Tendencje rynkowe dotyczące produkcji filmowej w oparciu o nowoczesne technologie AI

KALINA BIENIAS

Uniwersytet Śląski w Katowicach
Szkoła Filmowa im. Krzysztofa Kieślowskiego

Streszczenie

Artykuł poświęcony jest analizie zastosowania tzw. sztucznej inteligencji (AI) w branży filmowej oraz jej wpływu na procesy twórcze i produkcyjne. Skupiono się na dynamicznym rozwoju technologii generatywnych AI w obszarach tworzenia obrazu, filmu i tekstu, z uwzględnieniem ich rosnącego znaczenia w ostatnich latach. Omówiono potencjalne zagrożenia, jakie sztuczna inteligencja stwarza dla tradycyjnej twórczości artystycznej, oraz wyrażane obawy o możliwość zastąpienia realnych twórców przez maszyny. Tekst artykułu składa się z trzech głównych części. Część pierwsza opisuje ewolucję technologii generatywnych, podkreślając ich coraz większą rolę w różnych dziedzinach, w tym w produkcji filmowej. W części drugiej omówiono kluczowe problemy związane z prawami autorskimi i etyką, szczególnie w kontekście wykorzystywania istniejących dzieł do szkolenia modeli AI bez zgody ich autorów. Część trzecia stanowi analizę konkretnych programów i narzędzi opartych na AI, które mogą wspierać twórców filmowych na różnych etapach produkcji. W podsumowaniu artykułu postawiona jest teza, że rozwój technologii AI, choć budzi liczne kontrowersje, jest nieunikniony i stanowi kolejny przełom w branży kreatywnej, porównywalny do wprowadzenia dźwięku i koloru w filmie. Zwrócono uwagę na możliwe konsekwencje tej rewolucji, w tym na zmiany w strukturze rynku pracy w sektorze artystycznym, ale także na nadzieje związane z tym, że narzędzia AI mogą stać się cennym wsparciem, a nie zagrożeniem dla twórczości ludzkiej.

Słowa-klucze: sztuczna inteligencja, technologie generatywne, AI, produkcja filmowa

Abstract

Market Trends in Film Production Based on Modern AI Technologies

The article is dedicated to analyzing the application of artificial intelligence (AI) in the film industry and its impact on creative and production processes. It focuses on the dynamic development of generative AI technologies in the fields of image, film, and text creation, highlighting their growing significance in recent years. Potential threats posed by AI to traditional artistic creativity are discussed, along with concerns about the possibility of machines replacing human creators. The article consists of three main parts. The first part describes the evolution of generative technologies, emphasizing their increasing role in various fields, including film production. The second part addresses key issues related to copyright and ethics, particularly in

the context of using existing works to train AI models without the consent of their authors. The third part analyzes specific AI-based programs and tools that can support filmmakers at different stages of production. In the conclusion, the article asserts that the development of AI technologies, despite raising numerous controversies, is inevitable and represents another breakthrough in the creative industry, comparable to the introduction of sound and color in film. The article highlights the potential consequences of this revolution, including changes in the labor market structure within the artistic sector, as well as the hope that AI tools can serve as valuable support rather than a threat to human creativity.

Keywords: artificial intelligence, generative technologies, AI, film production

I. WPROWADZENIE

W ostatnich kilku latach możemy zaobserwować gigantyczny rozwój technologii opartej na sztucznej inteligencji. Chociaż efekty tego postępu widzimy dopiero teraz, to rozwój systemów generatywnych AI[1] (w pracy będę skupiać się na systemach AI dotyczących obrazu, filmu, tekstu) rozpoczął się już dużo wcześniej. Dlaczego zatem przedtem nie zwracaliśmy uwagi na powstające oprogramowania? Można rzec, że dopiero niedawno sztuczna inteligencja weszła w swój wiek nastoletni — wiek dorastania. Tak jak człowiek w tym okresie: zwraca na siebie uwagę, a co najważniejsze, rozwija się błyskawicznie. Wielu zastanawia się dzisiaj, jaki świat czeka nas, gdy sztuczna inteligencja osiągnie swoją umowną pełnoletność.

Każda technologia ma swoją dobrą i złą stronę. Nowe rewolucje i wynalazki zawsze budziły niepokój, jednakże żadne z wcześniejszych oprogramowań nie miało wystarczającego potencjału, by zagrozić kreatywności człowieka i zastąpić go w pracy twórczej. Programy, które zmieniły rynek, takie jak Adobe Photoshop dawały nowe możliwości techniczne do rozwoju kreatywności człowieka. Czynności związane z fotomontażem, retuszem, obróbką zdjęć zostały przeniesione z ciemni do programu komputerowego. Oszczędzało to czas i pieniądze, oraz ułatwiało cały proces związany z szeroko pojętą obróbką materiału zdjęciowego. Maszyna miała zadanie ułatwić życie człowiekowi, a nie zastąpić jego miejsce.

Obecnie wielu twórców obawia się zastąpienia sztuki człowieka przez sztukę stworzoną przez maszyny. Gigantyczny skok rozwojowy, który nastąpił w ostatnich latach budzi dość duży niepokój w środowiskach artystycznych. Znajdujemy się w momencie, kiedy padają pytania: czy AI odbierze pracę artystom? Czy powinno się

[1] W pracy używane jest określenie AI (ang. Artificial Intelligence) czyli Sztuczna inteligencja (SI), ponieważ to anglosaskie określenie upowszechniło się w nazwach własnych funkcjonujących obecnie modeli językowych i podobnych generatorów. Co warto jednocześnie podkreślić, owe modele obdarzane mianem "inteligentnych" bazują jedynie na ograniczonych zbiorach danych, poza które obecnie nie są w stanie wykroczyć — nie są zatem zdolne do działań utożsamianych z inteligencją.

rozwijać takie technologie? Czy w przyszłości filmy będą w całości tworzone przez sztuczną inteligencję?

