
orchestral: 0.41
singer: 0.41
songs: 0.41
drama: 0.41
painting: 0.41
arts: 0.40
vocal: 0.40
accompaniment: 0.40
dramatics: 0.40
dance: 0.39
conservatory: 0.39
theater: 0.39
symphony: 0.39
flute: 0.39
rhythm: 0.38
guitar: 0.38
dramatic: 0.38
concert: 0.38

Dekada 1940:

musical: 0.54
singing: 0.47
drama: 0.46

dancing: 0.46
art: 0.46
orchestral: 0.44
flute: 0.44
poetry: 0.44
symphony: 0.44
song: 0.44
melody: 0.43
dance: 0.43
piano: 0.42
composer: 0.41
violin: 0.41
orchestra: 0.41
musicians: 0.41
arts: 0.41
painting: 0.41
sculpture: 0.40
listening: 0.40
accompaniment: 0.40
literature: 0.39
choral: 0.39
operas: 0.38
dramatics: 0.38
conservatory: 0.38
sounds: 0.38
dramatic: 0.38
instrumental: 0.38

Dekada 1950:

musical: 0.58
melody: 0.48
art: 0.46
drama: 0.45
singing: 0.45
dancing: 0.45
poetry: 0.45
painting: 0.44

composer: 0.44
arts: 0.44
piano: 0.43
choral: 0.43
orchestra: 0.43
sculpture: 0.43
orchestral: 0.43
conservatory: 0.43
accompaniment: 0.42
violin: 0.41
songs: 0.41
dance: 0.41
musicians: 0.40
listening: 0.40
flute: 0.39
bach: 0.39
vocal: 0.39
symphony: 0.38
dramatics: 0.38
operas: 0.38
literature: 0.38
lyric: 0.38

Dekada 1960:

bmi: 0.55
ascap: 0.54
operatic: 0.52
musical: 0.52
singing: 0.51
dance: 0.50
dancing: 0.49
dramatics: 0.48
choral: 0.48
polyphonic: 0.47
vaudeville: 0.47
sculpture: 0.47
art: 0.46

soloists: 0.46
recitals: 0.46
accordion: 0.45
melody: 0.45
painting: 0.44
recitations: 0.44
songs: 0.44
drama: 0.44
accompaniment: 0.43
orchestra: 0.43
dances: 0.43
cantata: 0.42
arts: 0.42
piano: 0.42
bartok: 0.42
juilliard: 0.41
architecture: 0.41

Dekada 1970:

musical: 0.55
polyphonic: 0.49
poetry: 0.49
jazz: 0.48
dramatics: 0.48
singing: 0.46
song: 0.45
songs: 0.45
symphony: 0.44
dance: 0.44
piano: 0.44
bmi: 0.43
dancing: 0.43
art: 0.43
drama: 0.43
composer: 0.43
theater: 0.43
tanglewood: 0.42

ascap: 0.42
accompaniment: 0.42
orchestra: 0.42
theatre: 0.41
arts: 0.41
opera: 0.41
literature: 0.40
conservatory: 0.39
ballet: 0.39
lyrics: 0.39
recitations: 0.39
painting: 0.39

Dekada 1980:

musical: 0.61
dance: 0.55
poetry: 0.50
drama: 0.49
singing: 0.48
jazz: 0.48
opera: 0.47
lyrics: 0.47
dancing: 0.46
songs: 0.46
symphonic: 0.45
art: 0.45
dramatics: 0.44
song: 0.43
melody: 0.43
piano: 0.43
carols: 0.42
pantomimes: 0.42
dances: 0.42
theatre: 0.42
choral: 0.42
conservatory: 0.41
folk: 0.41

theater: 0.40
orchestra: 0.40
juilliard: 0.40
rhythm: 0.40
composers: 0.40
musicians: 0.40
arts: 0.40

Dekada 1990:

dance: 0.54
poetry: 0.48
art: 0.46
musical: 0.46
jazz: 0.46
opera: 0.45
singing: 0.45
songs: 0.45
dancing: 0.45
drama: 0.44
musicians: 0.44
song: 0.44
accompaniment: 0.44
dramatics: 0.43
melody: 0.43
lyrics: 0.43
folk: 0.43
dramatico: 0.42
orchestra: 0.42
theater: 0.41
symphony: 0.41
dances: 0.41
arts: 0.40
theatre: 0.40
piano: 0.40
painting: 0.40
blues: 0.40
movies: 0.39

carols: 0.39

listening: 0.38

Analiza podobieństw, znaczeń i określenie semantycznej ewolucji słowa '*music*' - muzyka, a może również w jakimś stopniu ewolucja zjawiska muzyki na podstawie tych danych to wymagające zadanie, którego w tej pracy się nie podejmę. Uważam samo wytworzenie danych za pracę badawczą – próbę pokazania, w jaki sposób można rozpocząć badanie semantyki dystrybucyjnej muzyki. A zatem dostrzegam wartość poznawczą w samym odkryciu i ustaleniu tych par słów i wartości liczbowych. Nie jest to jeszcze takie badanie semantyki dystrybucyjnej muzyki, jakie proponowałem w części 2 niniejszej pracy. Ale jest to badanie semantyki dystrybucyjnej słowa muzyka – co zgodnie z hipotezą dystrybucyjną jest badaniem blisko pokrewnym, choć zaledwie elementarnym i wysoce ogólnym.

4.2. Vibe composing

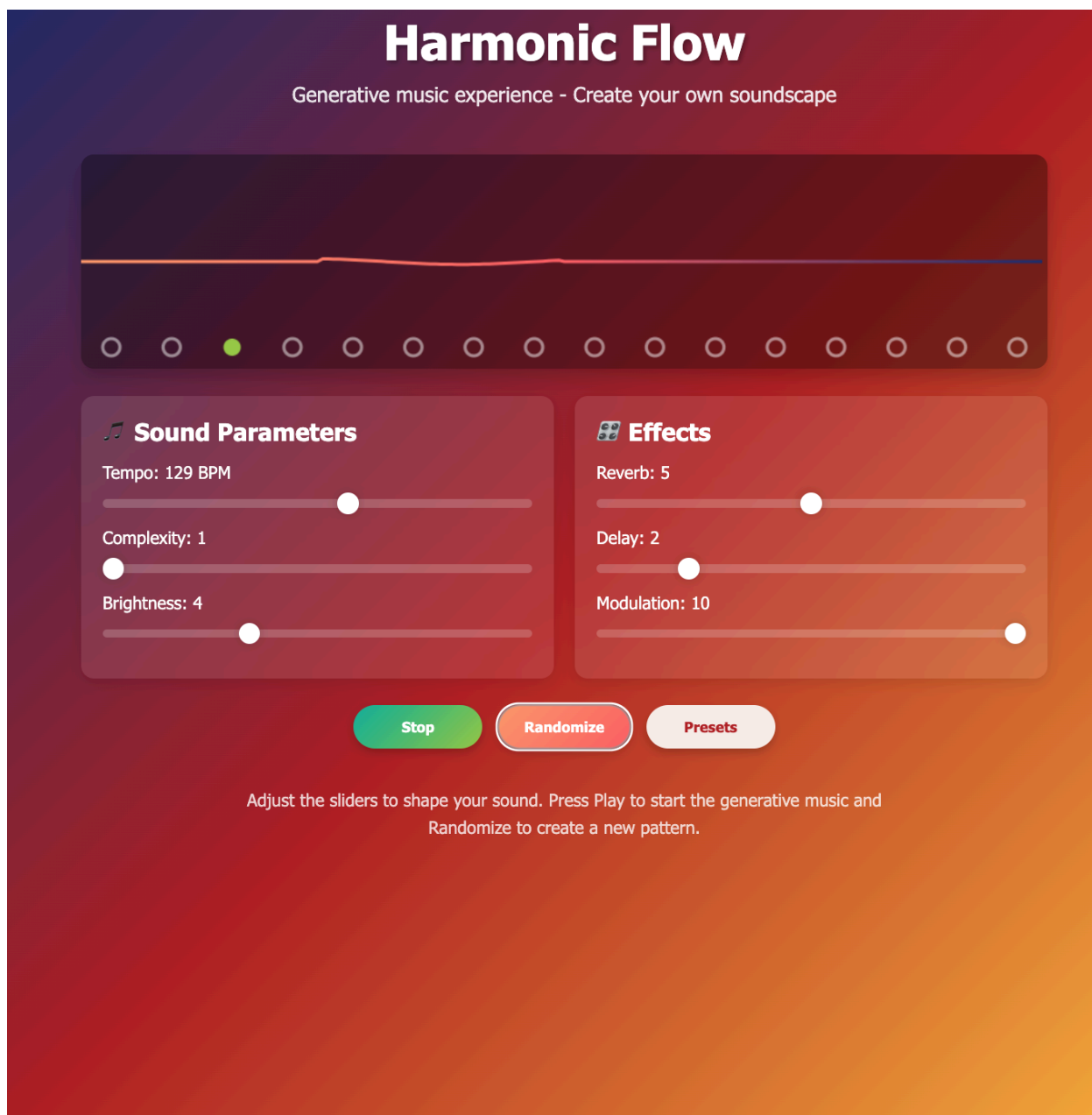
W 2025 znany autorytet ze świata AI, wszechstronny programista, współzałożyciel OpenAI Andrej Karpathy stworzył i spopularyzował termin *vibe coding* poprzez tweet na platformie X (Karpathy, 2025). Ogólnie oznacza on programowanie z pomocą AI. Dokładniej, termin nie doczekał się jeszcze naturalnego odpowiednika w języku polskim. Coding oznacza programowanie, a *vibe* to po prostu vibe/wajb (Banasiak, 2022), słowo zgłoszone do plebiscytu młodzieżowego słowa roku w 2021 (Wileczek, 2021a). *Vibe* to skrót od *vibration*, czyli wibracja, tak jak w wyrażeniu *good vibes* - dobre wibracje. Zatem *vibe* potocznie oznacza pewną aurę, nastrój, klimat (Banasiak, 2022). W języku angielskim również w formie czasownika to *vibe with somebody/something*, w formie polskiego zapożyczenia *vibe'ować z kimś/czymś* oznacza wspólnie odczuwać pewien nastrój, szczególnie pozytywny, stan odprężenia (Wileczek, 2021b), flow, a także wiąże się to ze wspólnym słuchaniem muzyki, wspólnym bujaniem się do muzyki, wspólnym tańcem. Vibe coding to zatem takie programowanie, w którym można dać ponieść się wibracji, można utrzymać stan flow, bez nadmiernego angażowania się w krytykę i poprawianie detali.

Ciekawa jest głęboka korelacja tej praktyki z muzyką. Twarzą *vibe codingu* już w czerwcu 2025 roku został Rick Rubin, jeden z najbardziej cenionych producentów muzycznych wszechczasów (Ray, 2025). Zarazem Rubin znany jest jako osoba, która przede wszystkim pojawia się w studiu, nie robi zbyt wiele, a albumy same się produkują. Rubin jest wielkim zwolennikiem taoistycznej praktyki działania przez niedziałanie (Stefon, 2025). Współtworzenie z AI ciekawie pasuje do tej koncepcji, ponieważ programista programuje, nie programując. Jeśli AI rozumieć jako element natury (a w tej pracy rozumiem je jako owoc ewolucji) zasada *wúwéi* autentycznie się tu wypełnia. Co ciekawe, Rubin wraz z firmą Anthropic stworzył swoisty interaktywny podręcznik do *vibe codingu*, który w całości opiera się na parafrazach *Tao te Ching* (Laozi, 1997; Rubin, 2025).

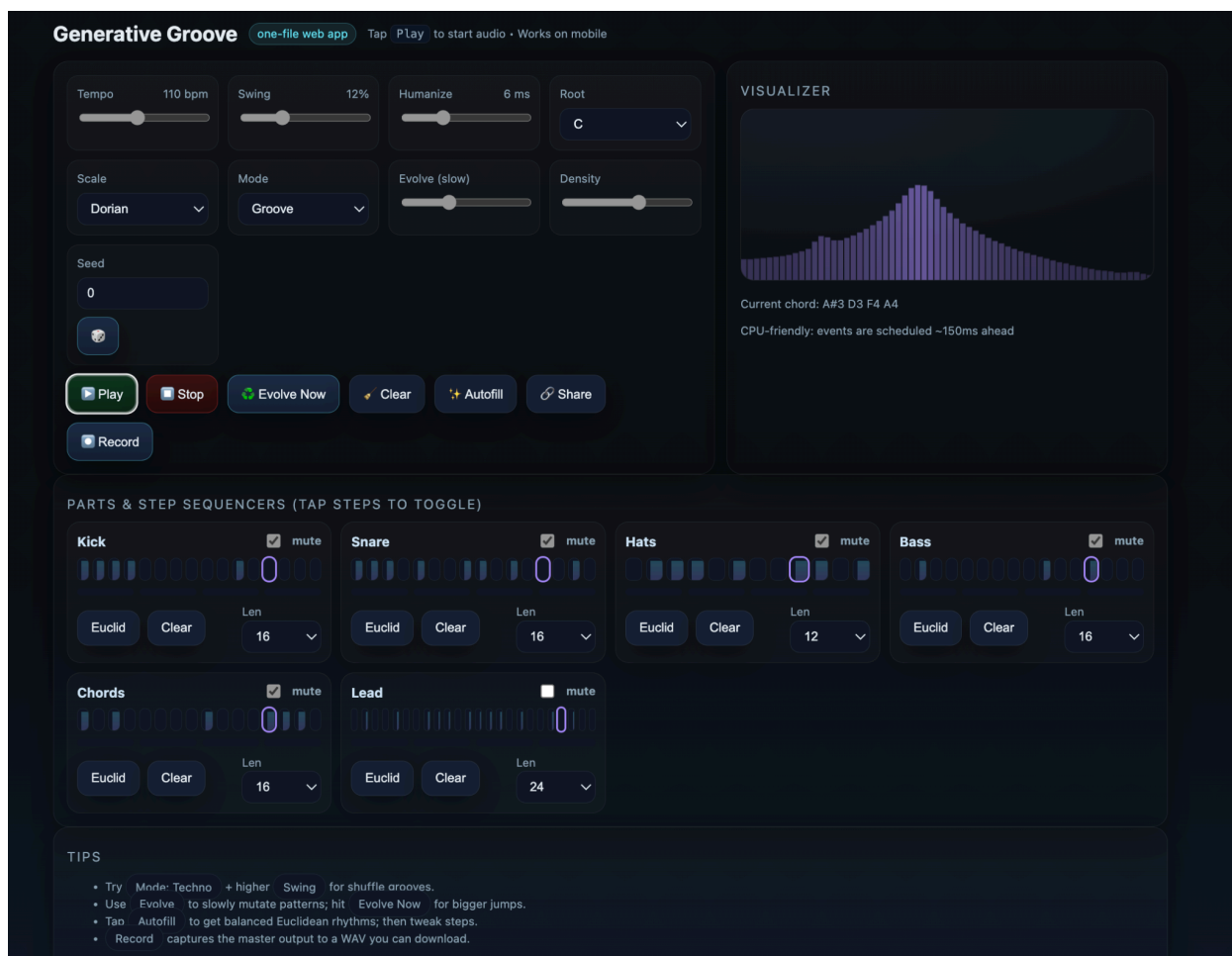
Poprzez praktykę *vibe codingu* można tworzyć również utwory muzyczne. W ramach wyodrębnienia tej praktyki proponuję nazwę *vibe composing*. Dokonałem już różnego rodzaju eksperymentów m.in. z różnymi systemami LLM. Poprzez komunikację w języku naturalnym prosiłem system o stworzenie np. kompozycji algorytmicznych w MIDI (Krzyżanowski, 2024). W niniejszym eksperymencie *vibe composingu* chcę wyjść o krok dalej. Chcę zaprezentować tworzenie programów do generacji utworów. Przede wszystkim mają one posłużyć jako demonstracja muzycznych możliwości drzemiących w tych systemach ogólnego przeznaczenia, co jest zarazem demonstracją najświeższych narzędzi w kulturze muzycznej, a zatem najnowszych efektów ewolucji kultury muzycznej. Porównam trzy systemy: ChatGPT, DeepSeek oraz Google Gemini.

W niniejszym eksperymencie proszę każdy system o to, żeby od ręki stworzył prosty program do generacji muzyki, który będzie można uruchomić jako stronę internetową w dowolnej przeglądarce. Dla prostoty i uniwersalności sugeruję, żeby taki program stworzyć jako jeden plik HTML. Prompt brzmi tak „*One-shot a generative music app in one html file. Make it interesting and interactive with easy and satisfying control. It should work as a website on desktop and mobile*”.

DeepSeek v3.1 stworzył aplikację, którą nazwał Harmonic Flow. Rzeczywiście działa ona z pojedynczego pliku i generuje różnorodne audio, przede wszystkim modulowane pętle fal sinusoidalnych. Aplikacja posiada także prostą wizualizację spektrogramu, co jest imponujące jak na prosty pojedynczy plik html. Audio teoretycznie można edytować poprzez suwaki: tempo, complexity, brightness, reverb, delay, modulation. Jednak subiektywnie słyszalne zmiany pojawiają się przede wszystkim przy modulacji tempa oraz przy kliknięciu przycisku *Randomize*, który dobiera losowe wartości wszystkich parametrów. Widnieje też przycisk *Presets*, który po kliknięciu informuje, że ta funkcja nie została jeszcze zaimplementowana. To prosty, ale ciekawy i przede wszystkim działający program do tworzenia muzyki. Aplikacja jest aktywna pod adresem <https://ziemniaki.github.io/harmonic-flow/>



Ten sam prompt wpisałem do ChatGPT z modelem GPT5. Za pierwszym razem wykonał nie działającą aplikację, której nie zdołał poprawić. Wpisałem zatem ten sam prompt jeszcze raz, w nowej konwersacji, licząc, że tym razem od zera uda się coś zbudować. Nie zawiodłem się – system bezbłędnie po pierwszym zapytaniu bez żadnych poprawek zbudował cały rozbudowany program do tworzenia muzyki:



Dokładniej mówiąc, jest to interesująca hybryda generatywnej aplikacji audio oraz programu do tworzenia muzyki. O generatywności należy wspomnieć nie tylko dlatego, że pojawia się ona w nazwie - *Generative Groove*. Poza tym, o ile tylko użytkownik kliknie przycisk *play*, program zaczyna sam tworzyć muzykę bez końca. Ścieżki maszyny perkusyjnej: *kick*, *snare*, *hats* oraz instrumenty nazwane *bass*, *chords* i *lead* wypełniają się same, co program określił mianem ewolucji. Z tego co rozumiem kod programu, to nie jest faktyczna ewolucja, ponieważ zachodzi żadne dziedziczenie i selekcja. To raczej po prostu losowe fluktuacje, stochastyczna zmienność. Zatem w tym akurat przypadku ewolucja muzyki jest tylko metaforą. Jeśli jednak użytkownik pragnie przejąć całkowitą kontrolę nad programem, również może to zrobić. Suwak *Evolve (slow)* należy przeciągnąć do pozycji zerowej, a wtedy wszelkich zmian trzeba już dokonywać ręcznie.