II. SKALA ROZWOJU GENERATYWNEJ SZTUCZNEJ INTELIGENCJI NA PRZESTRZENI OSTATNICH 14 LAT

Coś, co wydaje się opinii publicznej skokiem w rozwoju jest zazwyczaj efektem stopniowych przemian w ramach danej technologii. To powolny proces, na który wpływ mają pojedyncze dokonania, których znaczenie można odpowiednio rozpatrywać w skali makro. Decydujące dla wzrostu popularności programów AI okazało się bowiem osiągnięcie przez owe programy pewnego poziomu użyteczności i wielotorowej przydatności. Konkretni użytkownicy programów lub platform generatywnych oraz w bardziej ogólnym znaczeniu, osoby mniej lub bardziej entuzjastycznie do nich nastawione mogą mieć wrażenie, jakoby dane rozwiązania pojawiły się zupełnie niespodziewanie. Dopiero z perspektywy czasu można dostrzec, które odkrycia miały kluczowy wpływ na rozwój generatywnej technologii.

W 2013 roku firma Google przedstawiła dwa nowe algorytmy, które opierały się na reprezentowaniu słów przez wektory liczbowe[2]. Nowa technika była znana w branży pod nazwą „Word2vec". Zespół naukowców wytrenował algorytm tak, aby przewidywać następne lub brakujące słowo w treści tekstu. Model szkolono na treściach z Internetu, co było dużym przełomem. Wcześniej trenowano na danych oznaczonych i pogrupowanych przez ludzi, co oczywiście było czasochłonne i skutkowało niewystarczającą ilością danych.

W 2017 roku Google Brain stworzył model nowej architektury *głębokiego uczenia - deep lerning architecture* pod nazwą „Transformer Architecture" opublikowane

[2] T. Mikolov, K. Chen, G. Corrado, J. Dean, *Efficient Estimation of Word Representations in Vector Space*, http://arxiv.org/abs/1301.3781 [3 sierpień 2024 r.].

w pracy naukowej „Attention Is All You Need”[3]. Głębokie uczenie należy do dziedziny uczenia maszynowego, które rozwija algorytmy opierając się na funkcjonowaniu i strukturze ludzkiego mózgu. Architektura odnosi się tutaj do struktury skomplikowanej sieci neuronowej. Sieć neuronowa jest po prostu funkcją matematyczną, która zmienia wektor wejściowy na wektor wyjściowy[4]. Transformer przyczynił się do rozwoju modeli językowych, w szczególności modeli GPT oraz modeli generowania obrazu, które będą opierać się na rozumieniu tekstu. Nowe podejście do przetwarzania danych było podwaliną do szybkiego rozwoju kolejnych modeli językowych. Skrót GPT odnosi się do *Generative pre-trained transformer*, czyli Wstępnie Przeszkolony Transformator Generatywny. Szkolenie dotyczy trenowania modelu na ogromnej ilości danych, aby mógł on - w przypadku ChatGPT- rozumieć i generować tekst. Chat GPT należy do grupy tzw. large language models (LLM) czyli dużych modeli językowych. Grupa programów LLM zaliczana jest do sztucznej inteligencji.

W 2019 roku opublikowano kolejny model firmy OpenAI: GPT-2, ale dopiero rok 2020 w kontekście modelu firmy OpenAI jest uważany za bardziej przełomowy, ponieważ opublikowano model GPT-3. Był to nie tylko model, który rozumiał język i jego znaczenie. GPT-3 przez zastosowanie wspomnianej wcześniej architektury transformatorowej i wykorzystanie olbrzymiej ilość danych, potrafił wykonywać zadania wcześniej nieosiągalne. Trzecia generacja potrafiła tworzyć kreatywne teksty, spójniejsze odpowiedzi, wykazywała się lepszym rozumieniem tekstu i kontekstu. Warto dodać, że eksplozja nowych modeli generujących zdjęcia jak i obrazy nastąpiła po roku 2020 - po przełomie Chatu GPT-3. W roku 2021 mieliśmy aż 21 nowych modeli językowych oraz 17 nowych modeli związanych z obrazem!

[3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. Gomez, L. Kaiser, I. Polosukhin, *Attention Is All You Need*, http://arxiv.org/abs/1706.03762 [3 wrzesień 2024 r.].

[4] I. Białynicki-Birula, I. Białynicka-Birula, *Komputer się uczy: sieci neuronowe* [w:] *Modelowanie rzeczywistości*, Warszawa 2002, s. 100-109.

Programy generujące obrazy opierają się na osiągnięciach modeli językowych. Muszą potrafić zrozumieć tekst i kontekst, aby wygenerować obraz opisany przez użytkownika. Są to tak zwane: „*text-to-image models*", czyli od modele przetwarzające tekst na obraz.

W 2009 roku powstał projekt ImageNet. Była to wielka baza obrazów, skatalogowanych i opisanych przez ludzi, stworzona do badań nad rozwojem oprogramowań rozpoznających obrazy. Obecnie baza składa się z 20 000 kategorii oraz 14 milionów zdjęć. Odegrała kluczową rolę w rozwoju widzenia komputerowego *(ang. computer vision),* czyli dziedzinie sztucznej inteligencji, która rozwija umiejętność interpretowania i rozumienia zdjęć lub obrazów przez maszyny[5].

W 2014 roku stworzono nowy model głębokiego uczenia - *GAN Architecture*. Nowa struktura sieci neuronowej przyczyniła się do gigantycznego rozwoju szeroko pojętego generowania obrazu i dźwięku. Na początku GAN generowało bardzo prymitywne obrazy, np. proste kształty. Stanowiło to dużą obietnice i potencjał do tworzenia generatywnych wyników. Po zaledwie 4 latach rozwoju udało się osiągnąć tworzenie realistycznych obrazów, twarzy oraz obiektów. W 2018 roku naukowcy z firmy Nvidia zaprezentowali StyleGan, który był przełomem w generowaniu realistycznych obrazów[6].

[5] *ImageNet Data,* https://www.image-net.org/download.php [28 lipca 2024 r.].

[6] T. Karras, T. Aila, S. Laine, J. Lehtinen, *Progressive Growing Of Gans For Improved Quality, Stability, And Variation,* https://openreview.net/pdf?id=Hk99zCeAb [4 sierpnia 2024 r.].

Ilustracja 1: https://arxiv.org/abs/1812.04948

Obecna architektura *GAN* ma szerokie możliwości w kontekście przetwarzania obrazu: tworzenie nowych obrazów, generowanie obrazów na podstawie treści, restaurowanie obrazów, podnoszenie rozdzielczości zdjęć również poprzez usuwania szumu w obrazie, zmienianie stylu czy koloryzacji obrazów, przekształcanie zdjęć 2D w modele 3D, morfowanie obrazów (używane np. w efektach specjalnych i animacjach), dodawanie nowych elementów do obrazów[7]. W kontekście przetwarzania audio istnieją następujące możliwości: tworzenie nowych utworów (na podstawie danych wprowadzonych podczas treningu, czyli fragmentów utworów muzycznych), morfowanie dźwięku, restaurowanie dźwięku, uzupełnianie brakujących dźwięków, zmiany stylu muzycznego, generowanie efektów dźwiękowych, synteza mowy, rozpoznawanie mowy, generowanie odpowiedzi[8].

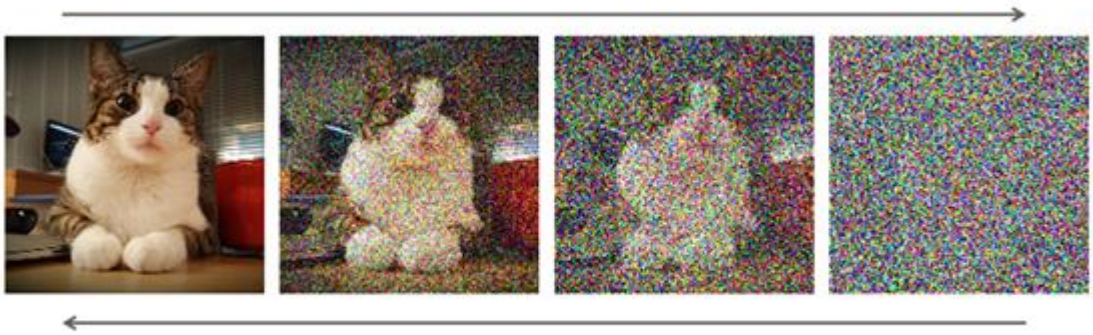
W 2021 roku OpenAI opracowało *CLIP architecture* (*Contrastive Language-Image Pretraining*), która jest w stanie równocześnie łączyć dane wizualne i tekstowe. CLIP

[7] X. Ouyang, Y. Chen, K. Zhu, G. Agam, *Image restoration refinement with Uformer GAN*, https://openaccess.thecvf.com/content/CVPR2024W/NTIRE/papers/Ouyang_Image_Restoration_Refinement_with_Uformer_GAN_CVPRW_2024_paper.pdf [4 sierpnia 2024 r.].

[8] S. Li, Y. Sung, *INCO-GAN: Variable-Length Music Generation Method Based on Inception Model-Based Conditional GAN*, https://doi.org/10.3390/math9040387 [4 sierpnia 2024r.].

może służyć jako wyszukiwarka – na podstawie podanego tekstu może wyszukać pasujące obrazy. Innym zastosowaniem jest tworzenie opisów do istniejących obrazów[9].

W 2021 roku następuje też znaczący rozwój *Modeli Dyfuzji* (*Difussion Models*). Słowo „dyfuzja" oznacza mieszanie się lub rozprzestrzenianie się cząsteczek. W przypadku tych modeli mamy do czynienia ze skomplikowanym procesem dodawania coraz większej ilości szumu do zdjęć, a potem z procesem redukowania szumu. Po usunięciu szumu model jest w stanie powrócić do początkowego obrazu. Prawdopodobnie dlatego w nazwie znajduje się słowo dyfuzja - dodajemy nowe „cząsteczki", dodajemy szum. W modelach dyfuzji istotny jest więc skomplikowany proces nauki usuwania szumu i oczyszczania obrazu, co pozwala na późniejsze generowanie nowych obrazów. *Difussion Models* mogą generować obrazy na podstawie tekstu, animować obrazy (wprowadzać ruch w zdjęciach), tworzyć filmy i modele 3D[10].