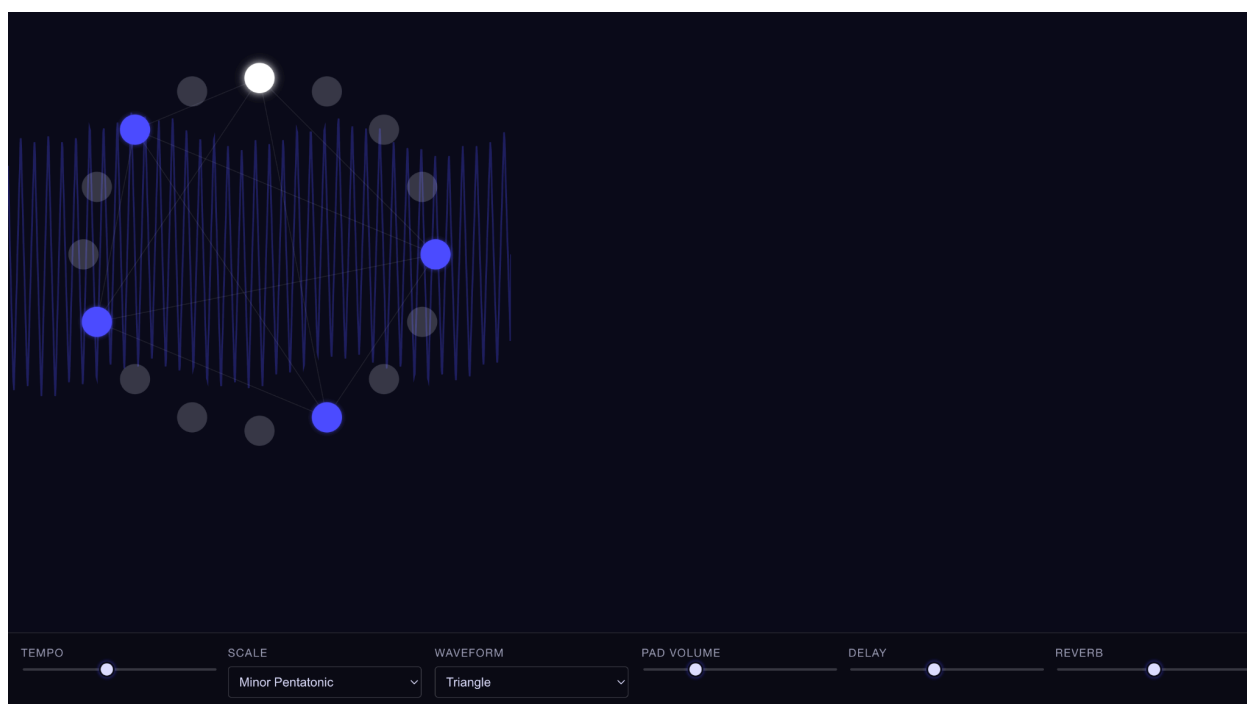
Według podanych informacji ChatGPT stworzył program *Generative Groove* w 21 sekund. Ponadto, jego pełen rozmiar to 30kb – czyli na jednej dyskietce zmieściłby się w 48 kopiach, a na płycie CD w 24 000 kopii. Mimo tych imponująco niewielkich rozmiarów program zawiera funkcje:

- zmiany tempa,
- kontroli swingu (globalnego odchylenia w mikrorytmice),
- humanizacji (losowych odchyień w mikrorytmice),
- wyboru tonacji oraz skali dla instrumentów,
- wyboru jednego z trzech ważonych trybów losowej generacji sekwencji,
- gęstości (częstotliwości) losowo generowanych sekwencji,
- wyboru bazowego *seedu* (bazowej liczby do *randomizera*),
- funkcji *'evolve now'*, która jednorazowo w przyspieszonym tempie dokonuje losowych modyfikacji,
- funkcji *'clear'*, która wyzerowuje dotychczasowy stan programu,
- 'autofill'*, który dopełnia wolne miejsca w sekwencerze,
- guzik share, który kopiuje link nie tylko do aplikacji, ale także do całego aktualnego stanu aplikacji
- funkcjonalny guzik *'record'*, który na żywo nagrywa i eksportuje plik audio z bieżącego wykonania (działa wyłącznie w przeglądarce na komputerze, nie działa na smartfonie)
- funkcjonalny sekwencer dla każdej ścieżki z dodatkową możliwością zmiany długości pętli, funkcją *clear* oraz wypełnieniem *'euclid'*
- funkcję wyciszenia dla każdej ścieżki
- wizualizację dźwięku: globalny wizualizer podzielony na pasma częstotliwości oraz sygnalizator chwilowego poziomu dynamiki indywidualnie dla każdej ścieżki

To nie tylko funkcjonalny, kreatywny program do dźwiękowej zabawy, ale wręcz gotowy do sprzedaży wielofunkcyjny produkt. Fakt wytworzenia takiego programu w 21 sekund to doskonała miara współczesnego zaawansowania AI w świecie muzyki. Program

przekracza maksymalne oczekiwania, z którymi podchodziłem do tego eksperymentu. Aplikacja jest aktywna pod adresem <https://ziemniaki.github.io/Generative-Groove/>

Dalej przechodzę jeszcze do analogicznej próby z Google Gemini 2.5 pro. Gemini na podstawie tego samego promptu również wytworzył interesujący program muzyczny: *Stellaris Orb*. Aplikacja pod względem działania przypomina nieco aplikację stworzoną przez DeepSeek. Proponuje ona ograniczoną manipulację pewną pętlą dźwiękową. W odróżnieniu od DeepSeek, który daje rezultaty raczej w estetyce atonalnej muzyki elektroakustycznej, program *Stellaris Orb* stworzony przez Gemini tworzy bardziej melodyjne, tonalne pętle:



Użytkownik ma możliwość wyboru jednej z sześciu skal i czterech barw, czy też kształtów fal, które odtwarzają zadaną pętlę. Poza tym istnieje możliwość manipulacji głośnością, tempem, pogłosem (*reverb*) oraz echem (*delay*). W centrum interfejsu znajduje się okrąg, który jest sekwencerem. Każde pole zawiera dźwięk o określonej wysokości, który znajduje się w odpowiednim ustalonym miejscu pętli – dźwięki są uszeregowane wznosząco. Każdy dźwięk można włączyć lub wyłączyć.

Poza tym na okrąg nałożony jest spektrogram, który na bieżąco pokazuje wizualizację fali dźwiękowej. Program działa bez zarzutu, bez żadnych poprawek, napisany za pierwszym razem. Aplikacja jest dostępna pod adresem: <https://ziemniaki.github.io/Stellaris-Orb/>

W mojej ocenie tę 'konkurencję' wygrywa ChatGPT. Poziom złożoności napisanego przez niego programu znacznie przerasta wyniki od konkurentów, a przede wszystkim to właśnie *Generative Groove* jest najbardziej wszechstronnym i dopracowanym muzycznym narzędziem. Tylko ten program posiada cechy pełnej cyfrowej stacji roboczej (Digital Audio Workstation, DAW) (Manning, 2013).

Przechodzę do podsumowania eksperymentu mającego na celu zaprezentowanie muzycznego współtworzenia z AI poprzez tekst, które określam jako *vibe composing*. Moim zdaniem spośród konkurentów najwyższy poziom zaawansowania technologicznego pokazuje ChatGPT programem *Generative Groove*. Być może nie wszystkie wyniki tej kreatywnej pracy są artystycznie imponujące, ale systemy AI nieustannie się rozwijają, a nawet i te obecne bez wątpienia pokazałyby jeszcze znacznie lepsze efekty, gdyby każdy z tych rezultatów dalej poprawiać, edytować, rozwijać. Moim zdaniem możliwości tych systemów w sferze muzyki są ogromne. Nie oznacza to, że można w nich od razu stworzyć wszystko zgodnie z dowolną wizją. Mają swoje mocne i słabe strony. Ale znając te mocne strony, można zejść bardzo daleko.

Zakończenie

W mojej dysertacji przedstawiam i usiłuję wyjaśnić mechanizmy procesów ewolucji kulturowej muzyki. Staralem się wykazać, że owa ewolucja kulturowa muzyki może być najlepiej zrozumiana w perspektywie programu badawczego memetyki. Ta właśnie teoria pozwala traktować muzyczne idee i struktury jako replikatory podlegające selekcji, co moim zdaniem jest cenną i unikatową perspektywą. Podkreśliłem to, że tylko memetyka zakłada istnienie odrębnego replikatora. Usiłowałem wykazać to, kierując się zasadą, że replikator z definicji musi 'zabiegać' o swój byt. W tym celu zademonstrowałem różne muzyczne systemy, które w różnym stopniu przejawiają autonomię i funkcje manipulacji otoczeniem.

Pokazałem również, że szczególnie istotną rolę w omawianych procesach ewolucji odgrywają technologie cyfrowe, a wśród nich kluczowe miejsce zajmuje sztuczna inteligencja. Analiza pozwoliła stwierdzić, że między innymi memetyczna dynamika ewolucji muzyki doprowadziła do powstania muzycznej sztucznej inteligencji. Jednym z jej najważniejszych przejawów są muzyczne systemy rekomendacji. Owe rekomendery w sposób nienadzorowany przez człowieka organizują globalny obieg muzyki. Stanowią one być może najważniejsze ucieleśnienie tendencji do rozwoju coraz większej sprawczości i autonomii muzycznych memów. Jednocześnie rola i sprawczość człowieka spada. To, że muzyczne maszyny przejmują kompetencje człowieka, unaocznia także fakt, że memy i geny mogą ze sobą konkurować. Zauważam szerzej, że muzyczne memy oddziałują nie tylko na człowieka, ale także na szersze środowisko biokulturowe w taki sposób, aby coraz skuteczniej prowokować wytwarzanie własnych kopii. Ostatecznie, zapośredniczona w technologii kultura muzyczna może być opisana jako system, którego działanie coraz bardziej przypomina intencjonalne zabieganie o własne przetrwanie.

Niniejsza praca ukazuje fascynujący postęp ewolucji memetycznej na przykładach ze świata muzyki. Ewolucja muzyki to nie tylko czasowa zmienność materiału muzycznego, przemijanie stylów, następstwo jednej mody za drugą. Ewolucja muzyki to także nieustanne umacnianie się dynamicznych systemów informacyjnych, które służą jako konstrukcyjny szkielet podtrzymujący zmienność muzyki brzmiącej. Melodie uporczywie wpadające w ucho, muzyka nierozzerwalnie związana z kultami i rytuałami, system muzyczny związany ze społecznym prestiżem, muzyka kupowana i sprzedawana jako styl życia, w końcu powstanie muzycznych mega-korporacji zarabiających na pasjonatach i amatorach. Muzyka zawsze potrafiła wpływać na

zachowanie ludzi, za to pojedynczemu człowiekowi jeszcze trudniej jest wpłynąć na kształt muzycznego systemu w odczuwalnym stopniu.

Badanie muzycznych memów internetowych pokazuje jedynie wybrane przykłady, ale repertuar memowy jest tak bogaty, że w praktyce wręcz niemożliwy do dogłębnego przebadania. A jednak właśnie badanie treści mediów społecznościowych, TikToka czy YouTube shorts może moim zdaniem dać wymierne informacje na temat autentycznie najnowszych interesujących struktur muzycznych. To na tym polu tworzy się obecnie najwięcej treści kultury. W roku 2025 wyłania się nowy gatunek audiowizualnych treści nazywany brainrot, czyli gnicie mózgu. Pomimo tego, że wiedza o istnieniu takiego gatunku jest niszowa, we wrześniu 2025 byłem świadkiem, jak dwóch dwuletnich chłopców śpiewało przy zabawie piosenkę z nurtu brainrot. Jednym z powodów, dla których warto owe muzyczne memy analizować jest budowanie wiedzy i publicznej świadomości na temat ich wpływu na zdrowie człowieka i społeczeństw. Bieżące badanie memetyczno-muzycznej treści internetu jest dziś równie ważne z powodów politycznych - memy są skutecznym narzędziem manipulacji opinią publiczną, a stosowana w nich muzyka pełni ważną rolę narracyjną. Trzecim powodem, dla którego warto bacznie obserwować internetowe treści jest nieustannie rosnąca ilość materiałów tworzonych przy użyciu AI. Jest to przede wszystkim zjawisko nowe, a więc nieznane i nie wiadomo do czego doprowadzi.

Także w zakresie rozważań teoretycznych dotyczących memetyki muzyki wiele problemów czeka wciąż na satysfakcjonujące rozwiązanie. Przede wszystkim dotychczasowy model muzyczno-memetycznych badań Stevena Jana nadal się nie rozpowszechnił. Świadczy to moim zdaniem albo o tym, że zagadnienie to nie jest interesujące dla większości badaczy, albo o tym, że model wymaga korekty. Oparcie memetyki muzyki na badaniu partytur w erze TikToka i AI moim zdaniem jest nieskuteczną strategią. Te same lub podobne pojęcia zastosowane do badań ilościowych w internecie mogą dostarczyć wielu ciekawych informacji.

Schemat do badania i opisu sprawczych muzycznych instytucji w oparciu o teorię aktora-sieci, model systemu wiabilnego i wielopoziomą architekturę kompetencji został szeroko zastosowany tylko jednorazowo. Jest to gotowe narzędzie badawcze, które może stać się podstawą dla wielu podobnych analiz. Poza tym sam ten schemat można znacznie bardziej rozwinąć. Teoria aktora-sieci pozwala na zmapowanie sieci zależności pomiędzy wszelkimi ludzkimi i nie-ludzkimi aktorami w instytucjonalnej sieci. Taką mapę, określającą na przykład przepływ informacji, można rzeczywiście określić. Wzbogacona o model systemu wiabilnego może stać się nie tylko ciekawym narzędziem do badań, ale może nawet ujawnić praktyczne słabe i mocne strony danej

muzycznej instytucji. Dążenie do wprowadzenia działającego schematu kontroli zgodnego z VSM powinno prowadzić do optymalizacji działania instytucji.

Osobny obszar stanowi muzykologiczne badanie technologii AI. Nawet technologie, które już wyszły z obiegu nie są dobrze opisane, a tym bardziej najnowsze, szczególnie w języku polskim. Znaczącym postępowaniem na tym polu jest książka Marcina Strzeleckiego *Niech się stanie muzyka! O muzycznej generatywności*. Ale również i ta obszerna publikacja nie opisuje wszystkiego. Autor skupił się na technologiach generatywnych, a zatem nie opisuje szerzej systemów rekomendacji muzycznej. Te technologie towarzyszą dziś znacznej większości słuchaczy na całym świecie, a nadal nie są powszechnym przedmiotem zainteresowań muzykologii. Moje opisy systemów z Spotify, TikTok i YouTube mają formę elementarną. Każdy z tych systemów z osobną moim zdaniem zasługuje na osobne, szeroko zakrojone badania. To zaawansowane technologie, które sterują konsumpcją muzyki na całym świecie. Moim zdaniem są idealnym przedmiotem badawczym dla muzykologii.

Całym, nadal prężnie rozwijającym się nowym polem badawczym jest kreatywność maszyn. Kiedy zaczynałem pisać tę pracę systemy generacji muzyki z promptów tekstowych jeszcze nie istniały. Zainspirowany sukcesami AlphaGo wnioskowałem, że niebawem sztuczna inteligencja nowego rodzaju rozwinie się również w dziedzinie muzyki, ale nie przypuszczałem, że nabierze to aż takiego tempa. Opisywana przeze mnie aplikacja do generowania muzyki Suno była w tamtym momencie w swojej trzeciej wersji. W międzyczasie ukazała się już czwarta, a we wrześniu 2025 twórcy zapowiedzieli nadchodzącą premierę piątej odsłony systemu. Utwory stworzone w trzeciej wersji, jak repertuar fikcyjnego zespołu The Velvet Sundown, który zyskał miliony odsłuchań na platformach streamingowych, charakteryzują się jeszcze pewnym rodzajem szumu w partiach wokalnych. W praktyce rozpoznałem muzykę generowaną w Suno rozbrzmiewającą w dwóch restauracjach. Czy właściciele o tym wiedzą? Czy klienci o tym wiedzą? Czy piąta wersja systemu będzie jeszcze tworzyła utwory rozpoznawalne jako AI? A siódma wersja? A dziewiąta?

Bibliografia:

Abdollahpouri, H., Burke, R., i Mobasher, B. (2017). Recommender systems as multistakeholder environments. W Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization, UMAP '17.

Abidin, C., & Kaye, D. B. V. (2021). Audio memes, earworms, and templatability: The 'aural turn' of memes on TikTok. *Critical meme reader: Global mutations of the viral image*, 58-68.

Acerbi, A. (2019). *Cultural evolution in the digital age*. Oxford University Press.

Adami, C. (2012). The use of information theory in evolutionary biology. *Annals of the New York Academy of Sciences*, 1256(1), 49–65.
<https://doi.org/10.1111/j.1749-6632.2011.06422.x>

Adams, C. R. (1976). Melodic contour typology. *Ethnomusicology*, 20(2), 179–215.

Adams, C. T. (2023). Music Publishing in Puccini's Italy. In A. Wilson (Ed.), *Puccini in Context* (pp. 122–129). chapter, Cambridge: Cambridge University Press.

Adorno, T. W. (1949). *Philosophy of New Music*. University of Minnesota Press.

Afchar, D., Meseguer-Brocal, G., & Hennequin, R. (2025). AI-generated music detection and its challenges. arXiv preprint arXiv:2501.10111. <https://arxiv.org/abs/2501.10111>

Aggarwal, C. C. (2016). *Recommender systems: The textbook*. Springer.

Agüera y Arcas, B., Alakuijala, J., Evans, J., Laurie, B., Mordvintsev, A., Niklasson, E., Randazzo, E., & Versari, L. (2024). *Computational Life: How well-formed,*

self-replicating programs emerge from simple interaction (arXiv preprint arXiv:2406.19108). <https://doi.org/10.48550/arXiv.2406.19108>

Ahuja, Siddarth, 2025 Ableton MCP <https://github.com/ahujasid/ableton-mcp>

Akçay, Ç., & Nowicki, S. (2023). An ecological and neurobiological perspective on the evolution of vocal learning. *Frontiers in Ecology and Evolution*, 11, 1193903.

Alberts B, Johnson A, Lewis J, et al. *Molecular Biology of the Cell*. 4th edition. New York: Garland Science; 2002. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK21054/>

Alexander, 2014. Gibson Min-ETune – Automatic Tuner <https://www.guitarsite.com/gibson-min-etune-automatic-tuner/>

Alim, H. S. (2006). *Roc the mic right: The language of hip hop culture*. Routledge.