Ilustracja 1: https://developer.nvidia.com/blog/improving-diffusion-models-as-an-alternative-to-gans-part-1/

[9] A. Sikarwar, A. Pawar, A. Sethiya, A. Surana, S. Rajput, *CLIP Model Image Search: A Deep Learning Approach for Finding Relevant Images*, https://www.researchgate.net/publication/370300801_CLIP_Model_Image_Search_A_Deep_Learning_Approach_for_Finding_Relevant_Images [15 sierpnia 2024r.].

[10] L. Zhang, X. Chen, Y. Wang, Y. Lu, Y. Qiao, *Brush Your Text: Synthesize Any Scene Text on Images via Diffusion Model*, https://doi.org/10.1609/aaai.v38i7.28550 [30 lipca 2024r.].

Przykładem modeli dyfuzji są: DALL-E (płatny, opracowany przez OpenAI), Imagen (opracowany przez Google, nie udostępniony publicznie) oraz Stable Difussion (StabilityAI), który jako jedyny udostępniony jest na zasadzie open source - kod źródłowy jest dostępny do pobrania dla wszystkich. Na koniec warto dodać, że technologia sieci neuronowych nie różni się zbytnio od tej sprzed 20 lat.

To, co pozwoliło na prawdziwy rozwój generatywnej technologii sztucznej inteligencji to przede wszystkim dostęp do dużej ilości danych, które wykorzystuje się w treningu modeli. Wynalezienie Internetu oraz jego rozwój pozwoliły naukowcom tworzyć bazy danych, takie jak ImageNet, które przyczyniły się do efektywniejszego uczenia maszynowego. Internet sam w sobie był przecież źródłem dostępu do darmowych danych wykorzystywanych przy szkoleniu modeli. Kolejnym dość nieoczywistym czynnikiem, który niewątpliwie wpłynął na rozwój powyższych technologii było duże zainteresowanie w społeczeństwie grami komputerowymi. Gry zyskiwały z czasem na jakości, co spowodowało popyt na rozwój technologii kart graficznych w komputerach, które obecnie wykorzystywane są w rozwoju sieci neuronowych. Dzięki ogólnemu rozwojowi informatyki nastąpił postęp dotyczący zwiększenia mocy obliczeniowych, co również pozwoliło na coraz lepiej i efektywniej rozwijać nowe skomplikowane algorytmy.

III. AI A PRAWO AUTORSKIE

W środowisku artystycznym został wywołany chaos informacyjny. Twórcy z całego świata, ci znani i mniej znani zaczęli domniemywać, że to właśnie na ich pracach, oczywiście bez wyrażenia zgody – trenowano algorytmy[11]. Jak wiadomo z poprzedniego rozdziału, trzeba przepuścić ogromną ilość danych, aby wyszkolić program. Co jednak warte jest zaznaczenia, proces uzyskiwania tak wielu zgód od

[11] D. Milmo, *Thom Yorke and Julianne Moore join thousands of creatives in AI warning*, https://www.theguardian.com/film/2024/oct/22/thom-yorke-and-julianne-moore-join-thousands-of-creatives-in-ai-warning [22 października 2024 r.].

poszczególnych twórców byłby nie tylko szalenie czasochłonny, ale sama idea rzetelnych działań korporacji w tym zakresie posiada znamiona myślenia życzeniowego. Jak twierdzi raport analityczny[12] Dylana Patel'a, i Afzala Ahmad'a jeden dzień pracy chatuGPT może kosztować nawet 700 000 dolarów. Być może tańsze i szybsze okaże się wypłacanie artystom odszkodowań, o ile uda im się udowodnić swoje racje na drodze sądowej. Obecnie nie stworzono regulacji, które byłyby w stanie zapewnić artystom wystarczającej ochrony przed AI. Dlatego takie sprawy, zważając na obecne przepisy prawne, mogą być ciężkie lub zwyczajnie niemożliwe do wygrania.

Internet jest obecnie największym miejscem do dzielenia się swoimi artystycznymi pracami. Dzielenia, do którego dochodzi często bez naszej zgody. Nietrudno przecież za pomocą przysłowiowego kliknięcia myszki pobrać obrazy i teksty z Internetu bez żadnych konsekwencji i zmartwień o prawa autorskie. A jak udowodnić potem korporacji, że korzystała z zasobów objętych prawem autorskim? Nie mamy wglądu ani do procesu szkolenia AI, ani do danych, na których odbywają trening. Docieramy do momentu w historii, w którym prawo mające za zadanie chronić twórczość człowieka nie nadąża za nowinkami technologicznymi. Oczywiście firmy zajmujące się rozwojem AI czerpią z tego korzyści.

W pisemnym zeznaniu Sama Altmana – dyrektora generalnego firmy OpenAi – z przesłuchania w sprawie: *Nadzór nad SI: Zasady dotyczące Sztucznej Inteligencji*[13] (tłum. własne) z dnia 16.05.2023 roku można znaleźć fragment o zasadach trenowania AI:

[12] D. Patel, *The Inference Cost Of Search Disruption – Large Language Model Cost Analysis*, https://www.semianalysis.com/p/the-inference-cost-of-search-disruption [28 lipca 2024 r.].

[13] *Oversight of A.I.: Rules for Artificial Intelligence | United States Senate Committee on the Judiciary*, https://www.judiciary.senate.gov/committee-activity/hearings/oversight-of-ai-rules-for-artificial-intelligence [28 lipca 2024 r.].