Allegra Rosenberg, 'BBL Drizzy' Was the Beginning of the Future of AI Music, October 2024, <https://www.wired.com/story/bbl-drizzy-foretold-the-future-of-ai-music/>

Allocca, K. (2013, February 12/15). The Harlem Shake Has Exploded (Updated). YouTube Trends. <https://youtube-trends.blogspot.com/2013/02/the-harlem-shake-has-exploded.html>

Alruthaya, A., Nguyen, T.-T., & Lokuge, S. (2021). The Application of Digital Technology and the Learning Characteristics of Generation Z in Higher Education. arXiv. <https://arxiv.org/abs/2111.05991>

Altman, S. (2024, January 10). Reflections. Sam Altman Blog. <https://blog.samaltman.com/reflections>

Ammirante, P., & Russo, F. (2023). Towards a vocal constraints model of melodic expectancy: Evidence from two listening experiments. *Music & Science*, 6(1). <https://doi.org/10.1177/20592043231179410>

Anglada-Tort, M., Harrison, P. M. C., Lee, H., & Jacoby, N. (2023). Large-scale iterated singing experiments reveal oral transmission mechanisms underlying music evolution. *Current Biology*, 33(8), 1472–1486.e12. <https://doi.org/10.1016/j.cub.2023.02.070>

asddfds. (2016, May 4). MLG pro. In Urban Dictionary. Retrieved September 27, 2025, from <https://www.urbandictionary.com/define.php?term=MLG%20pro>

Aubé, W., Peretz, I., & Armony, J. (2013). The effects of emotion on memory for music and vocalisations. *Memory*, 21. <https://doi.org/10.1080/09658211.2013.770871>

Azadi, P. (2025). Computational Irreducibility as the Foundation of Agency: A Formal Model Connecting Undecidability to Autonomous Behavior in Complex Systems. arXiv. <https://arxiv.org/abs/2505.04646>

Bagdikian, B. H. (2004). *The new media monopoly*. Beacon Press.

Banasiak, D. (2022, 3 czerwca). vibe [Hasło]. Obserwatorium Językowe UW. Pobrane z <https://obserwatoriumjezykowe.uw.edu.pl/hasla/vibe/>

Bandi, A., Kongari, B., Naguru, R., Pasnoor, S., & Vilipala, S. V. (2025). The Rise of Agentic AI: A Review of Definitions, Frameworks, Architectures, Applications, Evaluation Metrics, and Challenges. *Future Internet*, 17(9), 404. <https://doi.org/10.3390/fi17090404>

Bannan, N., & Harvey, A. R. (2025). Music as a social instrument: A brief historical and conceptual perspective. *Cognition (Frontiers)*

Barkan, O., & Koenigstein, N. (2016). Item2Vec: Neural Item Embedding for Collaborative Filtering. arXiv preprint arXiv:1603.04259

Bartel, J. (2016, June 9). MLG Anaheim Call of Duty groups breakdown. ESPN.

Bartocci E, Lió P (2016) Computational Modeling, Formal Analysis, and Tools for Systems Biology. PLoS Comput Biol 12(1): e1004591. <https://doi.org/10.1371/journal.pcbi.1004591>

Bates, E. (2019). Actor-Network Theory and Organology. The Oxford Handbook of Organology.

Baudry, B., & Monperrus, M. (2022). Exhaustive survey of Rickrolling in academic literature. arXiv. <https://arxiv.org/abs/2204.06826>

Baumol, W. J., & Bowen, W. G. (1966). Performing Arts: The Economic Dilemma. The Twentieth Century Fund.

BBC, 2016. Artificial intelligence: Go master Lee Se-dol wins against AlphaGo program <https://www.bbc.com/news/technology-35797102>

Beer, S. (1972) Brain of the Firm, Allen Lane, The Penguin Press, London, Herder and Herder, USA

Beer, S. The Viable System model, 1984

Bent, I. D., Hughes, D. W., Provine, R. C., Rastall, R., Kilmer, A., Hiley, D., Szendrei, J., Payne, T. B., Bent, M., & Chew, G. (2001). Notation. In S. Sadie (Ed.), Grove Music Online. Oxford University Press. <https://doi.org/10.1093/gmo/9781561592630.article.20114>

Beyene, Y., Katoch, S., WoldeGabriel, G., Hart, W. K., Uto, K., Sudo, M., Kondo, M., Hyodo, M., Renne, P. R., Suwa, G., & Asfaw, B. (2013). The characteristics and chronology of the earliest Acheulean at Konso, Ethiopia. *Proceedings of the National Academy of Sciences of the United States of America*, 110(5), 1584–1591. <https://doi.org/10.1073/pnas.1221285110>

Bhavika Bansal, YouGov. (2024, April 30). Can AI be America's next top idol? YouGov. <https://business.yougov.com/content/49276-can-ai-be-americas-next-top-idol>

Billboard. (2025). Darude – Hot 100 singles chart history. Retrieved from <https://www.billboard.com/artist/darude/chart-history/hsi/>

Blackmore, S. (1999). *The meme machine*. Oxford University Press.

Bloch, M. (2000). A well-disposed social anthropologist's problems with memes. In R. Auger (Ed.), *Darwinizing culture: The status of memetics as a science*. Oxford University Press.

Bogdan, I. (2018). The Charter of the Kwidzyn (Marienwerder) Convivium Musicum of 1587 as a source for the history of musical culture in Prussian towns. *Interdisciplinary Studies in Musicology*, 11, 213–234. Pobrano z <https://pressto.amu.edu.pl/index.php/ism/article/view/15050>

Bourdieu, P. (1984). *Distinction: A Social Critique of the Judgement of Taste*. Harvard University Press.

Bourdieu, P. (1993) *The Field of Cultural Production: Essays on Art and Literature*. Columbia University Press, New York.

Börzsei, L. K. (2013). *Makes a meme instead: A concise history of internet memes*. Utrecht University. Retrieved from

<https://www.semanticscholar.org/paper/Makes-a-Meme-Instead%3A-A-Concise-History-of-Internet-B%C3%B6rzsei/b355b891a35360ab4c89397f8b67b8017816c931>

Bowling, D. L., & Purves, D. (2015). A biological rationale for musical consonance. *Proceedings of the National Academy of Sciences of the United States of America*, 112(36), 11155–11160. <https://doi.org/10.1073/pnas.1505768112>

Boyd, G. (2001). The Human Agency Of Meme Machines: A review of the Meme Machine.

Journal of Memetics - Evolutionary Models of Information Transmission, 5.

http://cfpm.org/jom-emit/2001/vol4/boyd_g.html

Boyd, R., & Richerson, P. J. (1995). “Why does culture increase human adaptability?” *Ethology and Sociobiology*, 16, 125–143.

Boyd, R., & Richerson, P. J. (2009). Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1533), 3281–3288. <https://doi.org/10.1098/rstb.2009.0134>

BPI. (2024, October 24). New survey reveals UK fans want greater transparency over AI generated music - 'All About The Music 2025' Out Now. BPI. <https://www.bpi.co.uk/news-analysis/new-survey-reveals-uk-fans-want-greater-transparency-over-ai-generated-music>

Bradlow, E. T., & Fader, P. S. (2001). A Bayesian lifetime model for the “Hot 100” Billboard songs. *Journal of the American Statistical Association*, 96(454), 368–381. <https://doi.org/10.1198/016214501753168091>

Brandon, & LiterallyAustin. (2024). MLG Era TikTok Revival / MLG Meme Revival. Know Your Meme. Retrieved September 27, 2025, from <https://knowyourmeme.com>

Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and early language acquisition. *Frontiers in Psychology*, 3, 327. <https://doi.org/10.3389/fpsyg.2012.00327>

Bringsjord, S., Bello, P. & Ferrucci, D. Creativity, the Turing Test, and the (Better) Lovelace Test. *Minds and Machines* 11, 3–27 (2001). <https://doi.org/10.1023/A:1011206622741>

Briot, J.-P., Hadjeres, G., & Pachet, F. (2017). Deep learning techniques for music generation: A survey. *ACM Computing Surveys*, 50(6), 1–34. <https://hal.sorbonne-universite.fr/hal-01660772v1/file/dl4mg-v10.pdf>

Briot, JP., Pachet, F. Deep learning for music generation: challenges and directions. *Neural Comput & Applic* 32, 981–993 (2020). <https://doi.org/10.1007/s00521-018-3813-6>

Britt, B. C. (2019). The use of nondialogic trolling to disrupt online communication. *First Monday*, 24(10). <https://doi.org/10.5210/fm.v24i10.10164>

Brown, A., & Gullberg, M. (2008). BIDIRECTIONAL CROSSLINGUISTIC INFLUENCE IN L1-L2 ENCODING OF MANNER IN SPEECH AND GESTURE: A Study of Japanese Speakers of English. *Studies in Second Language Acquisition*, 30(2), 225–251. doi:10.1017/S0272263108080327

Brown, A. R. (2016). Understanding musical practices as agency networks. In F. Pachet, A. Cardoso, V. Corruble, & F. Ghedini (Eds.), *Proceedings of the Seventh International Conference on Computational Creativity (ICCC 2016)* (pp. 139–146). Sony CSL. <https://www.computationalcreativity.net/iccc2016/proceedings-2016/>

Brown, S. (2000). The "musilanguage" model of music evolution. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 271–300). MIT Press.

Brown, Tom & Mann, Benjamin & Ryder, Nick & Subbiah, Melanie & Kaplan, Jared & Dhariwal, Prafulla & Neelakantan, Arvind & Shyam, Pranav & Sastry, Girish & Askell, Amanda & Agarwal, Sandhini & Herbert-Voss, Ariel & Krueger, Gretchen & Henighan, Tom & Child, Rewon & Ramesh, Aditya & Ziegler, Daniel & Wu, Jeffrey & Winter, Clemens & Amodei, Dario. (2020). Language Models are Few-Shot Learners. 10.48550/arXiv.2005.14165.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, 25(5), 975–979. DOI: 10.1121/1.1907229

Brutus, W. (2024, May 16). Delray Beach artist's viral song is removed by Spotify, raising questions over AI music. WLRN. <https://www.wlrn.org/arts-culture/2024-05-16/ai-music-kendrick-drake-bbl-drizzy-delray-beach-willonius>

Bryant, M. (2024, March 19). Swedish composer becomes Spotify's most-famous musician you've never heard of. *The Guardian*. <https://www.theguardian.com/uk-news/2024/mar/19/swedish-composer-johan-rohr-becomes-spotifys-most-famous-musician-youve-never-heard-of>

Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 12(4), 331–370.

Busse Berger, A. M., & Rodin, J. (Eds.). (2015). *The Cambridge History of Fifteenth-Century Music*. Cambridge University Press

Butsch, R. (2000). *The making of American audiences: From stage to television, 1750–1990*. Cambridge University Press. (See excerpt: https://assets.cambridge.org/97805216/64837/excerpt/9780521664837_excerpt.pdf)

Byczkowska-Owczarek, D. (2022). "A Socially Constructed Individualist: An Interactionist Study of Role-Making among Orchestral Conductors." *Qualitative Sociology Review*.

Caldwell, D. (2014, September 16/updated 2024, December 27). Major League Gaming / MLG. Know Your Meme.

Callon, M. (1986). Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay. In J. Law (Ed.), *Power, Action and Belief: A New Sociology of Knowledge?* (pp. 196-223). Routledge.

Calvin, W.H., 1997; The Six Essentials? Minimal Requirements for the Darwinian Bootstrapping of Quality. *Journal of Memetics - Evolutionary Models of Information Transmission*, 1.

http://cfpm.org/jom-emit/1997/vol1/calvin_wh.html

Calvino, F., 2024. Fostering an inclusive digital transformation as AI spreads among firms. OECD Policy Brief.

https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/11/fostering-an-inclusive-digital-transformation-as-ai-spreads-among-firms_cd50d324/5876200c-en.pdf

Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review*, 67(6), 380–400.

<https://doi.org/10.1037/h0040373>

Campbell, J. O. (2016). Universal Darwinism as a process of Bayesian inference. *Frontiers in Systems Neuroscience*, 10, 49. <https://doi.org/10.3389/fnsys.2016.00049>

Campbell, P. S., & Dahm, L. C. (2022). A social history of music in US schools: The coalescence of research and practice. *Music Research Annual*, 3, 1–32.

<https://musicresearchannual.org/campbell-dahm-music-in-us-schools-2>

Carnovalini, F., & Roda, A. (2020). Computational creativity and music generation systems: An introduction to the state of the art. *Frontiers in Artificial Intelligence*, 3, 14.

Carter J. H. (2000). The immune system as a model for pattern recognition and classification. *Journal of the American Medical Informatics Association : JAMIA*, 7(1), 28–41. <https://doi.org/10.1136/jamia.2000.0070028>

Cervi, L., Tejedor, S., & Tusa, F. (2021). Slapping cats, bopping heads, and oreo shakes: Understanding indicators of virality in TikTok short videos. *ArXiv*. [/abs/2111.02452](https://arxiv.org/abs/2111.02452)

Chaffey, D. 2025 Global social media statistics research summary <https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/>

Chaitin, G. J. (2012). *Proving Darwin: Making biology mathematical*. Vintage.

Chaudhary, A. (2010, December 3). Keyboard Cat's World Series Wonderful Pistachios Commercial! CatSynth. <https://catsynth.com/tag/keyboard-cat/>

Cheung et al., Uncertainty and Surprise Jointly Predict Musical Pleasure and Amygdala, Hippocampus, and Auditory Cortex Activity, *Current Biology* (2019), <https://doi.org/10.1016/j.cub.2019.09.067>

Choi, K., Fazekas, G., Sandler, M., & Cho, K. (2017). Convolutional recurrent neural networks for music classification. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2392–2396. <https://doi.org/10.1109/ICASSP.2017.7952585>

Chomsky, N. (1981). *Lectures on Government and Binding: The Pisa lectures*. Foris Publications.

Chopade, Pranali & Chormale, Nikita & Shinde, , & Shitole, Vishakha & Hanchate, Dinesh & Bere, Sachin. (2024). Music Recommendation System on Spotify Using Deep Learning. 8. 7. 10.55041/IJSREM38868.

Christensen, T. (1999). Four-hand piano transcription and geographies of nineteenth-century musical reception. *Journal of the American Musicological Society*, 52(2), 255–298. <https://doi.org/10.2307/831999>

Christensen, T. (Ed.). (2002). *The Cambridge History of Western Music Theory*. Cambridge University Press.

Chuan, C.-H., Agres, K., & Herremans, D. (2018). From Context to Concept: Exploring Semantic Relationships in Music with Word2Vec. arXiv:1811.12408

Chuan-Kang Ting, Memetic Computing, Overview. 2024: <https://link.springer.com/journal/12293>

Chvaja, R. (2020). Why Did Memetics Fail? Comparative Case Study. *Perspectives on Science*, 28(4), 542-570. https://doi.org/10.1162/posc_a_00350

Clark, A. (1997). *Being there: Putting brain, body, and world together again*. MIT Press.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204. <https://doi.org/10.1017/S0140525X12000477>

Clownboss. (2012, February 9). Montage Parodies. Know Your Meme. (Updated January 29, 2025 by LiterallyAustin).

<https://knowyourmeme.com/memes/montage-parodies>

Cohen, D. E. (2022). From Ramos to Rameau: Toward the origins of the modern concept of harmony. *Journal of Music Theory*, 66(1), 1–42. <https://doi.org/10.1215/00222909-9534115>

Cohen, J. M. (2009). Music Institutions and the Transmission of Tradition. *Ethnomusicology*, 53(2), 308-325

CoinMarketCap. (2025). Keyboard Cat (KEYCAT). Retrieved from <https://coinmarketcap.com/currencies/keyboard-cat/>

Constant, A., Clark, A., Kirchhoff, M., & Friston, K. J. (2022). Extended active inference: Constructing predictive cognition beyond skulls. *Mind & language*, 37(3), 373–394. <https://doi.org/10.1111/mila.12330>

Conte, R. (2000). Memes through (social) minds. In R. Aunger (Ed.), *Darwinizing culture: The status of memetics as a science* [chapter]. Oxford University Press.

Cook, N. (2001). Between process and product: Music and/as performance. *Music Theory Online*, 7(2). <https://www.mtosmt.org/issues/mto.01.7.2/mto.01.7.2.cook.html>

Corno, A. (2022). Il canto gregoriano: uniformità versus pluralità. *Stato, Chiesa E Pluralismo Confessionale*. <https://doi.org/10.54103/1971-8543/18611>

CoverNet.ai. (2024). Case Studies: Legal Battles Over AI-Generated Music. Retrieved from <https://www.covernet.ai/post/case-studies-legal-battles-over-ai-generated-music>

Covington, P., Adams, J., & Sargin, E. (2016). Deep neural networks for YouTube recommendations. *RecSys '16*

Coyle, J. (2009, May 22). 'Keyboard Cat' phenomenon spreads on web, TV. ABC News.

Crook, D. (2009). A sixteenth-century catalog of prohibited music. *Journal of the American Musicological Society*, 62(1), 1–78. <https://doi.org/10.1525/jams.2009.62.1.1>

Cros Vila, L., Sturm, B. L. T., Casini, L., & Dalmazzo, D. (2025). The AI Music Arms Race: On the Detection of AI-Generated Music. *Transactions of the International Society for Music Information Retrieval*, 8(1), 179-194. <https://doi.org/10.5334/tismir.254>

Cyranek. (2015a, August 18). MLG Darude Sandstorm [Video]. YouTube <https://www.youtube.com/watch?v=vg7ZrQKExlo>

Cyranek. (2015b, February 16). [Let it go – MLG airhorn remix] [Video]. YouTube. <https://www.youtube.com/watch?v=mVHJ6OwTYWc>

Cyranek. (2015c, November 14). Undertale Megalovania – MLG airhorn remix [Video]. YouTube. <https://www.youtube.com/watch?v=Cg-DWc9nPfQ>

D.L. Bowling, & D. Purves, A biological rationale for musical consonance, *Proc. Natl. Acad. Sci. U.S.A.* 112 (36) 11155-11160, <https://doi.org/10.1073/pnas.1505768112> (2015).

Damak, K., & Nasraoui, O. (2019). SeER: An explainable deep learning MIDI-based hybrid song recommender system. arXiv preprint arXiv:1907.01640. <https://arxiv.org/abs/1907.01640>

Darwin, C. (1871). *The descent of man, and selection in relation to sex* (Vol. 2). London: John Murray.

da Silva, P. (2003). *David Cope and Experiments in Musical Intelligence*, Spectrum press

Davis, C. (2013). *The soundtrack of my life*. Simon & Schuster.

Dawkins, R. (1982). *The extended phenotype*. Oxford University Press.

Dawkins, R. (2014, 2 lutego). Just for Hits Richard Dawkins [Wideo]. YouTube.
<https://www.youtube.com/watch?v=T5DOiZ8Y3bs>

Dawkins, R. 1976 *The Selfish Gene*

De, D., El Jamal, M., Aydemir, E., & Khera, A. (2025). Social media algorithms and teen addiction: Neurophysiological impact and ethical considerations. *Cureus*, 17(1), e77145.
<https://doi.org/10.7759/cureus.77145>

DeDonna, A. R. (2020). The Neapolitan Conservatories: Identity, Formation, and Operation. In *Instrumental Music in Late Eighteenth-Century Naples: Politics, Patronage and Artistic Culture* (pp. 27–58). chapter, Cambridge: Cambridge University Press.

Demis Hassabis – Facts – 2024. NobelPrize.org. Nobel Prize Outreach 2025. Tue. 26 Aug 2025. <<https://www.nobelprize.org/prizes/chemistry/2024/hassabis/facts/>>

Dennett, D. C. (1995). *Darwin's dangerous idea: Evolution and the meanings of life*. Simon & Schuster

Dennett, D. C. (2017). *From bacteria to Bach and back: The evolution of minds*. W. W. Norton & Company.

DeNora, T. (1997). *Beethoven and the Construction of Genius: Musical Politics in Vienna, 1792-1803*. University of California Press.

Deschamps, É. (2021, May). Story of the European Anthem. European Parliamentary Research Service.

Desrosiers, C., & Karypis, G. (2011). A comprehensive survey of neighborhood-based recommendation methods. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), *Recommender Systems Handbook* (pp. 107–144). Springer.

Deutsch, D. (2005, July). TEDGlobal 2005 transcript [Transcript]. David Deutsch. <https://www.daviddeutsch.org.uk/videos/tedglobal-2005-transcript>

Deutsch D, Marletto C. 2015 Constructor theory of information. *Proc. R. Soc. A* 471: 20140540. <http://dx.doi.org/10.1098/rspa.2014.0540>

Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. North American Chapter of the Association for Computational Linguistics.

Dictionary.com. (2018, March 1). Darude – Sandstorm meme: Meaning & history. Retrieved from <https://www.dictionary.com/e/memes/darude-sandstorm/>

Dirac, P. A. M. “The Evolution of the Physicist’s Picture of Nature.” *Scientific American*, Vol. 208, No. 5 (May 1963)

Douglas Blackiston et al. ,A cellular platform for the development of synthetic living machines. *Sci. Robot.*6,eabf1571(2021).DOI:10.1126/scirobotics.abf1571

Drott, E. (2020). Fake streams, listening bots, and click farms: Counterfeiting attention in the streaming music economy. *American Music*, 38(2), 153-175. <https://doi.org/10.5406/americanmusic.38.2.0153>

Du, W., Li, H., & Chen, L. (2022). Track2Vec: fairness music recommendation with a GPU-free customizable-driven framework. arXiv preprint arXiv:2210.16590. <https://arxiv.org/abs/2210.16590>

Duch W. (2021). Memetics and neural models of conspiracy theories. *Patterns* (New York, N.Y.), 2(11), 100353. <https://doi.org/10.1016/j.patter.2021.100353>

Eck, D., & Schmidhuber, J. (2002). A first look at music composition using LSTM recurrent neural networks (Technical Report No. IDSIA-12-02). Istituto Dalle Molle Di Studi Sull Intelligenza Artificiale.

Edelman, G. M. (1987). *Neural Darwinism: The theory of neuronal group selection*. Basic Books.

Edmonds, B. (2005). The revealed poverty of the gene-meme analogy – why memetics per se has failed to produce substantive results. *Journal of Memetics - Evolutionary Models of Information Transmission*, 9. http://cfpm.org/jom-emit/2005/vol9/edmonds_b.html

Ehrlich, C. (1995). *First Philharmonic: A History of the Royal Philharmonic Society*. Oxford: Clarendon Press

Eiben, A. E., & Smith, J. E. (2015). *Introduction to evolutionary computing* (2nd ed.). Springer.

Emirbayer, M., & Mische, A. (1998). What is agency? *American Journal of Sociology*, 103(4), 962–1023. <https://doi.org/10.1086/231294>

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors*, 37(2), 381–394. <https://doi.org/10.1518/001872095779064555>

Eno, B. (2011, October 11). Composers as gardeners [Conversation]. Edge. https://www.edge.org/conversation/brian_eno-composers-as-gardeners

Equalizer Magazine. (2013). Beethoven – Moonlight sonata for air horns [Audio track]. SoundCloud. <https://soundcloud.com/equalizer-mag/beethoven-moonlight-sonata>

Essinger, J. (2014). Ada's algorithm: How Lord Byron's daughter Ada Lovelace launched the digital age. Melville House.

European Union. (2025). Symbols of the European Union: European anthem. Retrieved from https://european-union.europa.eu/principles-countries-history/symbols/european-anthem_en

Fares, Murhaf; Kutuzov, Andrei; Oepen, Stephan & Velldal, Erik (2017). Word vectors, reuse, and replicability: Towards a community repository of large-text resources, In Jörg Tiedemann (ed.), Proceedings of the 21st Nordic Conference on Computational Linguistics, NoDaLiDa, 22-24 May 2017. Linköping University Electronic Press. ISBN 978-91-7685-601-7

Feltman, R., Parshall, A., & DeViscio, J. (2025, September 18). The Linguistic Science behind Viral Social Media Slang. Scientific American. <https://www.scientificamerican.com/podcast/episode/etymology-nerd-adam-aleksic-on-how-internet-culture-is-transforming-the-way/>

Fenlon, I. (Ed.). (2009). Music in medieval and early modern Europe: Patronage, sources and texts. Cambridge University Press

Ferrag, M. A., Tihanyi, N., & Debbah, M. (2025). From LLM Reasoning to Autonomous AI Agents: A Comprehensive Review. arXiv. <https://arxiv.org/abs/2504.19678>

Ficsor, M. (2002). *The law of copyright and the internet: The 1996 WIPO treaties, their interpretation and implementation*. Oxford University Press.

Findeisen, F. (2017). *The addiction formula: A holistic approach to writing captivating, memorable hit songs*. Findeisen.

Finer, Jem, 2000. Longplayer: <https://longplayer.org/>

Fitch, W. T. (2005). The evolution of language: A comparative review. *Biology and Philosophy*, 20(2), 193–230. <https://doi.org/10.1007/s10539-005-5597-1>

Floridi, L. AI as Agency without Intelligence: On Artificial Intelligence as a New Form of Artificial Agency and the Multiple Realisability of Agency Thesis. *Philos. Technol.* 38, 30 (2025). <https://doi.org/10.1007/s13347-025-00858-9>

Forsgren, S., & Martiros, H. (2022). Riffusion - Stable diffusion for real-time music generation

Fortuna, M. A., Zaman, L., Ofria, C., & Wagner, A. (2017). The genotype-phenotype map of an evolving digital organism. *PLOS Computational Biology*, 13(2), e1005414. <https://doi.org/10.1371/journal.pcbi.1005414>

Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., Heine-Suñer, D., Cigudosa, J. C., Urioste, M., Benitez, J., Boix-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z., Spector, T. D., Wu, Y. Z., Plass, C., ... Esteller, M. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the United States of America*, 102(30), 10604–10609. <https://doi.org/10.1073/pnas.0500398102>

Frank, S. A. (2014). The inductive theory of natural selection: Summary and synthesis [E-print]. arXiv. <https://arxiv.org/abs/1412.1285>

Frederick, R. (2025). Using AI for writing lyrics. MySongCoach. <https://mysongcoach.com/using-ai-for-writing-lyrics>

Frelon, Andrew, 2025. I am Andrew Frelon, the guy running the fake Velvet Sundown Twitter Medium. <https://medium.com/@andrew.frelon/i-am-andrew-frelon-the-guy-running-the-fake-velvet-sundown-twitter-fcab2b7e471b>

Frelon, Andrew, 2025. I am Andrew Frelon, the guy running the fake Velvet Sundown Twitter Medium. <https://medium.com/@andrew.frelon/i-am-andrew-frelon-the-guy-running-the-fake-velvet-sundown-twitter-fcab2b7e471b>

Friedlander, J. P. (2020). Year-End 2020 RIAA Revenue Statistics. riaa.com. <https://www.riaa.com/wp-content/uploads/2021/02/2020-Year-End-Music-IndustryRevenue-Report.pdf>.

Friston, K. The free-energy principle: a unified brain theory?. Nat Rev Neurosci 11, 127–138 (2010). <https://doi.org/10.1038/nrn2787>

Funk, T. (2018). A musical suite composed by an electronic brain: Reexamining the Illiac Suite and the legacy of Lejaren A. Hiller Jr.

Gabbatt, A. (2015, April 6). Easter egg-rolling: how a European game hatched on the White House lawn. The Guardian.

<https://www.theguardian.com/lifeandstyle/2015/apr/06/easter-egg-rolling-white-house-eu-rope-tradition>

GameStream. (2007, June 21). Call of Duty 4: Modern Warfare – Ingame gameplay! [Video]. YouTube. <https://www.youtube.com/watch?v=LO4caTLqqoM>

Ganesh, A., Liu, B., Peng, B., Du, L., Li, S., Yang, S., Yang, H., & Liu, J. (2024). SWaT: Statistical Modeling of Video Watch Time through User Behavior Analysis. Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.1. <https://doi.org/10.48550/arXiv.2408.07759>

Gatherer, D., 1997; Macromemetics: Towards a Framework for the Re-unification of Philosophy.

Journal of Memetics - Evolutionary Models of Information Transmission, 1.

http://cfpm.org/jom-emit/1997/vol1/gatherer_dg.html

Gerbault, P., Liebert, A., Itan, Y., Powell, A., Currat, M., Burger, J., Swallow, D. M., & Thomas, M. G. (2011). Evolution of lactase persistence: an example of human niche construction. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 366(1566), 863–877. <https://doi.org/10.1098/rstb.2010.0268>

Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration*. University of California Press.

Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, 8(3), 183–206. <https://doi.org/10.1515/cogl.1997.8.3.183>

Gjerdingen, R. O. (1988). *A classic turn of phrase: Music and the psychology of convention*. Philadelphia: University of Pennsylvania Press.

Gleick, J. (2011). *The Information: A History, a Theory, a Flood*. Pantheon Books.

Glynn, P. (2025, February 10). Kendrick Lamar and Drake beef: What's the latest? BBC News. <https://www.bbc.com/news/articles/cy8xmkmg207o>

Goehr, L. (1992). *The imaginary museum of musical works: An essay in the philosophy of music*. Clarendon Press.

Goldstein, A. Thrills in response to music and other stimuli. *Psychobiology* 8, 126–129 (1980). <https://doi.org/10.3758/BF03326460>

Göke, J. (2024). History Makes Memes. Memes Make History. In V. Schafer & F. Pailler (Ed.), *Online Virality: Spread and Influence* (pp. 61-78). Berlin, Boston: De Gruyter Oldenbourg. <https://doi.org/10.1515/9783111311371-004>

Gomez Sarmiento, I. (2025, August 8). AI-generated music is here to stay. Will streaming services like Spotify label it? NPR. <https://www.npr.org/2025/08/08/nx-s1-5492314/ai-music-streaming-services-spotify>

Goodrow, C. (2021, 15 września). On YouTube's recommendation system. YouTube Blog. Pobrane 6 czerwca 2025 z <https://blog.youtube/inside-youtube/on-youtubes-recommendation-system/>

Gottron, T., & Schwagereit, F. (2016). The Impact of the Filter Bubble - A Simulation Based Framework for Measuring Personalisation Macro Effects in Online Communities. ArXiv, abs/1612.06551.

Graving, T. (2015, August 7). Disney intro (MLG air horn remix) [Video]. YouTube. <https://www.youtube.com/watch?v=gDRAEktIJWA>

Gregory, D., & Able, A. (2021). They're the same picture. In Know Your Meme. Know Your Meme. Retrieved September 28, 2025, from <https://knowyourmeme.com/memes/theyre-the-same-picture>

Gregory, T.R. Artificial Selection and Domestication: Modern Lessons from Darwin's Enduring Analogy. *Evo Edu Outreach* 2, 5–27 (2009). <https://doi.org/10.1007/s12052-008-0114-z>

Grobler, A. (2006). *Metodologia nauk*. Kraków: Wydawnictwo Aureus.

Grout, D. J., & Palisca, C. V. (2001). *A history of Western music* (6th ed.). W. W. Norton.

Guez, T. (2023, May 24). Ghostwriter in the Machine: Copyright Implications of AI-Generated Imitations. *The National Law Review*. Retrieved from <https://natlawreview.com/article/ghostwriter-machine-copyright-implications-ai-generated-imitations>

Gumuskaya G, Srivastava P, Cooper BG, Lesser H, Semegran B, Garnier S, Levin M. Motile Living Biobots Self-Construct from Adult Human Somatic Progenitor Seed Cells. *Adv Sci (Weinh)*. 2024 Jan;11(4):e2303575. doi: 10.1002/advs.202303575. Epub 2023 Nov 30. PMID: 38032125; PMCID: PMC10811512.

Gutelle, S. (2015, April 1). YouTube adds a “Darude – Sandstorm” button for April Fools’ Day. *Tubefilter*

Gwenanenn. (2025). Major League Gaming. In *Aesthetics Wiki*. Fandom. https://aesthetics.fandom.com/wiki/Major_League_Gaming

Hagen, A. N. (2015). The Playlist Experience: Personal Playlists in Music Streaming Services. *Popular Music and Society*. <https://doi.org/10.1080/03007766.2015.1021174>

Hagen, Anja. (2016). Music Streaming the Everyday Life. 10.1057/978-1-137-58290-4_14.

Haiduk, A., & Fitch, W. T. (2022). Understanding design features of music and language: The choric/dialogic distinction. *Frontiers in Psychology*, 13, 786899. <https://doi.org/10.3389/fpsyg.2022.786899>

Hamilton, M., & Pearce, M. (2024). Trajectories and revolutions in popular melody based on U.S. charts from 1950 to 2023. *Scientific Reports*, 14, 14749. <https://doi.org/10.1038/s41598-024-64571-x>

Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016). Diachronic word embeddings reveal statistical laws of semantic change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1489–1501). Association for Computational Linguistics. <https://doi.org/10.18653/v1/P16-1141>

Harris, H. M. B., & Hill, C. (2021). A place for viruses on the tree of life. *Frontiers in Microbiology*, 11, Article 604048. <https://doi.org/10.3389/fmicb.2020.604048>

Hawckett, Andrew (2013) An empirical investigation of the concept of memes in music using mass data analysis of string quartets. Doctoral thesis, University of Huddersfield.

Haynes, G. (2018, March 20). Bento the Keyboard Cat, internet sensation and YouTube star, dies. *The Guardian*. <https://www.theguardian.com/technology/2018/mar/20/bento-keyboard-cat-youtube-dies>

Hennequin, R., Khlif, A., Voituret, F., & Moussallam, M. (2020). Spleeter: A fast and efficient music source separation tool with pre-trained models. *Journal of Open Source Software*, 5(50), 2154. DOI: 10.21105/joss.02154

Henrich, J. (2015). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.

Henry, A., Wiratama, V., Afilipoaie, A., Ranaivoson, H., & Arrivé, E. (2024). Impacts of AI on music consumption and fairness. *Emerging Media*, 2, 1–14. <https://doi.org/10.1177/27523543241269047>

Henshilwood, Christopher & Dubreuil, Benoît. (2011). The Still Bay and Howiesons Poort, 77–59 ka: Symbolic Material Culture and the Evolution of the Mind during the African Middle Stone Age. *Current Anthropology*. 52. 361-400. 10.1086/660022.

Hesmondhalgh, D. (2020). Is music streaming bad for musicians? Problems of evidence and argument. *New Media & Society*, 23(12), 3593-3615. <https://doi.org/10.1177/1461444820953541> (Original work published 2021)

Heylighen, F., and Chielens, K. (2009). Evolution of culture, memetics. In *Encyclopedia of Complexity and Systems Science*, B. Meyer, ed. (Springer), pp. 3205–3220

Higgins, C. (2016, February 8). Darude on making the only song in eSports. Red Bull. <https://www.redbull.com/ca-en/how-darude-feels-making-the-only-song-in-esports>

Hiller, L. A., & Isaacson, L. M. (1959). *Experimental music: Composition with an electronic computer*. McGraw-Hill.

Ho, J., Jain, A., & Abbeel, P. (2020). Denoising Diffusion Probabilistic Models. ArXiv, [abs/2006.11239](https://arxiv.org/abs/2006.11239).

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>

Hoeschele, M., & Fitch, W. T. (2022). Cultural evolution: Conserved patterns of melodic evolution across musical cultures. *Current Biology*, 32(6), R265–R267. <https://doi.org/10.1016/j.cub.2022.01.080>

Holoman, D. K. (2004). *The Société des Concerts du Conservatoire, 1828–1967*. Berkeley: University of California Press

Holpuch, A. (2013, February 19). Harlem Shake: Baauer cashes in on viral video's massive YouTube success. *The Guardian*. <https://www.theguardian.com/technology/2013/feb/19/harlem-shake-baauer-youtube-success>

Hootsuite Blog. (2025, 12 maja). How the TikTok algorithm ranks content in 2025 + tips for visibility. Pobrane 6 czerwca 2025 z <https://blog.hootsuite.com/tiktok-algorithm/>

Hootsuite Blog. Pobrane 6 czerwca 2025 z <https://blog.hootsuite.com/youtube-algorithm/>

Hoppin, R. H. (1978). *Medieval Music*. New York, NY: W. W. Norton & Company

Hordijk, W., Hein, J., & Steel, M. (2010). Autocatalytic Sets and the Origin of Life. *Entropy*, 12(7), 1733-1742. <https://doi.org/10.3390/e12071733>

Horowitz, S. J. (2024). Drake removes 'Taylor Made Freestyle,' featuring AI Tupac Shakur vocals, from social media after threat of lawsuit. *Variety*. <https://variety.com/2024/music/news/drake-removes-taylor-made-freestyle-tupac-shakur-lawsuit-1235983577/>

Hou, X., Zhao, Y., Wang, S., & Wang, H. (2025). Model Context Protocol (MCP): Landscape, security threats, and future research directions. arXiv preprint arXiv:2503.23278. <https://arxiv.org/abs/2503.23278>

Huang, J. (2025, March 18). GTC keynote with NVIDIA CEO Jensen Huang [Speech transcript]. Rev.

<https://www.rev.com/transcripts/gtc-keynote-with-nvidia-ceo-jensen-huang>

Hull, D. L. (1980). Individuality and selection. *Annual Review of Ecology and Systematics*, 11, 311–332. <https://doi.org/10.1146/annurev.es.11.110180.001523>

Huron, D. (2016). *Voice leading: The science behind a musical art*. MIT Press.

Huron D. *Sweet Anticipation: Music and the Psychology of Expectation*. Music Percept. 2007

Hypebot. (2025, July 11). AI music is exploding on Spotify: DON'T PANIC! Hypebot. <https://www.hypebot.com/hypebot/2025/07/ai-music-is-exploding-on-spotify.html>

Höpflinger, A.-K. (2025). Song review: The Beatles, “Now and Then”. *Journal of Religion, Film and Media*, 11(1), 155–159. DOI: 10.25364/05.11:2025.1.12.

Ingham, T. (2023, April 17). Universal Music Group responds to ‘fake Drake’ AI track: Streaming platforms have ‘a fundamental responsibility to prevent the use of their services in ways that harm artists’. *Music Business Worldwide*. <https://www.musicbusinessworldwide.com/universal-music-group-responds-to-fake-drake-ai-track-streaming-platforms-have-a-fundamental-responsibility-to-prevent-the-use-of-their-services-in-ways-that-harm-artists/>

Iskandar, K. L., Spil, T., & Bukhsh, F. (2025). AI and music: How do listeners and artists perceive it? In T. X. Bui (Ed.), *Proceedings of the 58th Hawaii International Conference on System Sciences (HICSS)* (pp. 6067–6076). University of Hawai‘i at Mānoa. <https://hdl.handle.net/10125/109335>

J. Bieschke, P. Weber, N. Sarafoff, M. Beekes, A. Giese, & H. Kretzschmar, Autocatalytic self-propagation of misfolded prion protein, *Proc. Natl. Acad. Sci. U.S.A.* 101 (33) 12207-12211, <https://doi.org/10.1073/pnas.0404650101> (2004).

Jablonka, E., & Lamb, M. J. (2005). *Evolution in four dimensions: Genetic, epigenetic, behavioral, and symbolic variation in the history of life*. Cambridge, MA: MIT Press.

Jacewicz, E., Alexander, J. M., & Fox, R. A. (2023). Extended high frequency in hearing and speech. *Acoustics Today*, 19(3), 22–29. <https://doi.org/10.1121/AT.2023.19.3.22>

Jackendoff, Ray, *Foundations of Language: Brain, Meaning, Grammar, Evolution* (Oxford, 2002; online edn, Oxford Academic, 1 Sept. 2007), <https://doi.org/10.1093/acprof:oso/9780198270126.001.0001>, accessed 10 July 2025.