„Nasze modele są trenowane na szerokim zakresie danych, które obejmują publicznie dostępne treści, licencjonowane treści i treści generowane przez ludzkich recenzentów. Tworzenie tych modeli wymaga nie tylko zaawansowanego projektowania algorytmów i znacznych ilości danych treningowych, ale także znacznej infrastruktury obliczeniowej do trenowania modeli, a następnie obsługi ich dla milionów użytkowników" (tłum. własne).

Na stronie internetowej OpenAI znajduje się tekst dotyczący kwestii prywatności oraz użycia danych internetowych do treningu algorytmów[14], w którym wspominają o korzystaniu z danych opublikowanych w sieci. Są to dość powierzchowne i ogólnikowe stwierdzenia, które w pewnym stopniu mogą stanowić zabezpieczenie interesów firmy OpenAI. W sieci przecież można znaleźć praktycznie wszystko, a w szczególności prace objęte prawami autorskimi, które zostały tam udostępnione bez zgody i wiedzy twórców (piractwo), na których następnie szkolono - również bez zgody - algorytmy. Zjawisko to niemniej budzi wątpliwości. Zastanawiający jest bowiem fakt, że założyciele wielkich korporacji bez najmniejszych skrupułów trenowali algorytmy na danych internetowych, a oczywistością jest, że sieć to ogromny zasób twórczości ludzkiej, szczególnie tej chronionej prawem autorskim. Bezkarność wspomnianych czynów zachęca wyłącznie do ich praktykowania, a można by pokusić się o stwierdzenie, że jest to po prostu kradzież.

W 2020 roku użytkownicy platformy Tweeter mogli natknąć się na post Shawna Presser'a o następującej treści: *„załóżmy, że chcesz wytrenować światowej klasy model GPT, tak jak OpenAI. Jak? Nie masz żadnych danych. Teraz masz. Teraz wszyscy mają. Przedstawiamy „books3", aka „all of bibliotik"*. W poście dodatkowo umieszczono informacje o liczbie pozycji książkowych zawartych w zbiorze —196 640 — oraz

[14] *Our approach to AI safety*, https://openai.com/index/our-approach-to-ai-safety/ [28 lipca 2024 r.].

o możliwości pobrania zbioru pod konkretnym adresem[15], który dzisiaj nie jest już aktywny[16].

Link prowadzi do strony nazwanej „The Eye" (Oko). Na stronie Oka możemy znaleźć opis, mówiący, że są organizacją non-profit zajmującą się archiwizacją cyfrowego dziedzictwa i pomysłów. „Jesteśmy cyfrowymi bibliotekarzami". Cyfrowi bibliotekarze w pliku o rozmiarze 37 GB zgromadzili 196 640 pozycji literackich objętych prawami autorskimi. Zbiór był możliwy do pobrania dla każdego, ale dedykowany był dla środowiska zajmującego się rozwijaniem modeli językowych. Ze zbioru Books3 korzystała chociażby Meta Marka Zuckerberga.

O problemie nielegalnego wykorzystania twórczości pisarzy doniosła duńska organizacja wspierająca branżę kreatywną w sieci - Right Alliance. Na podstawie naruszenia praw autorskich, Right Alliance położyła kres nielegalnemu udostępnianiu olbrzymiego zbioru literatury w sieci[17]. Sprawę następnie opisał serwis internetowy TorrentFreak[18], a w ślad za nim sprawę opisał The Antlantic[19]. Dziennikarz Alex Reisner napisał artykuł dla amerykańskiego magazynu, ale co najważniejsze, konkretnie dotarł do książek, które zostały nielegalnie zawarte w zbiorze danych. Books3 był jednym wielkim plikiem tekstowym, z którego Reisner wydobył numery ISBN (Międzynarodowy znormalizowany numer książki). Następnie sprawdzał jaki numer odpowiada danej pozycji. Dziennikarzowi udało się

---

[15] *404 Not Found*, https://the-eye.eu/public/AI/pile_preliminary_components/books3.tar.gz [2 wrzesienia 2024 r.].

[16] Zwrócę na ten problem uwagę w dalszej części niniejszego rozdziału.

[17] Victoria, *Rights Alliance removes the illegal Books3 dataset used to train artificial intelligence | RettighedsAlliancen*, https://rettighedsalliancen.com/rights-alliance-removes-the-illegal-books3-dataset-used-to-train-artificial-intelligence/ [30 lipieca 2024 r.].

[18] *Anti-Piracy Group Takes Prominent AI Training Dataset ' „Books3" Offline * TorrentFreak*, https://torrentfreak.com/anti-piracy-group-takes-prominent-ai-training-dataset-books3-offline-230816/ [30 lipieca 2024 r.].

[19] A. Reisner, *The Authors Whose Pirated Books Are Powering Generative AI*, https://www.theatlantic.com/technology/archive/2023/08/books3-ai-meta-llama-pirated-books/675063/ [30 lipca 2024 r.].

wydobyć z morza tekstu aż 183 000 pozycji. Reszta zbioru, 13640 książek, pozostała niezidentyfikowana. Na światło dzienne wyszło, że wykorzystano pracę np. Stephena Kinga, Zadie Smith, Michaela Pollana, Margaret Atwood, Sarah Silverman, Richarda Kadrey'a i innych. Reisner udostępnił również na stronie The Atlantic wyszukiwarkę, która umożliwia sprawdzenie, jacy autorzy i jakie konkretne dzieła umieszczono w zbiorze danych.