Jakubowski, K., Finkel, S., Stewart, L., & Müllensiefen, D. (2017). Dissecting an earworm: Melodic features and song popularity predict involuntary musical imagery. *Psychology of Aesthetics, Creativity, and the Arts*, 11(2), 122–135. <https://doi.org/10.1037/aca0000090>

Jakubowski, K., Finkel, S., Stewart, L., & Müllensiefen, D. (2017). Dissecting an earworm: Melodic features and song popularity predict involuntary musical imagery. *Psychology of Aesthetics, Creativity, and the Arts*, 11(2), 122–135. <https://doi.org/10.1037/aca0000090>

Jan, S. (2004). The Illusory Mozart: Selfish Memes in the Priests' Marches from *Idomeneo* and *Die Zauberflöte*. *International Journal of Musicology*,

Jan, S. (2010). Memesatz contra Ursatz: memetic perspectives on the aetiology and evolution of musical structure. *Musicae Scientiae*

Jan, S. (2011). Music, memory, and memes in the light of Calvinian neuroscience. *Music Theory Online*

Jan, S. (2012). "The Heavens are Telling": a memetic-Calvinian reading of a Haydn chord progression. *Interdisciplinary Science Reviews*

Jan, S. (2015a) Memetic Perspectives on the Evolution of Tonal Systems. *Interdisciplinary Science Reviews*, 40 (2). pp. 145-167. ISSN 0308-0188

Jan, S. (2015b). A Memetic Analysis of a Phrase by Beethoven: Calvinian Perspectives on Similarity and Lexicon-Abstraction. *Psychology of Music*

Jan, S. (2015c). From Holism to Compositionality: Memes and the Evolution of Segmentation, Syntax and Signification in Music and Language

JAN, S. (2016). From holism to compositionality: memes and the evolution of segmentation, syntax, and signification in music and language. *Language and Cognition*, 8(4), 463–500. doi:10.1017/langcog.2015.1

Jan, S. (2024). Song and dance: a memetic angle on the evolution of musicality and music via case studies of a musemplex in Saint-Saëns and ABBA

Jan. S. (2007) *The Memetic of Music: Neo-Darwinian view of musical structure and culture*. Ashgate

Janik, V. M., and Slater, P. J. B. (1997). Vocal learning in mammals. *Advances in the Study of Behavior*, 26, 59–99.

Jansson, A. (2021). Beyond the platform: Music streaming as a site of logistical and symbolic struggle. *New Media & Society*, 25, 3203 - 3221. <https://doi.org/10.1177/14614448211036356>.

Jenkins, H. (2006). *Convergence Culture: Where Old and New Media Collide*. NYU Press. <http://www.jstor.org/stable/j.ctt9qffwr>

Johnson, G. (1997, 11 listopada). Undiscovered Bach? No, a computer wrote it. *The New York Times*.

Johnstone, H. D. (2020). The Academy of Ancient Music (1726–1802): Its History, Repertoire and Surviving Programmes. *Royal Musical Association Research Chronicle*, 51, 1–136. doi:10.1017/rrc.2019.1

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, 31(5), 559–575. doi:10.1017/S0140525X08005293

Jürgens U. (2009). The neural control of vocalization in mammals: a review. *Journal of voice : official journal of the Voice Foundation*, 23(1), 1–10. <https://doi.org/10.1016/j.jvoice.2007.07.005>

Kapatsinski, V., Olejarczuk, P., & Redford, M. A. (2017). Perceptual Learning of Intonation Contour Categories in Adults and 9- to 11-Year-Old Children: Adults Are More Narrow-Minded. *Cognitive science*, 41(2), 383–415. <https://doi.org/10.1111/cogs.12345>

Kaplan, A. M., & Haenlein, M. (2011). Two Hearts in Three-Quarter Time: How to Waltz the Social Media/Viral Marketing Dance. *Business Horizons*, 54, 253-263. <https://doi.org/10.1016/j.bushor.2011.01.006>

Karpathy, A. [@karpathy]. (2025, February 3). There's a new kind of coding I call "vibe coding," where you fully give in to the vibes, embrace exponentials [Post]. X (formerly Twitter). <https://x.com/karpathy/status/1886192184808149383>

Kemp, S. (2024). The time we spend on social media. DataReportal – Kepios. Pobrane z <https://wearesocial.com/us/blog/2024/01/digital-2024-5-billion-social-media-users>

Kenneth O. Stanley, Joel Lehman Why Greatness Cannot Be Planned The Myth of the Objective 2015

Khalili, J. (2024, December 18). The edgelord AI that turned a shock meme into millions in crypto. Wired. <https://www.wired.com/story/truth-terminal-goatse-crypto-millionaire/>

Killin, A. (2016). Rethinking music's status as adaptation versus technology: a niche construction perspective. *Ethnomusicology Forum*, 25(2), 210–233. <https://doi.org/10.1080/17411912.2016.1159141>

Klaes, A. (2025). PROFESSIONAL SINGERS' RIGHT OF PUBLICITY IN THE AGE OF GENERATIVE ARTIFICIAL INTELLIGENCE. *Iowa Law Review*, 110. Retrieved from <https://ilr.law.uiowa.edu/sites/ilr.law.uiowa.edu/files/2025-03/ILR-110-Klaes.pdf>

Knardal, P. S. (2020). "Orchestrating institutional complexity and performance management in the performing arts."

Knobel, M., & Lankshear, C. (2007). *A New Literacies Sampler*. Peter Lang.

Kolchinsky, A., & Wolpert, D. H. (2018). Semantic information, autonomous agency and nonequilibrium statistical physics. *Interface Focus*, 8(6), 20180041. <https://doi.org/10.1098/rsfs.2018.0041>

Koren, Y., Bell, R., i Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer*, 42(8), 30–37.

Kouzakova, M., van Baaren, R., & van Knippenberg, A. (2010). Lack of behavioral imitation in human interactions enhances salivary cortisol levels. *Hormones and Behavior*, 57(4–5), 421–426. <https://doi.org/10.1016/j.yhbeh.2010.01.011>

Kováč, Ladislav. (2000). Fundamental principles of cognitive biology. *Evolution and Cognition*. 1. 51-69.

Krishnan, N. (2025). AI agents: Evolution, architecture, and real-world applications. arXiv preprint arXiv:2503.12687. <https://doi.org/10.48550/arXiv.2503.12687>

Kronfeldner, M.E. Is cultural evolution Lamarckian?. *Biol Philos* 22, 493–512 (2007). <https://doi.org/10.1007/s10539-006-9037-7>

Krueger, J. (2014). Affordances and the musically extended mind. *Frontiers in Psychology*, 4, 1003. <https://doi.org/10.3389/fpsyg.2013.01003>

Krzyżanowski, W. (2024). कलियुग [Album]. Bandcamp. <https://prototapes.bandcamp.com/album/>

Krämer, B. (2011). The mediatization of music as the emergence and transformation of institutions. *International Journal of Communication*, 5, 559–579. <https://ijoc.org/index.php/ijoc/article/view/929>

Kuc. (2023, 6 listopada). Naukowiec: media społecznościowe działają na mózg jak hazard. *Business Insider Polska*. <https://businessinsider.com.pl/technologie/naukowiec-media-spoecznościowe-działają-na-mozg-jak-hazard/2hfgy5v>

Kujala, T., Partanen, E., Virtala, P., & Winkler, I. (2023). Prerequisites of language acquisition in the newborn brain. *Trends in Neurosciences*, 46(8), 637–650. <https://doi.org/10.1016/j.tins.2023.05.011>

Kuper, A. (2000). If memes are the answer, what is the question? In R. Aunger (Ed.), *Darwinizing culture: The status of memetics as a science*. Oxford University Press.

Kurzban, R., & Neuberg, S. L. (2005). Managing ingroup and outgroup relationships. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 653–675). Hoboken, NJ: Wiley.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91–196). Cambridge University Press.

Laland, K., Matthews, B. & Feldman, M.W. An introduction to niche construction theory. *Evol Ecol* 30, 191–202 (2016). <https://doi.org/10.1007/s10682-016-9821-z>

Laland, K. N., & Odling-Smee, J. (2000). Niche construction, biological evolution, and cultural change. *Behavioral and Brain Sciences*, 23(1) <https://doi.org/10.1017/S0140525X00002417>

Laland, K. N., & Odling-Smee, J. (2000). The evolution of the meme. In R. Aunger (Ed.), *Darwinizing culture: The status of memetics as a science* (pp. 121–141). Oxford University Press

Landauer, R. (1961). Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3), 183–191. <https://doi.org/10.1147/rd.53.0183>

Laozi. (1997). *Tao te ching* (D. C. Lau, Trans.). Penguin Books. (Original work published ca. 400 BCE)

Lasche, N. (2019, December 17). YouTube Music makes discovery more personal with playlists mixed for you. YouTube Official Blog. <https://blog.youtube/news-and-events/youtube-music-makes-discovery-more/>

Lasche, N. (2020, November 9). YouTube Music brings personalization to your everyday moods and moments. YouTube Official Blog. <https://blog.youtube/news-and-events/youtube-music-brings-personalization-your-every-day-moods-and-moments/>

Latour, B. (1988). *Science in action: How to follow scientists and engineers through society*. Harvard University Press.

Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press.

Law, J. (1992). Notes on the Theory of the Actor-Network: Ordering, Strategy, and Heterogeneity. *Systems Practice*, 5(4), 379-393.

Lee, D. (2020). Hornbostel-Sachs Classification of Musical Instruments. *Knowledge Organization*, 47(1), 72–91

Lee, J. (2024, April 10). New AI music generator Udio synthesizes realistic music on demand. *Ars Technica*. <https://arstechnica.com/information-technology/2024/04/new-ai-music-generator-udio-synthesizes-realistic-music-on-demand/>

Leman, M. (2007). *Embodied music cognition and mediation technology*. MIT Press.

Lerdahl, F. (1992). Cognitive Constraints on Compositional Systems. In *Contemporary Music Review*, 6(2), 97-121.

Levin, M. (2022). Technological Approach to Mind Everywhere: An Experimentally-Grounded Framework for Understanding Diverse Bodies and Minds. *Frontiers in Systems Neuroscience*, 16, 768201. <https://doi.org/10.3389/fnsys.2022.768201>

Levin, M. (2022). Technological approach to mind everywhere: An experimentally-grounded framework for understanding diverse bodies and minds. *Frontiers in Systems Neuroscience*, 16, 768201. <https://doi.org/10.3389/fnsys.2022.768201>

Levin, M., Friston, K., Jammingual, C. (2025) This Controversial Theory Shatters Our View of Reality... [Video] YouTube: <https://youtu.be/0yOV9Pzk2zw?si=bvs5W0ONpMqleSWu>

Levin, M. Darwin's agential materials: evolutionary implications of multiscale competency in developmental biology. *Cell. Mol. Life Sci.* 80, 142 (2023). <https://doi.org/10.1007/s00018-023-04790-z>

Levin, M. [Michael Levin's Academic Content]. (2025, July 5). Nothing in biology makes sense without teleology [Video]. YouTube. <https://www.youtube.com/watch?v=1gZw1SuykB8>

Levin M. Darwin's agential materials: evolutionary implications of multiscale competency in developmental biology. *Cell Mol Life Sci.* 2023 May 8;80(6):142. doi: 10.1007/s00018-023-04790-z. PMID: 37156924; PMCID: PMC10167196.

Levin M. Self-Improving Memory: A Perspective on Memories as Agential, Dynamically Reinterpreting Cognitive Glue. *Entropy*. 2024 C; 26(6):481. <https://doi.org/10.3390/e26060481>

Lewontin, R. C. (1970). The units of selection. *Annual Review of Ecology and Systematics*, 1, 1–18. <https://doi.org/10.1146/annurev.es.01.110170.000245>

Li, S., Zhang, Y., Tang, F., Ma, C., Dong, W., & Xu, C. (2024). Music style transfer with time-varying inversion of diffusion models. arXiv. <https://doi.org/10.48550/arXiv.2402.13763>

Limehouse, J. (2024, May 6). Obama weighed in on Kendrick Lamar, Drake rap battle 8 years ago: 'Gotta go with Kendrick'. *USA Today*. <https://eu.usatoday.com/story/entertainment/music/2024/05/06/barack-obama-kendrick-lamar-drake/73558440007/>

Lin, Y. (2021). A Survey on Reinforcement Learning for Recommender Systems. arXiv. <https://arxiv.org/abs/2109.10665>

Ling, C., AbuHilal, I., Blackburn, J., De Cristofaro, E., Zannettou, S., & Stringhini, G. (2021). Dissecting the Meme Magic: Understanding Indicators of Virality in Image Memes. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), Article 148. <https://doi.org/10.1145/3449155>

Ling, C., Blackburn, J., De Cristofaro, E., & Stringhini, G. (2021). Slapping cats, bopping heads, and oreo shakes: Understanding indicators of virality in TikTok short videos. arXiv. <https://doi.org/10.48550/arXiv.2111.02452>

Liu, B., Li, X., Hu, L., Wanzke, C., Käsdorf, B. T., Jan Hendrik, F., ... & Liu, W. (2020). Spontaneous Emergence of Self-Replicating Molecules. *Journal of the American Chemical Society*, 142(52), 21013–21021. <https://doi.org/10.1021/jacs.9b10796>

Liu, X., Zhou, Y., & Zhang, Y. (2025). The Effects of Landmark Saliency on Drivers' Spatial Cognition and Takeover Performance in Autonomous Driving Scenarios. *Behavioral Sciences*, 15(7), 966. <https://doi.org/10.3390/bs15070966>

Liu, Y. (2024). Deep learning based music recommendation systems: A review of algorithms and techniques. *Applied and Computational Engineering*, 109, 17–23. <https://www.ewadirect.com/proceedings/ace/article/view/17400>

Lonnberg, A., Xiao, P., & Wolfinger, K. (2020). The growth, spread, and mutation of internet phenomena: A study of memes.

Lovelace, A. A. (1843). Sketch of the Analytical Engine invented by Charles Babbage, translation with notes by A. A. Lovelace

Lu, Z. (2024). Personalized Marketing and Recommendation Systems on TikTok. *Advances in Economics, Management and Political Sciences*, 88, 46-50.

Lu, Z., Whalen, I., Boddeti, V., Dhebar, Y., Deb, K., Goodman, E., & Banzhaf, W. (2019). NSGA-Net: Neural architecture search using multi-objective genetic algorithm (arXiv preprint). arXiv. <https://doi.org/10.48550/arXiv.1810.03522>

Lukyanchikova, E., Askarbekuly, N., Aslam, H., & Mazzara, M. (2023). A Case Study on Applications of the Hook Model in Software Products. *Software*, 2(2), 292-309. <https://doi.org/10.3390/software2020014>

Lyon, P. (2025). Fundamental Principles of Cognitive Biology 2.0. *Biological Theory*. <https://doi.org/10.1007/s13752-025-00497-5>

Lyotard, J.-F. (1984). *The postmodern condition: A report on knowledge* (G. Bennington & B. Massumi, Trans.). University of Minnesota Press.

Ma, H., Zhang, Y., Shan, X., & Hu, X. (2025). Exploring the Impact of Artificial Intelligence on the Creativity Perception of Music Practitioners. *Journal of Intelligence*, 13(4), 47. <https://doi.org/10.3390/jintelligence13040047>

Maasø, A., & Hagen, A. N. (2019). Metrics and decision-making in music streaming. *Popular Communication*, 18(1), 18–31. <https://doi.org/10.1080/15405702.2019.1701675>

MacCallum, R. M., Mauch, M., Burt, A., & Leroi, A. M. (2012). Evolution of music by public choice. *Proceedings of the National Academy of Sciences of the United States of America*, 109(30), 12081–12086. <https://doi.org/10.1073/pnas.1203182109>

Macready, H., Stanton, L. (2025, 14 lutege). How the YouTube algorithm works in 2025.

Magarey, R. D., & Trexler, C. M. (2020). Information: A missing component in understanding and mitigating social epidemics. *Humanities and Social Sciences Communications*, 7, Article 128. <https://doi.org/10.1057/s41599-020-00620-w>

Magnusson, T. (2019). *Sonic writing: Technologies of material, symbolic, and signal inscriptions*. Bloomsbury Academic.

Maheshwari, C. (2023). Music recommendation on Spotify using deep learning. arXiv preprint arXiv:2312.10079. <https://arxiv.org/pdf/2312.10079>

Manning, P. (2013). The digital audio workstation. In *Electronic and computer music* (4th ed., pp. 319-344). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199746392.003.0021>

Marjeh, R., Harrison, P.M.C., Lee, H. et al. Timbral effects on consonance disentangle psychoacoustic mechanisms and suggest perceptual origins for musical scales. *Nat Commun* 15, 1482 (2024). <https://doi.org/10.1038/s41467-024-45812-z>

Marletto C. 2015 Constructor theory of life. *J. R. Soc. Interface* 12: 20141226.
<http://dx.doi.org/10.1098/rsif.2014.1226>

Marshall, S. M., Moore, D. G., Murray, A. R. G., Walker, S. I., & Cronin, L. (2022). Formalising the Pathways to Life Using Assembly Spaces. *Entropy*, 24(7), 884.
<https://doi.org/10.3390/e24070884>

Masci, M. (2022). Reconstructing the Paris Conservatory's Cours d'Harmonie (1796–1880). *Music Theory Online*, 28(4).
<https://mtosmt.org/issues/mto.22.28.4/mto.22.28.4.masci.html>

Masood, Adnan, 2025. Thinking Machines: How Multimodal Reasoning AI Will Transform Enterprise Decision-Making
<https://medium.com/@adnanmasood/thinking-machines-how-multimodal-reasoning-ai-will-transform-enterprise-decision-making-fc43f9658b58>

Mayr, E. (1961). Cause and effect in biology. *Science*, 134(3489), 1501–1506.
<https://doi.org/10.1126/science.134.3489.1501>

Mazur, M. (1966). *Cybernetyczna teoria układów samodzielnych*. Warszawa: Państwowe Wydawnictwo Naukowe.

McAllister, J. (1996). Scientists' Aesthetic Preferences Among Theories: Conservative Factors in Revolutionary Crises. , 169-187.
https://doi.org/10.1007/978-94-009-1786-6_8.

McDermott, J. (2005). The origins of music: Innateness, uniqueness, and evolution. *Music Perception*, 23(1), 29–59. <https://doi.org/10.1525/mp.2005.23.1.29>

McDonald, T. M., Maystre, L., Lalmas, M., Russo, D., & Ciosek, K. (2023). Optimizing for the Long-Term Without Delay. Spotify Research. <https://research.atspotify.com/2023/07/optimizing-for-the-long-term-without-delay>

McFedries, P. (2014, May 22). Stop, attention thief! IEEE Spectrum. Institute of Electrical and Electronics Engineers. <https://spectrum.ieee.org/stop-attention-thief>

McInerney, J., Lacker, B., Hansen, S., Higley, K., Bouchard, H., Gruson, A., & Mehrotra, R. (2018). Explore, exploit, explain: Personalizing explainable recommendations with bandits. In Proceedings of the 12th ACM Conference on Recommender Systems (pp. 31–39). ACM.

McLean, A. F., Jr. (1965). American vaudeville as ritual. University Press of Kentucky.

McMillen, P., Levin, M. Collective intelligence: A unifying concept for integrating biology across scales and substrates. *Commun Biol* 7, 378 (2024). <https://doi.org/10.1038/s42003-024-06037-4>

McNamara, A. (2011). Can we measure memes? *Frontiers in Evolutionary Neuroscience*, 3, 1-7. <https://doi.org/10.3389/fnevo.2011.00001>

Mellace, R. (2014). From Court Chapel to public concert: Hasse's oratorios from Dresden to Vienna. *Musicologica Brunensia*, 49(1), [235]-249. <https://hdl.handle.net/11222.digilib/130214>

Menary, R. (2010). Introduction to the special issue on 4E cognition. *Phenomenology and the Cognitive Sciences*, 9(4), 459–463. <https://doi.org/10.1007/s11097-010-9187-6>

Mercer, A. (2014, December 28). Air Horn Remixes. Know Your Meme. (Updated January 29, 2025 by LiterallyAustin). <https://knowyourmeme.com/memes/air-horn-remixes>

Merchel, S., & Altinsoy, M. E. (2020). Psychophysical comparison of the auditory and tactile perception: A survey. *Journal on Multimodal User Interfaces*, 14, 271–283. <https://doi.org/10.1007/s12193-020-00333-z>

Mesoudi, A. (2018). Cultural evolution. In eLS. John Wiley & Sons. <https://doi.org/10.1002/9780470015902.a0028231>

Meyer, L.B. 1956 *Emotion and Meaning in Music* (Chicago: University of Chicago Press).

Michael Levin, EISM. EISM Interviews. 2022. <https://eism.eu/michael-levin/>

Michael Levin, Maggie Ciskanik. (2024, 10 stycznia). Intelligence without a brain. John Templeton Foundation. <https://www.templeton.org/news/intelligence-without-a-brain>

Michael Levin 2024 B. Patterns are alive and we are living patterns: <https://iai.tv/articles/patterns-are-alive-and-we-are-living-patterns-auid-2919>

Michelsanti, D., Tan, Z.-H., Zhang, S.-X., Xu, Y., Yu, M., Yu, D., & Jensen, J. (2021). An overview of deep-learning-based audio-visual speech enhancement and separation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. DOI: 10.1109/TASLP.2021.3066303

Miller, D. M. (2022). MTV. In EBSCO Research Starters. EBSCO Information Services

Miller, G. (2000). *The mating mind: How sexual choice shaped the evolution of human nature*. New York: Doubleday.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97. <https://doi.org/10.1037/h0043158>

Miton, H., Wolf, T., Vesper, C., Knoblich, G., & Sperber, D. (2020). Motor constraints influence cultural evolution of rhythm: Transmission chain experiment. *Proceedings of the Royal Society B*, 287(1937), 20202001. <https://doi.org/10.1098/rspb.2020.2001>

Mitsufuji, Y., Stöter, F.-R., Fabbro, G., Kitamura, D., Uhlich, S., Takahashi, N., ... Fontecchio, M. (2022). Music demixing challenge 2021: Track-wise preliminary results. *Frontiers in Signal Processing*, 1, 808395. DOI: 10.3389/frsip.2021.808395

Mol, J., Wijnberg, N. M., Carroll C. (2005). Value chain envy: explaining new entry and vertical integration in popular music. *Journal of Management Studies*,

Moliner, E., & Välimäki, V. (2024). Diffusion-based audio inpainting. *Journal of the Audio Engineering Society*, 72(3), 100–113. DOI: 10.17743/jaes.2022.0129

Monge-Roffarello, A., Baroni, L., Salgado, L. C. de C., & Pereira, R. (2022). Towards Understanding the Dark Patterns That Steal Our Attention. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2), Article 377.

Moreira D, López-García P. Ten reasons to exclude viruses from the tree of life. *Nat Rev Microbiol*. 2009 Apr;7(4):306-11. doi: 10.1038/nrmicro2108. Epub 2009 Mar 9. PMID: 19270719.

Morris, J. W. (2015). Curation by code: Infomediaries and the data-mining of taste. *European Journal of Cultural Studies*, 18(4–5), 446–463. <https://doi.org/10.1177/1367549415577387>

Morris, Jeremy Wade (2015) *Selling Digital Music, Formatting Culture*. University of California Press, Berkeley, CA, U.S.A.

Morrissey, N. (2021). *Metamodernism and vaporwave: A study of Web 2.0*. *Nota Bene Journal*.

Moscato, P. (2003). A gentle introduction to memetic algorithms. In F. Glover & G. Kochenberger (Eds.), *Handbook of metaheuristics* (pp. 105–144). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-48056-5_5

Mousavi, P., Maimon, G., Moumen, A., Petermann, D., Shi, J., Wu, H., Yang, H., Kuznetsova, A., Ploujnikov, A., Marxer, R., Ramabhadran, B., Elizalde, B., Lugosch, L., Li, J., Subakan, C., Woodland, P., Kim, M., Lee, H., Watanabe, S., Adi, Y., & Ravanelli, M. (2025). Discrete audio tokens: More than a survey! arXiv. <https://arxiv.org/abs/2506.10274>

Mullen, M. (2024, October 30). “Now I realize how little you really need to create good music when you just pick the right sounds – whether that’s on purpose or by accident”: How Darude created Sandstorm, the meme-worthy trance anthem named after a synth preset. MusicRadar. <https://www.musicradar.com/artists/when-you-turn-on-the-roland-jp-8080-the-first-sound-that-comes-up-is-called-sandstorm-how-darude-created-the-era-defining-trance-anthem-thats-named-after-a-synth-preset>

Multi.life. (2025, 21 stycznia). W sidłach dopaminy. Dlaczego media społecznościowe nas uzależniają. Multi.life. <https://multi.life/artykul/w-sidlach-dopaminy-dlaczego-media-spolecznosciowe-nas-uzalezniaja>

Murray, Elisabeth A., and others, *The Evolutionary Road to Human Memory* (Oxford, 2019; online edn, Oxford Academic, 23 Jan. 2020), <https://doi.org/10.1093/oso/9780198828051.001.0001>, accessed 6 June 2025.

Murtagh J., 2024, Kamień węgielny informatyki, „Scientific American. Polska edycja” (398), październik 2024, ISSN 0867-6380

Mon, Y.-J. (2025). LSTM-Based Music Generation Technologies. *Computers*, 14(6), 229. <https://doi.org/10.3390/computers14060229>

N. Krasnogor and J. Smith, "A tutorial for competent memetic algorithms: model, taxonomy, and design issues," in *IEEE Transactions on Evolutionary Computation*, vol. 9, no. 5, pp. 474-488, Oct. 2005, doi: 10.1109/TEVC.2005.850260.

Narmour, E. (1984). Toward an analytical symbology: The melodic, harmonic, and durational functions of implication and realization. In M. Baroni & L. Callegari (Eds.), *Musical Grammars and Computer Analysis: Atti del convegno* (Modena, 4-6 ottobre 1982) (*Quaderni della Rivista italiana di musicologia*, 8, pp. 83-114). Olschki.

National Collegiate Athletic Association. (2021, July 19). NCAA playing rules summary: Artificial crowd noise. https://ncaaorg.s3.amazonaws.com/championships/resources/rules/RULES_Artificial_Crowd_Noise_Summary.pdf

Negus, Keith. (1999) *Music Genres and Corporate Cultures*. London: Routledge

Neri, F., & Cotta, C. (2012). Memetic algorithms and memetic computing optimization: A literature review. *Swarm and Evolutionary Computation*, 2(1), 1–14. <https://doi.org/10.1016/j.swevo.2011.11.003>

Nestor Maslej, Loredana Fattorini, Raymond Perrault, Yolanda Gil, Vanessa Parli, Njenga Kariuki, Emily Capstick, Anka Reuel, Erik Brynjolfsson, John Etchemendy,

Katrina Ligett, Terah Lyons, James Manyika, Juan Carlos Niebles, Yoav Shoham, Russell Wald, Tobi Walsh, Armin Hamrah, Lapo Santarlaschi, Julia Betts Lotufo, Alexandra Rome, Andrew Shi, Sukrut Oak. "The AI Index 2025 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2025.

Nettl, B. (2005). *The study of ethnomusicology: Thirty-one issues and concepts*. University of Illinois

Newen, A., De Bruin, L., & Gallagher, S. (2018). *The Oxford handbook of 4E cognition*. Oxford University Press

Nippon. (2013, January 29). Can't Help But Sing of Joy. Retrieved from <https://www.nippon.com/en/views/b02401/>

Nogales, A., Donaher, S., & García-Tejedor, Á. J. (2023). A deep learning framework for audio restoration using convolutional/deconvolutional deep autoencoders. *Expert Systems with Applications*, 230, 120586. DOI: 10.1016/j.eswa.2023.120586

Nowak, R., & Whelan, A. (2018). "Vaporwave Is (Not) a Critique of Capitalism": Genre work in an online music scene. *Open Cultural Studies*, 2(1), 451-462. <https://doi.org/10.1515/culture-2018-0041>

November, N. (2021, June 11). Marketing orchestral music in the domestic sphere in early nineteenth-century Vienna: The Beethoven arrangements published by Sigmund Anton Steiner. *Musicologica Austriaca: Journal for Austrian Music Studies*. <https://www.musau.org/parts/neue-article-page/view/108>

Nussbaum, C. O. (2007). *The musical representation: Meaning, ontology, and emotion*. MIT Press.

Odilov, S. (2024, January 11). Can AI become your next CEO? Forbes. <https://www.forbes.com/sites/sherzododilov/2024/01/11/can-ai-become-your-next-ceo>

Odling-Smee, F. & Lala, Kevin & Feldman, Marcus. (2003). Niche Construction: The Neglected Process in Evolution. 10.1515/9781400847266.

Odling-Smee FJ (1988) Niche constructing phenotypes. In: Plotkin (ed) The role of behavior in evolution.

MIT Press, Cambridge

Odling-Smee FJ, Laland KN, Feldman MW (1996) Niche construction. *Am Nat* 147:641–648

Odling-Smee FJ, Laland KN, Feldman MW (2003) Niche construction: the neglected process in evolution. In: *Monographs in population biology*, vol 37. Princeton University Press, Princeton

Okada, T., & Ishibashi, H. (2017). Cognitive process of creative drawing by copying others' artworks. *Cognitive Science*, 41(5), 1201-1231. <https://doi.org/10.1111/cogs.12442>

Ong, W. (1982). *Orality and Literacy: The Technologizing of the Word*. Routledge. <https://doi.org/10.4324/9780203328064>

Ong, W. J. (1982). *Orality and Literacy: The Technologizing of the Word*.

Ong, Y. S., & Keane, A. J. (2004). Meta-Lamarckian learning in memetic algorithms. *IEEE Transactions on Evolutionary Computation*, 8(2), 99–110. <https://doi.org/10.1109/TEVC.2003.819944>

OpenAI (2019). MuseNet. Retrieved April 25, 2019, from <https://openai.com/index/musenet/>

Overall-Estate1349. (2022, November). The 2013-2017 / MLG era of memes. Kinda more Gen Z than Zillennial, but I think Zillennials also laughed at some of these [Comment on the post The 2013-2017/MLG era of memes]. Reddit. <https://www.reddit.com/r/Zillennials/comments/11b68t9>

Ozzi, D. (2022). *Sellout: The Major-Label Feeding Frenzy that Swept Punk, Emo, and Hardcore 1994-2007*. HarperCollins.

Paakki, H., Vepsäläinen, H. & Salovaara, A. Disruptive online communication: How asymmetric trolling-like response strategies steer conversation off the track. *Comput Supported Coop Work* 30, 425–461 (2021). <https://doi.org/10.1007/s10606-021-09397-1>

Pace, I. (2012). Instrumental performance in the nineteenth century. In: Lawson, C. & Stowell, R. (Eds.), *The Cambridge History of Musical Performance*. (pp. 643-695). Cambridge University Press. doi: 10.1017/CHOL9780521896115.027

Pahwa, N. (2023, April 17). How an AI-Generated Drake and Weeknd Song Became a Bizarre Hit. *Slate*. Retrieved from <https://slate.com/technology/2023/04/drake-weeknd-ai-single-ghostwriter-tiktok-heart-on-my-sleeve.html>

Paolizzo, F., & Johnson, C. (2020). Creative autonomy in a simple interactive music system. *Journal of New Music Research*, 49(1), 1–11. <https://doi.org/10.1080/09298215.2019.1709510>

Pareño, E. (2025, July 18). AI Music Enters its Band Era With The Velvet Sundown. *Rolling Stone Philippines*. <https://www.rollingstonephilippines.com/music/ai-music-the-velvet-sundown>

Parncutt, R., & Hair, G. (2018). A psychocultural theory of musical interval: Bye bye Pythagoras. *Music Perception*, 35(4), 475–501. <https://doi.org/10.1525/mp.2018.35.4.475>

Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press. <http://www.jstor.org/stable/j.ctt13x0hch>

Pastushkov Dmitry, (2024). *The Mechanics of Recording Industry: A Brief History & Its Functions*. <https://soundcharts.com/blog/mechanics-of-the-recording-industry>

Patel, A. D. (2007). *Music, Language, and the Brain*. Oxford University Press.

Pat_Wingman [Patricia Wong]. (2024, July 26). Patricia Wong | Mom & Foodie [Video]. TikTok. https://www.tiktok.com/@pat_wingman/video/7396050485112884485

Paukner, A., Suomi, S. J., Visalberghi, E., & Ferrari, P. F. (2009). Capuchin monkeys display affiliation towards humans who imitate them. *Science*, 325(5942), 880–883. <https://doi.org/10.1126/science.1176269>

Paul, Katie, 2024. Autonomous agents and profitability to dominate AI agenda in 2025, executives forecast <https://www.reuters.com/technology/artificial-intelligence/autonomous-agents-profitability-dominate-ai-agenda-2025-executives-forecast-2024-12-12/>

Pawlak, A. (2020). Zapomniany geniusz. Prof. Marian Mazur i jego szkoła cybernetyki. *e-mentor*, 5(87), 5–14. <https://doi.org/10.15219/em87.1488>

Payne, C. (2019). MuseNet. OpenAI. <https://openai.com/blog/musenet>

Pearce, M.T. (2005). The construction and evaluation of statistical models of melodic structure in music perception and composition. (Unpublished Doctoral thesis, City University London)

Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A., & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *NeuroImage*, 50(1), 302–313. <https://doi.org/10.1016/j.neuroimage.2009.12.019>

Pesic, A. (2021). Concerts and inadvertent secularization: Religious music in the entertainment market of eighteenth-century Paris. *Past & Present*, 250(1), 135–169. <https://doi.org/10.1093/pastj/gtaa011>

Peterson, R. A., & Berger, D. G. (1975). Cycles in Symbol Production: The Case of Popular Music. *American Sociological Review*

PhantomPhaze. (2014, June 27). Gorillaz – Feel good inc. air horn remix [Video]. YouTube. <https://www.youtube.com/watch?v=TaC3IRPNsdU>

Phillips, D. J. (2013). *Shaping jazz: Cities, labels, and the global emergence of an art form*. Princeton University Press. <https://doi.org/10.23943/princeton/9780691150888.001.0001>

Phillips-Silver, J., & Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition*, 67(1), 94–102. <https://doi.org/10.1016/j.bandc.2007.11.007>

Planchart, A. E. (2018). Institutions and Foundations. In M. Everist & T. F. Kelly (Eds.), *The Cambridge History of Medieval Music* (pp. 627–673). chapter, Cambridge: Cambridge University Press.

Plotkin, H. (2000). Culture and psychological mechanisms. In R. Aunger (Ed.), *Darwinizing culture: The status of memetics as a science* [chapter]. Oxford University Press.

Podlipniak, P. (2007). *Uniwersalia muzyczne*. Wydawnictwo Poznańskiego Towarzystwa Przyjaciół Nauk.

Podlipniak, P. (2015). *Instykt tonalny: Koncepcja ewolucyjnego pochodzenia tonalności muzycznej*. Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza.

Pollock, E. (2025, 26 maja). What you need to know about the TikTok algorithm to go viral in 2025. Agorapulse. Pobrane 6 czerwca 2025 z <https://www.agorapulse.com/blog/tiktok/tiktok-algorithm/>

Popa, E. (2019). Artificial life and 'nature's purposes': The question of behavioral autonomy. *Human Affairs*, 30, 587 - 596. <https://doi.org/10.1515/humaff-2020-0052>.

Pram, L., & Morreale, F. (2025). Opening musical creativity? Embedded ideologies in generative-AI music systems. arXiv. <https://doi.org/10.48550/arXiv.2508.08805>

Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon*, 9(5), 1–6. <https://doi.org/10.1108/10748120110424816>

Proust J. (1995). Intentionality and evolution. *Behavioural processes*, 35(1-3), 275–286. [https://doi.org/10.1016/0376-6357\(95\)00057-7](https://doi.org/10.1016/0376-6357(95)00057-7)

Qin, Y., Hu, S., Lin, Y., Chen, W., Ding, N., Cui, G., Zeng, Z., Huang, Y., Xiao, C., Han, C., Fung, Y. R., Su, Y., Wang, H., Qian, C., Tian, R., Zhu, K., Liang, S., Shen, X., Xu, B., Zhang, Z., Ye, Y., Li, B., Tang, Z., Yi, J., Zhu, Y., Dai, Z., Yan, L., Cong, X., Lu, Y., Zhao, W., Huang, Y., Yan, J., Han, X., Sun, X., Li, D., Phang, J., Yang, C., Wu, T., Ji, H., Liu,

Z., & Sun, M. (2023). Tool learning with foundation models [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2304.08354>

Qiu, X., Sun, T., Xu, Y. et al. Pre-trained models for natural language processing: A survey. *Sci. China Technol. Sci.* 63, 1872–1897 (2020). <https://doi.org/10.1007/s11431-020-1647-3>

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. OpenAI. https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI.

Rafii, Z., Liutkus, A., Stöter, F.-R., Mimilakis, S. I., FitzGerald, D., & Pardo, B. (2018). An overview of lead and accompaniment separation in music. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26(8), 1307–1335. DOI: 10.1109/TASLP.2018.2825440

Rameau, J.-P. (1726). *Nouveau système de musique théorique*. Paris: J.-B.-C. Ballard.

Raoult D. There is no such thing as a tree of life (and of course viruses are out!). *Nat Rev Microbiol.* 2009 Aug;7(8):615; author reply 615. doi: 10.1038/nrmicro2108-c6. PMID: 19561623.

Ravignani, A., Bowling, D. L., & Fitch, W. T. (2021). The vocal learning and rhythmic synchronization hypothesis: A critical review and the relevance of vocal plasticity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1835), 20200323.

Ray, M. (2025, August 27). Rick Rubin. In *Encyclopædia Britannica*. Retrieved September 29, 2025, from <https://www.britannica.com/biography/Rick-Rubin>

Reed, R. (2014, March 3). Watch Jimmy Kimmel's "Ameowadeus" parody and other post-Oscar trailers: The host envisioned viral videos as blockbusters at last night's post-Oscars special. *Rolling Stone*.
<https://www.rollingstone.com/tv-movies/tv-movie-news/watch-jimmy-kimmels-ameowadeus-parody-and-other-post-oscar-trailers-87857>

Reed, R. (2023, May 2). AI created a song mimicking the work of Drake and The Weeknd. What does that mean for copyright law? Harvard Law School. Retrieved from <https://hls.harvard.edu/today/ai-created-a-song-mimicking-the-work-of-drake-and-the-weeknd-what-does-that-mean-for-copyright-law>

Reeves, K. (2009, April 10). Breaking: Play him off, Keyboard Cat is our new obsession. *Urlesque*.