Sprawa dotycząca „Books3" została nagłośniona przez inne zagraniczne wydawnictwa. Polscy redaktorzy natomiast niezmiennie milczą. Postanowiłam więc skorzystać z pracy amerykańskiego dziennikarza i starałam się wyszukać polskich pisarzy. Udało mi się dotrzeć do 24 polskich artystów, których prace znalazły się w nielegalnym zbiorze danych. Jako pierwsza w Polsce przedstawiam listę polskich artystów i ich konkretne dzieła, które zostały ukradzione, i wykorzystane do szkolenia algorytmów. Lista ta, nad czym ubolewam, jest zapewne niepełna. Proces odszukiwania konkretnego autora odbywa się poprzez wpisanie imienia i nazwiska. Przez długi czas starałam się wprowadzać dane polskich pisarzy, którzy jak przypuszczałam, mogli znaleźć się w pirackim pliku. Na pewno nie udało mi się odnaleźć wszystkich Polaków, ale moje odkrycie wystarczy by ukazać skalę zjawiska i faktu, że kradzież w imię rozwoju technologicznego dotyczy nas wszystkich. Jestem przekonana, że żyjący artyści z poniższej listy nie są świadomi, o kradzieży ich prac. A przecież mogliby pójść śladem amerykańskich autorów, którzy w sądzie próbują dowodzić swoich praw.

Lista przedstawiona w postaci tabeli

AUTOR	**DZIEŁA**
Jerzy Pilch	„Tysiąc spokojnych miast" „Moje pierwsze samobójstwo" „Pod Mocnym Aniołem"
Jerzy Kosiński	„Wystarczy być" „Randka w ciemno" „Cockpit" „Oral Pleasure: Kosinski as Storyteller" "Przechodząc obok – eseje wydane pośmiertnie" „Gra" „Kroki" „Diabelskie drzewo" „Malowany ptak"
Józef Czapski	„Na nieludzkiej ziemi" „Stracony czas: Wykłady o Prouście z sowieckiego Gułagu"
Miron Białoszewski	„Pamiętnik z powstania warszawskiego"
Andrzej Stasiuk	„Dukla"
Cyprian Kamil Norwid	Zbiór wierszy
Bronisław Malinowski	„Argonauci zachodniego Pacyfiku" „Wolność i cywilizacja" „Mit w psychologii pierwotnej" „Życie seksualne dzikich w północno-zachodniej Melanezji"
Ryszard Kapuściński	„Jeszcze dzień życia"

	„Szachinszach” „Cesarz” „Ten Inny” „Heban” „Wojna futbolowa” „Podróże z Herodotem”
Stanisław Lem	„Cyberiada” „Fiasko” „Kongres Futorologiczny” „Opowieści o pilocie Pirxie” „Mortal Engines” „Powrót z gwiazd” „Śledztwo” „Dzienniki gwiazdowe”
Henryk Sienkiewicz	„Relatos”
Bolesław Prus	„Lalka”
Wisława Szymborska	„Monolog psa zaplątanego w dzieje” „Lektury nadobowiązkowe”
Zygmunt Miłoszewski	„Ziarno prawdy. Cykl Teodor Szacki. Tom 2” „Uwikłanie. Teodor Szacki”
Sławomir Mrożek	„Słoń”
Andrzej Sapkowski	„Chrzest ognia” „Krew elfów” „Ostatnie życzenie” „Narrenturm” „Sezon burz” „Miecz przeznaczenia” „Pani Jeziora” „Ostatnie życzenie”

	„Maladie i inne opowiadania” „Czas pogardy” „Wieża Jaskółki”
Czesław Miłosz	„Góry Parnasu”
Witold Gombrowicz	„Przewodnik po filozofii w sześć godzin i kwadrans” „Bakakaj” „Kosmos” „Dziennik” „Ferdydurke” „Wspomnienia polskie” „Pornografia” „Trans-Atlantyk”
Tadeusz Różewicz	„New Poems” (wersja angielska)
Zbigniew Herbert	„The Collected Poems: 1956-1998” “The Collected Prose: 1948-1998”
Zygmunt Bauman	„Wspólnota. W poszukiwaniu bezpieczeństwa w niepewnym świecie” „Życie na przemiał” „Kultura w płynnej nowoczesności” „Czy bogactwo nielicznych służy wszystkim?” „Globalizacja: i co z tego dla ludzi wynika” „Płynny lęk” „Płynne życie” „Płynna nowoczesność” „Płynne czasy” „Żyjąc w czasie pożyczonym: Rozmowy z Citlali Rovirosa-Madrazo” „Nowoczesność i ambiwalencja”