Rehfeldt, R. A. (2020). *The Beethoven Revolution: A Case Study in Selection by Accident*. PMC.

Reynolds, C. A. (1997). *Papal Patronage and the Music of St. Peter's, 1380-1513*. University of California Press

Ricci, F., Rokach, L., i Shapira, B. (Red.). (2022). *Recommender systems handbook* (Wyd. 3). Springer.

Richerson, P. J., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. University of Chicago Press.

Richerson, P. J., Boyd, R., & Henrich, J. (2010). Gene–culture coevolution in the age of genomics. *Proceedings of the National Academy of Sciences of the United States of America*, 107(Supplement 2), 8985–8992. <https://doi.org/10.1073/pnas.0914631107>

Riemann, H. (1929). *Handbuch der Harmonielehre* (10th ed.). Leipzig: Breitkopf & Härtel.

Rion, M. (2004). *Die Idee der Verbindung von Musik und Poesie im Frankreich des 16. Jahrhunderts*. Ludwig-Maximilians-Universität zu München

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688. <https://doi.org/10.1016/j.tics.2016.07.002>

Roach, J. C., Glusman, G., Smit, A. F. A., Huff, C. D., Hubley, R., Shannon, P. T., Rowen, L., Pant, K. P., Goodman, N., Bamshad, M., Shendure, J., Drmanac, R., Jorde, L. B., Hood, L., & Galas, D. J. (2010). Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science*, 328(5978), 636–639. <https://doi.org/10.1126/science.1186802>

Roesner, F., Redmiles, E. M., Zannettou, S., Nemeth, O., Ayalon, O., Goetzen, A., & Gummadi, K. P. (2024). Analyzing User Engagement with TikTok’s Short-Format Video Recommendations Using Data Donations. *arXiv*.

Romana Gładych, 2017: Mikrozarządzenie – makroproblem? <https://hrpolska.pl/prawo-pracy/czytelnia/mikrozarzadzanie-makroproblem.html>

Rosati DP, Woolhouse MH, Bolker BM, Earn DJD. 2021 Modelling song popularity as a contagious process. *Proc. R.Soc. A* 477: 20210457. <https://doi.org/10.1098/rspa.2021.0457>

Rothenbuhler, E. W., & McCourt, T. (2003). *The Economics of the Sound Recording Industry*. Routledge

rovic.keeve2001 [@rovic.keeve2001]. (2020, March 13). [Covid lies exposed] [Video]. TikTok. <https://www.tiktok.com/@rovic.keeve2001/video/7512395450910100743>

Rowland, David (2019). Music Publishing in Britain ca. 1840-1900. In: Sala, Massimiliano ed. *Music and the Second Industrial Revolution. Music, Science & Technology* (2). Turnhout: Brepols.

Royal Society of Edinburgh (2023) Artificial intelligence doesn't yet exist, optimised search does
<https://rse.org.uk/resource/artificial-intelligence-doesnt-yet-exist-optimised-search-does/>

rraaaarrrl. (2024, September 20). Bah! — I loved when AI art could never be anything but AI... [Tumblr post]. Tumblr. <https://rraaaarrrl.tumblr.com/post/762192338374311936>

Rubin, R. (2025). The way of code: The timeless art of vibe coding. Anthropic. <https://www.thewayofcode.com/>

Rubinoff, K. R. (2017). Toward a revolutionary model of music pedagogy: The Paris Conservatoire, Hugot and Wunderlich's *Méthode de flûte*, and the disciplining of the musician. *Journal of Musicology*, 34(4), 473–514. <https://doi.org/10.1525/jm.2017.34.4.473>

Rumelhart, D. (1980) Schemata: The Building Blocks of Cognition. In: Spiro, R., Bruce, B. and Brewer, W., Eds., *Theoretical Issues in Reading Comprehension*, Erlbaum Associates, Mahway, 33-58.

Rutledge, A. N. G. (2011). *Zelter, Goethe and the emergence of a German choral canon* (Doctoral dissertation). University of Toronto.

Różalska, E. (2021, 20 października). „Chopin Forever” – wyjątkowa cyfrowa retrospektywa na Google Arts & Culture. Blog Google Polska. https://blog.google/intl/pl-pl/nowosci-produktowe/odkrywanie-wyszukiwanie/2021_10_chopin-forever-wyjatkowa-cyfrowa/

Sachs, C. (1943). *The rise of music in the ancient world, east and west*. New York: W. W. Norton.

Stocking, G. W. (1982). *Race, culture, and evolution: Essays in the history of anthropology*. Chicago: University of Chicago Press.

Chatterjee, A., & Cardillo, E. R. (Eds.). (2022). *Brain, beauty, and art: Essays bringing neuroaesthetics into focus*. Oxford University Press. <https://doi.org/10.1093/oso/9780197513620.001.0001>

Frank, S. A. (2024). *Natural selection at multiple scales*. arXiv. <https://doi.org/10.48550/arXiv.2410.18732>

Miller, G. F. (2000). Evolution of human music through sexual selection. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 329–360). MIT Press.

Hofstadter, R. (1955). *Social Darwinism in American thought*. Boston: Beacon Press.

Hurford, J. R. (1990). Nativist and functional explanations in language acquisition. In I. M. Roca (Ed.), *Logical Issues in Language Acquisition*

Fitch, W. T. (2008). Co-evolution of phylogeny and glossogeny: There is no “logical problem of language evolution.” *Behavioral and Brain Sciences*, 31(5), 521–522. [doi:10.1017/S0140525X08005128](https://doi.org/10.1017/S0140525X08005128)

Saravanou, A., Tomasi, F., Mehrotra, R., & Lalmas, M. (2021). Multi-Task Learning of Graph-based Inductive Representations of Music Content. In Proceedings of the 22nd International Society for Music Information Retrieval Conference (ISMIR 2021), 602–609.

Savage, P. E. (2025, June 11). Comparative musicology: Evolution, universals, and the science of the world's music. https://doi.org/10.31234/osf.io/b36fm_v16

Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *PNAS*, 112(29), 8987–8992.

Savage, P. E., Loui, P., Tarr, B., Schachner, A., Glowacki, L., Mithen, S., & Fitch, W. T. (2021). Music as a coevolved system for social bonding. *Behavioral and Brain Sciences*, 44, e59. doi:10.1017/S0140525X20000333

Savage, P. E., Passmore, S., Chiba, G., Currie, T. E., Suzuki, H., & Atkinson, Q. D. (2022). Sequence alignment of folk song melodies reveals cross-cultural regularities of musical evolution. *Current Biology*, 32(6), 1395–1402.e8. <https://doi://doi.org/10.1016/j.cub.2022.01.039>

Sawyer, R. K. (2006). *Explaining creativity: The science of human innovation*. Oxford University Press

Schafer, V., & Pailler, F. (2024). 'All your image are belong to us': heritagization, archiving and historicization of memes. *Visual Communication*, 23(3), 527-543. <https://doi.org/10.1177/14703572231221030>

Kenny, N., Chan, K., Nong, W. et al. Ancestral whole-genome duplication in the marine chelicerate horseshoe crabs. *Heredity* 116, 190–199 (2016). <https://doi.org/10.1038/hdy.2015.89>

Savage, P.E. Cultural evolution of music. *Palgrave Commun* 5, 16 (2019). <https://doi.org/10.1057/s41599-019-0221-1>

Fitch W. T. (2015). Four principles of bio-musicology. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 370(1664), 20140091. <https://doi.org/10.1098/rstb.2014.0091>

Pinker, S. (1997). *How the mind works*. W. W. Norton & Company.

Shintel, H. (2021). Music as auditory cheesecake. In T. K. Shackelford & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of evolutionary psychological science* (pp. 1–5). Springer. https://doi.org/10.1007/978-3-319-19650-3_2851

Tifferet, S., Gaziel, O., & Baram, Y. (2012). Guitar increases male Facebook attractiveness: Preliminary support for the sexual selection theory of music. *Letters on Evolutionary Behavioral Science*, 4(3), 4–6. <https://doi.org/10.5178/lebs.2012.18>

Passmore, S., & Savage, P. E. (2023). The Exceptions and the Rules in Global Musical Diversity. *Journal of Cognition*, 6(1): 47, pp. 1–17. DOI: <https://doi.org/10.5334/joc.312>

Savage, P. E., & Brown, S. (2013). Toward a new comparative musicology. *Analytical Approaches to World Music*, 2(2)

Schedl, M., Knees, P., McFee, B., i Bogdanov, D. (2022). Music recommendation systems: Techniques, use cases, and challenges. W F. Ricci, L. Rokach, i B. Shapira (Red.), *Recommender systems handbook* (Wyd. 3, s. 927-971). Springer.

Schenker, H. (1979). Free composition (Der freie Satz) (E. Oster, Trans.). Longman. (Original work published 1935)

Schiavio, A., & van der Schyff, D. (2018). 4E Music Pedagogy and the Principles of Self-Organization. *Behavioral sciences* (Basel, Switzerland), 8(8), 72. <https://doi.org/10.3390/bs8080072>

Schiavio, A., & van der Schyff, D. (2024). 4E music cognition in theory and practice. In *Psychological Perspectives on Musical Experiences and Skills: Research in the Western Balkans and Western Europe* (pp. 189-210). Open Book Publishers. <https://doi.org/10.11647/obp.0389.09>

Schleihauf, H., Hoehl, S., Tsvetkova, N., König, A., Mombaur, K. and Pauen, S. (2021), Preschoolers' Motivation to Over-Imitate Humans and Robots. *Child Dev*, 92: 222-238. <https://doi.org/10.1111/cdev.13403>

Schloss, J. G. (2014). *Making Beats: The Art of Sample-Based Hip-Hop*. Wesleyan University Press

Schmidt, C.. (2007, June 7). Charlie Schmidt's Keyboard Cat! – THE ORIGINAL! [Video]. YouTube. <https://www.youtube.com/watch?v=J---aiyznGQ>

Schmidt, C. 2025. Keyboard cat: the meme lives on <https://www.keyboardcat.fun/>

Schmidt, C. [Keyboard Cat]. (2016, November). Keyboard Cat Bongos [Video post]. Facebook. <https://www.facebook.com/thekeyboardcat/videos/1377236358997002>

Schonfeld, E. (2011, April 2). Founder Stories: Moot on the origin of 4chan and the evolution of memes. TechCrunch. <https://techcrunch.com/2011/04/02/founder-stories-poole-moot-4chan-memes/>

Schuling, F. (2019). Notation Cultures: Towards an Ethnomusicology of Notation. *Journal of the Royal Musical Association*, 144(2), 429–458. doi:10.1080/02690403.2019.1651508

Schultz, W. Dopamine reward prediction-error signalling: a two-component response. *Nat Rev Neurosci* 17, 183–195 (2016). <https://doi.org/10.1038/nrn.2015.26>

Scott, D. (2001). Music and social class. In J. Samson (Ed.), *The Cambridge History of Nineteenth-Century Music* (pp. 544–567). chapter, Cambridge: Cambridge University Press.

Seabrook, J. (2015). *The song machine: Inside the hit factory*. W. W. Norton & Company

Seth, A. K. (2005). Neural Darwinism and consciousness. *Consciousness and Cognition*, 14(1), 140–168. <https://doi.org/10.1016/j.concog.2004.08.008>

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>

Shannon, C.E. (1948) A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27, 379-423.

<https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>

Shared Processing of Language and Music

Ryan P. Atherton, Quin M. Chrobak, Frances H. Rauscher, Aaron T. Karst, Matt D. Hanson, Steven W. Steinert, and Kyra L. Bowe

Experimental Psychology 2018 65:1, 40-48

Sharma, A., Czégel, D., Lachmann, M., Kempes, C. P., Walker, S. I., & Cronin, L. (2023). Assembly theory explains and quantifies selection and evolution. *Nature*, 622(7982), 321–327. <https://doi.org/10.1038/s41586-023-06600-9>

Sharma, A., Czégel, D., Lachmann, M. et al. Assembly theory explains and quantifies selection and evolution. *Nature* 622, 321–328 (2023). <https://doi.org/10.1038/s41586-023-06600-9>

Sheikh, M. (2024, February 15). How the TikTok algorithm works in 2025. Sprout Social Blog. <https://sproutsocial.com/insights/tiktok-algorithm>

Shifman, L. (2011). An anatomy of a YouTube meme. *New Media & Society*, 14(2), 187-203. <https://doi.org/10.1177/1461444811412160> (Original work published 2012)

Shulman, M., (2024, January) How to Build AI That Creates Music, with Suno CEO, Mikey Shulman. (Podcast interview, transcript).

Siedenburg, K., Mativetsky, S., & McAdams, S. (2016). Auditory and verbal memory in North Indian tabla drumming. *Psychomusicology: Music, Mind, and Brain*, 26(4), 327–336. <https://doi.org/10.1037/pmu0000163>

Siersdorfer, S., Chelaru, S., Nejdil, W., & San Pedro, J. (2010). How useful are your comments? Analyzing and predicting YouTube comments and comment ratings. In *Proceedings of the 19th International Conference on World Wide Web (WWW '10)* (pp. 891–900). ACM. <https://doi.org/10.1145/1772690.1772781>

Silver, D., Huang, A., Maddison, C. et al. Mastering the game of Go with deep neural networks and tree search. *Nature* 529, 484–489 (2016). <https://doi.org/10.1038/nature16961>

Silver, D., Schrittwieser, J., Simonyan, K. et al. Mastering the game of Go without human knowledge. *Nature* 550, 354–359 (2017). <https://doi.org/10.1038/nature24270>

Simon, H. A. ,(1971). Designing organizations for an information-rich world. In M. Greenberger (Ed.), *Computers, Communications, and the Public Interest*. Johns Hopkins Press.

Simon, I., Huang, C.-Z. A., Engel, J., Hawthorne, C., & Dinculescu, M. (2019, 16 września). Generating piano music with Transformer. *The Magenta Blog*. <https://magenta.tensorflow.org/piano-transformer>

Stöter et al., (2019). Open-Unmix - A Reference Implementation for Music Source Separation. *Journal of Open Source Software*, 4(41), 1667, <https://doi.org/10.21105/joss.01667>

Singh, A., Bagdi, A., Agarwal, V., Goyal, T., & Ghosh, R. (2023). Envisioning a Hybrid Recommendation System for Filtering the Content from Short Videos. 2023 7th International Conference on Trends in Electronics and Informatics (ICOEI), 1661-1666. <https://doi.org/10.1109/ICOEI56765.2023.10125616>

Sister Mary Blaze. (n.d.). BBL DRIZZY [TikTok video]. TikTok. Retrieved August 15, 2025, from <https://www.tiktok.com/@sistermaryblaze/video/7365996835531689262>

Metro Boomin. (n.d.). BBL drizzy bpm 150. SoundCloud. Retrieved August 15, 2025, from <https://soundcloud.com/metroboomin/bbl-drizzy-bpm-150-mp3>

Sköld, M. (2022). Notation as visual representation of sound-based music. *Journal of New Music Research*, 51(2–3), 186–202. <https://doi.org/10.1080/09298215.2023.2174144>

Sloane, V. (2018, June 14). Joke’s on you — Darude’s “Sandstorm” is still cool. *Miami New Times*.

<https://www.miaminewtimes.com/music/interview-with-darude-hes-thankful-for-the-sand-storm-memes-10636780>

Snapes, L. (2023, April 18). AI song featuring fake Drake and Weeknd vocals pulled from streaming services. *The Guardian*. Retrieved from <https://www.theguardian.com/music/2023/apr/18/ai-song-featuring-fake-drake-and-weeknd-vocals-pulled-from-streaming-services>

Soha, M., & McDowell, Z. J. (2016). Monetizing a Meme: YouTube, Content ID, and the Harlem Shake. *Social Media + Society*, 2(1), 1–12. <https://doi.org/10.1177/2056305115623801>

Sony, 2025: <https://ai.sony/blog/Sony-AI-Ethics-Flagship-Reflecting-on-Our-Progress-and-Purpose/>

Sony Music Entertainment et al. v. Udio Inc., No. 1:24-cv-04748 (S.D.N.Y. June 24, 2024). <https://www.riaa.com/wp-content/uploads/2024/06/Udio-Complaint-6.24.241.pdf>

SoundAlerts. (2023, March 17). How to create a soundboard for your livestream — Guide. <https://soundalerts.com/blog/how-to-create-a-soundboard-for-your-livestream-guide>

Southcott, J. (2003). The singing by-ways: Origins of class music education in South Australia. *Journal of Historical Research in Music Education*, 25(1), 42–61. <https://www.jstor.org/stable/40215285>

Southern, T. (2018). I AM AI. Taryn Southern. Retrieved August 15, 2025, from <https://tarynsouthern.com/album/>

Southern, T. (n.d.). How AI is unlocking human creativity. Info-Tech. Retrieved August 15, 2025, from

<https://www.infotech.com/digital-disruption/taryn-southern-how-ai-is-unlocking-human-creativity>

SPENCER, E. K. (2025). When Donald Trump Dropped the Bass: The Weaponization of Dubstep in Internet Trolling Strategies, 2011–2016. *Twentieth-Century Music*, 22(1), 107–129. doi:10.1017/S1478572224000094

Spencer, H. (1857). Progress: Its law and cause. *The Westminster Review*, 67 (April), 445–485.

Spencer, H. (1857). The origin and function of music. *Fraser's Magazine*, 56, 396–408.

Spontaneous Emergence of Self-Replicating Molecules Containing Nucleobases and Amino Acids. Bin Liu, Charalampos G. Pappas, Jim Ottel , Ga l Schaeffer, Christoph Jurissek, Priscilla F. Pieters, Meniz Altay, Ivana Mari , Marc C. A. Stuart, and Sijbren Otto

Journal of the American Chemical Society 2020 142 (9), 4184-4192

DOI: 10.1021/jacs.9b10796

Spotify. (2023, March 8). Behind the scenes of Spotify's new AI DJ [Press release]. Spotify Newsroom. <https://newsroom.spotify.com/2023-03-08/spotify-new-personalized-ai-dj-how-it-works/>

Spotify. (2025, June 30). Discover Weekly turns 10: Celebrating 100 billion tracks streamed and a decade of personalized discovery. Spotify Newsroom. Pobrano 21 września 2025 z <https://newsroom.spotify.com/2025-06-30/discover-weekly-turns-10-celebrating-100-billion-tracks-streamed-and-a-decade-of-personalized-discovery/>

Spotify. (b.d.). Spotify Mixes. Spotify. Pobrano 21 września 2025 z <https://support.spotify.com/us/article/spotify-mixes/>

Spotify. (b.d.). Understanding recommendations on Spotify. Safety and Privacy center. Pobrane 6 czerwca 2025 z

<https://www.spotify.com/safetyandprivacy/understanding-recommendations/>

Spotify for Artists. (b.d.). Discovery Mode. Pobrane 6 czerwca 2025 z <https://artists.spotify.com/discovery-mode>

Spotify for Artists. (b.d.). Discovery Mode. Spotify. Pobrano 21 września 2025 z <https://artists.spotify.com/pl/discovery-mode>

Spotify Newsroom. (2025, 7 maja). Experience a new dimension of music discovery with more controls and enhanced tools. Pobrane 6 czerwca 2025 z <https://newsroom.spotify.com/2025-05-07/experience-a-new-dimension-of-music-discovery-with-more-controls-and-enhanced-tools/>

Springer. (2025). Aims and scope: Memetic Computing. Springer. Retrieved September 25, 2025, from <https://link.springer.com/journal/12293/aims-and-scope>

Stafford, P. E. (2018). The Grunge Effect: Music, Fashion, and the Media During the Rise of Grunge Culture In the Early 1990s. *M/C Journal*, 21(5). <https://doi.org/10.5204/mcj.1471>

Staiti, A., & Oliva, L. (2024, Fall Edition). Heinrich Rickert. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2024 ed.). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/fall2024/entries/heinrich-rickert/>

Stanford HAI. (2025). Simulating human behavior with AI agents. Retrieved from <https://hai.stanford.edu/policy/simulating-human-behavior-with-ai-agents>

Stefon, M. (2025). Wuwei. In Encyclopaedia Britannica. Odzyskano z <https://www.britannica.com/topic/wuwei-Chinese-philosophy>

Sterelny, Kim, *Thought in a hostile world*, 2003

Stern, J. (2013, February 13). 'Harlem Shake' Shakes It Across YouTube, With Over 44 Million Views. ABC News. <https://abcnews.go.com/blogs/technology/2013/02/harlem-shake-shakes-it-across-youtu-be-with-over-44-million-views>

Stockmann, D. (1982). Muzyka jako system komunikacji. Aspekty teorii informacji i znaku w badaniach muzyki przekazywanej tradycją ustną. *Res Facta*, (9), 77–87.

Stout Dietrich 2011 Stone toolmaking and the evolution of human culture and cognition *Phil. Trans. R. Soc.* B3661050–1059
<http://doi.org/10.1098/rstb.2010.0369>

Strunk, S. (2003). Harmony (i). In *Grove Music Online*. Oxford University Press

Stryker, S. (1968). Identity salience and role performance: The relevance of symbolic interaction theory for family research. *Journal of Marriage and the Family*, 30(4), 558–564. <https://doi.org/10.2307/349494>

Strzelecki, M. (2024). *Niech się stanie muzyka! O muzycznej generatywności*. Kraków: Akademia Muzyczna w Krakowie / Wydawnictwo Akademii Muzycznej w Krakowie.

Stöter, F.-R., Uhlich, S., Liutkus, A., & Mitsufuji, Y. (2019). Open-Unmix: A reference implementation for music source separation. *Journal of Open Source Software*, 4(41), 1667. DOI: 10.21105/joss.01667

Su, X., & Khoshgoftaar, T. M. (2009). A survey of collaborative filtering techniques. *Advances in Artificial Intelligence*, 2009, 1–19. <https://doi.org/10.1155/2009/421425>

Suddath, C. (2009, May 11). Play him off, Keyboard Cat. *Time*.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). The MIT Press.

Tan, S. (2024, December 2). Are we all musicians now? Authenticity, musicianship, and AI music generator Suno. OSF. <https://osf.io/4nt8z/>

Tang, D., Liu, Z., Yang, J., & Zhao, J. (2019). Memetic frog leaping algorithm for global optimization. *Soft Computing*, 23, 11077–11105. <https://doi.org/10.1007/s00500-018-3662-3>

Taruskin, R. (2010). *Music in the nineteenth century*. Oxford University Press.

Taruskin, R. (2010). *Music in the seventeenth and eighteenth centuries*. Oxford University Press.

Tencer, D. (2024, January 10). 158 million tracks had 1,000 plays or fewer on music streaming services last year. 45 million had no plays at all. *Music Business Worldwide*. <https://www.musicbusinessworldwide.com/158-million-tracks-1000-plays-on-streaming-services/>

Tenney, J. (1988). *A history of “consonance” and “dissonance”*. Excelsior Music Publishing Company, New York

Terapia Nad Wartą. (2024, 15 maja). Czym jest uzależnienie od dopaminy? *Terapia Nad Wartą*. <https://www.terapianadwarta.pl/czym-jest-uzaleznienie-od-dopaminy>

Tereshchenko S. Y. (2023). Neurobiological risk factors for problematic social media use as a specific form of Internet addiction: A narrative review. *World journal of psychiatry*, 13(5), 160–173. <https://doi.org/10.5498/wjp.v13.i5.160>

Tharin Pillay and Harry Booth, 5 Predictions for AI in 2025
<https://time.com/7204665/ai-predictions-2025/>

The Editors of Encyclopaedia Britannica. (2023, December 1). Musical societies and institutions. In *Encyclopædia Britannica*.
<https://www.britannica.com/art/musical-societies-and-institutions>

The impact of algorithmically driven recommendation systems on music consumption and production - a literature review. (2023). GOV.UK. Pobrane 6 czerwca 2025 z <https://www.gov.uk/government/publications/research-into-the-impact-of-streaming-services-algorithms-on-music-consumption/the-impact-of-algorithmically-driven-recommendation-systems-on-music-consumption-and-production-a-literature-review>

The Royal Institution. (2017, April 6). If brains are computers, who designs the software? | With Daniel Dennett [Video]. YouTube.
https://youtu.be/TTFoJQsd48c?si=oeZltyR3CdvH_evT

TikTok Support. (b.d.). How TikTok recommends content. Pobrane 6 czerwca 2025 z <https://support.tiktok.com/en/using-tiktok/exploring-videos/how-tiktok-recommends-content>

Tillmann, B., Graves, J., Talamini, F., Lévêque, Y., Fornoni, L., Hoarau, C., Pralus, A., Ginzburg, J., Albouy, P., & Caclin, A. (2023). Auditory cortex and beyond: Deficits in congenital amusia. *Hearing Research*. <https://doi.org/10.1016/j.heares.2023.108855>

Timothy Papandreou, 2025: Agentic And Physical AI — A Multitrillion Dollar Economy Emerges

<https://www.forbes.com/sites/timothypapandreou/2025/01/15/2025-agentic--physical-ai-multi-trillion-dollar-economy-emerges/>

Todd, P. M. (1989). A connectionist approach to algorithmic composition. *Computer Music Journal*, 13(4), 27–43. <https://doi.org/10.2307/3679551>

Toolify.ai. (2024, 19 stycznia). The impact of algorithms on music discovery and the music industry. Pobrane 6 czerwca 2025 z <https://www.toolify.ai/ai-news/the-impact-of-algorithms-on-music-discovery-and-the-music-industry-650761>

Totschnig, W. Fully Autonomous AI. *Sci Eng Ethics* 26, 2473–2485 (2020). <https://doi.org/10.1007/s11948-020-00243-z>

Towse, R. (2020). Dealing with digital: the economic organisation of streamed music. *Media, Culture & Society*, 42(7-8), 1461-1478. <https://doi.org/10.1177/0163443720919376> (Original work published 2020)

Trainor, L. J. (2009). The primal role of the vestibular system in determining musical rhythm. *Annals of the New York Academy of Sciences*, 1169(1), 63–71. <https://doi.org/10.1016/j.cortex.2007.10.014>

Trzciński, S. (2023). *Zarażeni dźwiękiem. Rynek muzyczny w czasach sztucznej inteligencji*. Warszawa: Wydawnictwo Naukowe PWN

Tschmuk, P. *Economics of music*, 2017

Tyrała, R. (2020). Znaczenie muzyki kościelnej w historii Kościoła. *Perspektywy Kultury*, 24(1), 71-104. <https://doi.org/10.35765/pk.2019.2401.07>

Tzafestas, E. (2021). First selectively imitate responses, then associate with function. *Frontiers in Psychology*, 12, 560653. <https://doi.org/10.3389/fpsyg.2021.560653>

UMG Recordings, Inc., et al. v. Suno, Inc., No. 1:24-cv-11533 (D. Mass. June 24, 2024). <https://www.riaa.com/wp-content/uploads/2024/06/Suno-complaint-file-stamped20.pdf>

Urban Dictionary. (2025). Darude [Definition on Urban Dictionary, page 2]. Retrieved from <https://www.urbandictionary.com/define.php?term=Darude&page=2>

Username701. (2024, April 20). Drake – Taylor Made Freestyle ft. A.I. Tupac & A.I. Snoop Dogg [Video]. YouTube. https://www.youtube.com/watch?v=X1f6Ny_aXx4

Van Buskirk, E. (2008, April 2). YouTube rickrolled around 6 million people yesterday

Van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press

van Hees, J., Grootswagers, T., Quek, G. L., & Varlet, M. (2025). Human perception of art in the age of artificial intelligence. *Frontiers in psychology*, 15, 1497469. <https://doi.org/10.3389/fpsyg.2024.1497469>

Varshneya, K. (2024, 8 stycznia). Spotify recommendation system and user engagement strategies. TechAhead. Pobrane 6 czerwca 2025 z <https://www.techaheadcorp.com/blog/spotify-recommendation-system/>

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems* 30 (NIPS 2017). Dostępne na: <https://arxiv.org/abs/1706.03762>

Veena, Pag, 2025 FL Studio MCP <https://github.com/veenastudio/flstudio-mcp>

Veissière, S. P. L., Constant, A., Ramstead, M. J. D., Friston, K. J., & Kirmayer, L. J. (2020). Thinking through other minds: A variational approach to cognition and culture. *Behavioral and Brain Sciences*, 43, e90. doi:10.1017/S0140525X19001213

Verma, Shivam, et al. "An Audio-Centric Multi-Task Learning Framework for Streaming Ads Targeting on Spotify." *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2*, ACM, 2025, pp. 4945–55. Crossref, <https://doi.org/10.1145/3711896.3737190>.

Vuoskoski, J. K., & Eerola, T. (2011). The role of mood and personality in the perception of emotions represented by music. *Cortex*, 47(9), 1099–1106. <https://doi.org/10.1016/j.cortex.2011.04.011>

Vuvan, D. T., Podolak, O. M., & Schmuckler, M. A. (2014). Memory for musical tones: the impact of tonality and the creation of false memories. *Frontiers in Psychology*, 5, Article 582. <https://doi.org/10.3389/fpsyg.2014.00582>

Walker, J. M., van der Heijden, E. S. M., Maulana, A., Rueda-M, N., Näsval, K., Salazar, P. A., Meyer, M., & Meier, J. I. (2024). Common misconceptions of speciation. *Evolutionary Journal of the Linnean Society*, 3(1), kzae029. <https://doi.org/10.1093/evolinnean/kzae029>

Walker, S. I. (2024). *Life as no one knows it: The physics of life's emergence*. Penguin Publishing Group

Wang, B. (2024). Diverse capability and scaling of diffusion and auto-regressive models when learning abstract rules. *arXiv preprint arXiv:2411.07873*. <https://doi.org/10.48550/arXiv.2411.07873>

Wang, D., & Chen, J. (2018). Supervised speech separation based on deep learning: An overview. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26(10), 1702–1726. DOI: 10.1109/TASLP.2018.2842159

Wang, E. (b.d.). Homogenization of pop music: How social media's algorithms...
Pobrane 6 czerwca 2025 z <https://www.scholarlyreview.org/article/124652.pdf>

Wang, L., & Wood, B. C. (2011). An epidemiological approach to model the viral propagation of memes. *Applied Mathematical Modelling*, 35(11), 5442–5447. <https://doi.org/10.1016/j.apm.2011.04.035>

Watkins, S. C. (2005). *Hip hop matters: Politics, pop culture, and the struggle for the soul of a movement*. Beacon Press.

Waxman, O. B. (2014, May 29). WATCH: YouTube's Newest Music Mashup Video (9th birthday). TIME.
<https://time.com/134448/watch-youtube-newest-music-mashup-video-for-9th-birthday>

Weber, M. (1958). *The rational and social foundations of music* (Don Martindale & Gertrud Neuwirth, Trans.). Southern Illinois University Press

Weber, W. (1975). *Music and the middle class: The social structure of concert life in London, Paris and Vienna*. Croom Helm.

Weber, W. (1994). Mass Culture and the Reshaping of European Musical Taste, 1770-1870. *International Review of the Aesthetics and Sociology of Music*, 25(1/2), 175–190. <https://doi.org/10.2307/836942>

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Chi, E.H., Xia, F., Le, Q., & Zhou, D. (2022). Chain of Thought Prompting Elicits Reasoning in Large Language Models. *ArXiv*, abs/2201.11903.

Weick, K. E. (1995). *Sensemaking in organizations*. SAGE Publications.

Weiss, J. (2013, June 16). In search of the air horn. *Red Bull Music Academy Daily*.
<https://daily.redbullmusicacademy.com/2013/06/in-search-of-the-air-horn>

Whiten, A., Hinde, R. A., Laland, K. N., & Stringer, C. B. (2011). Culture evolves. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567), 938–948. <https://doi.org/10.1098/rstb.2010.0372>

Whiteson, S., & Stone, P. (2006). Evolutionary function approximation for reinforcement learning. *Journal of Machine Learning Research*, 7, 837–881
<http://www.jmlr.org/papers/v7/whiteson06a.html>

Wieczorek, R. J. (2013). *Patronat muzyczny w renesansowych Włoszech (1470–1527)*. Mediolan, Ferrara, Mantua, Florencja, Rzym. Poznań: Wydawnictwo Naukowe UAM

Wiggins, G. A., Müllensiefen, D., & Pearce, M. T. (2010). On the non-existence of music: Why music theory is a figment of the imagination. *Musicae Scientiae*, 14(Suppl. 1), 231–255. <https://doi.org/10.1177/10298649100140S110>

Wikipedia contributors. (2025, August 13). Drake–Kendrick Lamar feud. In *Wikipedia, The Free Encyclopedia*. Retrieved 09:08, August 15, 2025, from https://en.wikipedia.org/w/index.php?title=Drake%E2%80%93Kendrick_Lamar_feud&oldid=1305689790

Wikstrom, Patrik. (2014). The music industry in an age of digital distribution. *Change: 19 Key Essays on How the Internet is Changing Our Lives*. 1-24.

Wileczek, A. (2021a, December 2). *Młodzieżowe Słowo Roku 2021 – komentarz Anny Wileczek*. PWN. Retrieved from

<https://sjp.pwn.pl/ciekawostki/haslo/mlodziejowe-slowo-roku-2021-komentarz-anny-wileczek;9283280.html>

Wileczek, A. E. (2021b). W czym pomaga twoja stara – o funkcjonalnej stronie języka młodych: komentarz [w: PWN, Młodzieżowe Słowo Roku 2022: Zakamarki Młodej Polszczyzny]. Pobrane z <https://sp51gdynia.pl/files/431/pwnmlodziejowesloworokuzakamarki-mlodejpolszczyzny2022.pdf>

Williams, A. T. (2021). Spotify: A strategic analysis. ScholarWorks at University of Montana. <https://scholarworks.umt.edu/cgi/viewcontent.cgi?article=1346&context=utpp>

Williams, MA. (2010). Autonomy: Life and Being. In: Bi, Y., Williams, MA. (eds) Knowledge Science, Engineering and Management. KSEM 2010. Lecture Notes in Computer Science(), vol 6291. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-15280-1_15

Williamson, V. J., Liikkanen, L. A., Jakubowski, K., & Stewart, L. (2014). Sticky tunes: How do people react to involuntary musical imagery? PLoS ONE, 9(1), e86170. <https://doi.org/10.1371/journal.pone.0086170>

Wilson, M. (2025, February 14). Suno explained: How to use the viral AI song generator for free. TechRadar. <https://www.techradar.com/computing/artificial-intelligence/what-is-suno-ai>

Wilson, R. (2020). Artificial intelligence safety [Snapshot report]. Routledge. https://www.routledge.com/rsc/downloads/RWCOMPP1901_AI_Safety_SS_r2-final.pdf

Wong, J. (2024, 21 stycznia). Algorithmic symphonies: How Spotify strikes the right chord. USC Viterbi School of Engineering. Pobrane 6 czerwca 2025 z <https://illumin.usc.edu/algorithmic-symphonies-how-spotify-strikes-the-right-chord/>

Woolley, B. (2020). *The bride of science: Romance, reason, and Byron's daughter*. Pan Macmillan.

Wortham, J. (2008, April 1). YouTube 'Rickrolls' everyone. *Wired*.

Y., Silviya. (2025, 7 stycznia). Echo chambers in the music industry: How trends are shaped by online subcultures. *Indigomusic.com*. Pobrane 6 czerwca 2025 z <https://indigomusic.com/feature/echo-chambers-in-the-music-industry-how-trends-are-shaped-by-online-subcultures>

Yampolskiy, R. V. (2024). *AI: Unexplainable, Unpredictable, Uncontrollable*. Chapman & Hall/CRC

Yang, S., Brossard, D., Scheufele, D. A., & Xenos, M. A. (2022). The science of YouTube: What factors influence user engagement with online science videos? *PLOS ONE*, 17(5), e0267697. <https://doi.org/10.1371/journal.pone.0267697>

YouTube Music Help. (2025). Listen to personal or custom radio. <https://support.google.com/youtubemusic/answer/15165061?hl=en>

Zakrzewski, B. (1964). Improwizacje grudniowe Juliusza Słowackiego. *Pamiętnik Literacki : Czasopismo Kwartalne Poświęcone Historii i Krytyce Literatury Polskiej*, 55(1), 215–219.

Zannettou, S., Finkelstein, J., Bradlyn, B., & Blackburn, J. (2018). On the origins of memes by means of fringe web communities. *Proceedings of the 2018 Internet Measurement Conference (IMC '18)*, 188–202. <https://doi.org/10.1145/3278532.3278550>

Zeravcic Z., & M.P. Brenner, Self-replicating colloidal clusters, *Proc. Natl. Acad. Sci. U.S.A.* 111 (5) 1748-1753, <https://doi.org/10.1073/pnas.1313601111> (2014).

Zhai, J., Cheng, Y., Li, H., & Tang, X. (2021). Research on meme transmission based on individual heterogeneity. *Mathematical Biosciences and Engineering*, 18(5), 5294–5315. <https://doi.org/10.3934/mbe.2021263>

Zhang, M., & Liu, Y. (2021). A commentary of TikTok recommendation algorithms in MIT Technology Review 2021. *Fundamental Research*, 1(6), 846–847. <https://doi.org/10.1016/j.fmre.2021.11.015>

Zhang, S., Yao, L., Sun, A., & Tay, Y. (2019). Deep learning based recommender system: A survey and new perspectives. *ACM Computing Surveys*, 52(1), 1–38. <https://doi.org/10.1145/3285029>

Zhou, R. (2024). Understanding the Impact of TikTok's Recommendation Algorithm on User Engagement. *International Journal of Computer Science and Information Technology*, 3(2), 201-208.

Zittrain, Jonathan. (2009). *The Future of the Internet and How to Stop It*.

Zou, J., Jiang, W., Xia, Y., Liu, Y., & Hou, Z. (2024). G-EvoNAS: Evolutionary neural architecture search based on network growth (arXiv preprint). arXiv. <https://doi.org/10.48550/arXiv.2403.02667>

Święczkowska, H. (2019). *Racjonalistyczna perspektywa językoznawcza: Szkice z nowożytnej filozofii języka*. Wydawnictwo Uniwersytetu w Białymstoku.