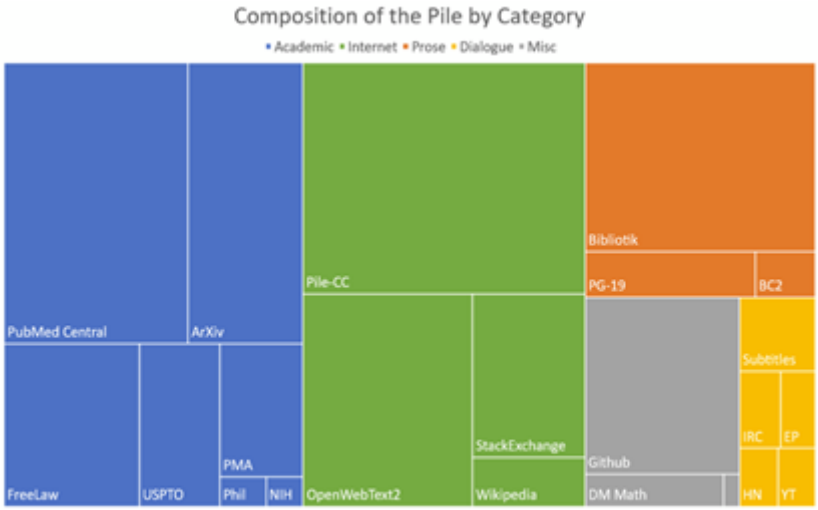
	„Nowoczesność i Zagłada” „Ponowoczesność jako źródło cierpień” „Retrotopia” „Sztuka życia” „Społeczeństwo w stanie oblężenia” „To nie jest dziennik” „Życie na przemiał” „Tiempos líquidos”
Bruno Schultz	„Ulica krokodyli” „Sanatorium pod klepsydrą”
Marek Hłasko	„Wszyscy byli odwróceni” „Cmentarze”
Gustaw Herling - Grudziński	„Inny Świat”
Dorota Masłowska	„Wojna polsko-ruska pod flagą biało-czerwoną”

IV. ARTYŚCI NA DRODZE SĄDOWEJ PRZECIWKO WYKORZYSTANIU SWOICH PRAC NA RZECZ ROZWOJU TECHNOLOGII AI

To jeszcze nie koniec afery dotyczącej zbioru Books3. Chociaż został on usunięty z pierwotnego źródła (nadal można znaleźć na różnych nielegalnych stronach link do pobrania) jest on częścią tzw. „The Pile” (z ang. Stos). The Pile zostało stworzone i udostępnione w 2020 roku przez firmę EleutherAI, na potrzeby szkolenia modeli LLM[20]. Zbiór zawierał 22 wysokiej jakości podzbiory, w tym Books3- pod nazwą Bibliotik. Poniżej fragment grafiki zamieszczonej przez EleutherAI w dokumencie opisującym skład zbioru[21].

[20] L. Gao i in., The Pile: An 800GB Dataset of Diverse Text for Language Modeling, 31 grudnia 2020 r., arXiv:2101.00027 [cs].

[21] *Ibidem*.



Ilustracja 2: https://arxiv.org/pdf/2101.00027

The Pile było wykorzystane przez wielkie korporacje takie jak: Meta, Apple, Microsoft. Uniwersytet Stanford oraz Pekińska Akademia Sztucznej Inteligencji również korzystały z tego zbioru danych do szkolenia swoich algorytmów. Gdy wyszło na jaw, że potajemnie wykorzystano prace artystów na potrzeby zysku dużych korporacji, poszkodowani zaczęli składać pozwy sądowe. W poprzednim roku artyści: Sarah Silverman, Richard Kadrey oraz Chrostopher Golden złożyli pozew przeciwko Meta Platforms, INC. Znany pod nazwą „Kadrey vs. Meta" lub „Sarah Silverman Case"[22]. Artyści oskarżyli Meta Platrofms, INC o:

1. Bezpośrednie naruszenie praw autorskich
2. Pośrednie naruszenie praw autorskich
3. Naruszenie ustawy DMCA (ang. *Digital Millennium Copyright Act*)
4. Czyn nieuczciwej konkurencji
5. Bezpodstawne wzbogacenie się
6. Niedbalstwo

[22] *Complaint – #1 in Kadrey v. Meta Platforms, Inc. (N.D. Cal., 3:23-cv-03417) – CourtListener.com*, https://www.courtlistener.com/docket/67569326/1/kadrey-v-meta-platforms-inc/ [2 września 2024 r.].

W listopadzie 2023 sąd na wniosek Meta Platforms, Inc, oddalił wszystkie roszczenia, z możliwością jedynie zmiany roszczeń dotyczących bezpośredniego naruszenia praw autorskich. Punkt 4,5,6 zostały odrzucone, ponieważ według sądu ustawa o amerykańskim prawie autorskim wykluczała roszczenia. Punkt 3 oddalono, ponieważ model językowy firmy Meta-LLaMA- nie generuje oraz nie rozpowszechnia treści książek artystów. Punkt 2 został oddalony, ponieważ powodowie nie przedstawili konkretnych dowodów ani szczegółowych porównań świadczących o znacznym podobieństwie między objętymi ochroną autorską publikacjami artystów, a wynikami generowanymi przez LLaMa. W konsekwencji oddalenia praktycznie wszystkich roszczeń dotyczących firmy Meta z powyższego postępowania, wysuwa się wniosek, że narzędzia AI nie mogą stanowić dzieł zależnych naruszających prawa autorskie materiałów, na których są trenowane.

Sprawa „Kadrey vs Meta" dalej jest w toku i połączyła się ze sprawą „Huckabee vs Meta Platforms, INC" [23]. Pozostaje nam tylko czekać na rozwój wydarzeń i ostateczne orzeczenie sądu. Nowe technologie przyniosły nowe wyzwania i to właśnie teraz formułowane jest podejście prawne dotyczące rozwoju sztucznej inteligencji. Powyższa sprawa dla wielu wydawała się być prosta do wygrania. Jak widzimy, nie jest to takie oczywiste, jak można by przypuszczać na pierwszy rzut oka.

Pozwy kilku artystów przeciwko dużym korporacjom ukazują dużą dysproporcję sił. Wielkie firmy posiadają środki oraz całe armie prawników, którzy będą wynajdywać jak najwięcej kruczków prawnych na korzyść swojego klienta. Udźwignięcie aspektu finansowego z oczywistych względów też jest łatwiejsze w przeciwieństwie do kilku artystów. Siły zaczynają się wyrównywać, kiedy to korporacja rozpoczyna proces przeciwko drugiej korporacji. 22 czerwca 2024 roku

[23] *Order on Stipulation – #107 in Kadrey v. Meta Platforms, Inc. (N.D. Cal., 3:23-cv-03417) – CourtListener.com*, https://www.courtlistener.com/docket/67569326/107/kadrey-v-meta-platforms-inc/ [2 września 2024 r.].

Universal Music Group, Warner Music, Sony Music i inni, pozwali Suno i Udido[24]. Są to dwa systemy AI generujące muzykę. Działanie Suno i Udido opiera się na wpisaniu przez użytkownika własnego tekstu oraz określenia stylu, w jakim ma zostać wygenerowana piosenka. Powodowie przedstawili konkretne argumenty bazując na przykładach muzycznych ze swoich wytwórni skonfrontowanych z generatywnymi wynikami oprogramowań. Sprawa jest dalej w toku, a świat muzyczny czeka na rozwój wydarzeń. Jest to pierwsze działanie sądowe przeciwko generowaniu muzyki przez sztuczną inteligencję. Jeśli giganci tacy jak Warner nie będą w stanie wywalczyć sprawiedliwości i zaprzestać kradzieży dzieł artystycznych, to kto podoła?

W dziedzinie sztuk wizualnych artyści: Sarah Andersen, Kelly McKernnan oraz Karla Ortiz pozwali 13 stycznia 2023 roku StabilityAI (twórcy Stable Difussion), Midjourney (twórca Midjourney), DevianArt (twórca DreamUp) oraz Runway AI (zaangażowny w tworzenie Stable Difussion)[25], o wykorzystanie ich prac do szkolenia algorytmów generujących grafiki, zdjęcia czy też obrazy. Podobnie jak w przypadku sprawy Kadrey vs Meta, powodowie muszą przygotować szczegółowe dowody użycia ich prac wiążące się ze znaczącym podobieństwem wygenerowanych przez AI obrazów.

Obecnie (stan na 21.10.2024 r.) w Stanach Zjednoczonych toczą się 34 rozprawy przeciwko wykorzystaniu prac artystów do rozwoju AI. Pozwani są również: Microsoft, Nvidia, OpenAI, Google, Antropic oraz Bloomberg[26].

---

[24] *Complaint – #1 in UMG Recordings, Inc. v. Suno, Inc. (D. Mass., 1:24-cv-11611) – CourtListener.com*, https://www.courtlistener.com/docket/68878608/1/umg-recordings-inc-v-suno-inc/ [2 września 2024 r.].

[25] *Complaint – #1 in Andersen v. Stability AI Ltd. (N.D. Cal., 3:23-cv-00201) – CourtListener.com*, https://www.courtlistener.com/docket/66732129/1/andersen-v-stability-ai-ltd/ [2 września 2024 r.].

[26] Informacje te są dostępne na stronie internetowej zajmującej się systematyzowaniem aktualności dotyczących pozwów przeciwko Sztucznej Inteligencji w Stanach Zjednoczonych, https://chatgptiseatingtheworld.com/2024/10/21/status-of-all-33-copyright-lawsuits-v-ai-oct-21-2024/ [21 października 2024 r.].

Przyszłe wyroki sądów odegrają kluczową rolę w kwestii interpretowania prawa w przypadku nowych technologii generatywnej sztucznej inteligencji. Do tego czasu powyższe firmy oraz nowo powstające start-upy będą czerpać korzyści z długotrwałych rozpraw. Większość użytkowników Internetu bardzo chwali sobie korzystanie z generatywnej AI. Można odnieść wrażenie, że tylko artyści próbują stanąć na drodze rozwoju sztucznej inteligencji. Użytkownicy są świadomi kradzieży własności intelektualnej i bardzo często potępiają na forach postępowanie firm z doliny krzemowej. Mimo, że większość wie, że rozwój AI odbywa się poprzez kradzież twórczości, to praktycznie każdy młody człowiek używa Chatu GPT oraz innych narzędzi, przyczyniając się tylko do rozwoju i popularyzowania tych technologii.

V. ZASTOSOWANIE NARZĘDZI AI DO USPRAWNIENIA PROCESU PRODUKCJI FILMOWEJ

X Muza przeżyła już kilka znaczących rewolucji. Rozwój technologiczny wymuszał zmiany w sposobie produkcji filmów oraz stale redefiniował kino. Pierwszym ważnym przełomem oraz wyzwaniem dla światowej branży filmowej, pozostaje bez wątpienia rewolucja dźwiękowa. Wielu aktorów musiało zrezygnować z pracy, co wiązało się nierzadko z przypadkami samobójstw tych artystów, którzy utracili zarówno dochód, jak i status wielkich gwiazd filmowych wśród publiczności (przypis- pitaval). Skończyła się era pracy na głośnych planach zdjęciowych, a rozpoczęła era ciszy na planie. Dodatkowy czynnik, jakim był dźwięk, wymusił również inne sposoby pracy na montażem. Dzięki temu twórcy mogli wprowadzić nowe środki wyrazu: słyszalne (nie werbalizowane za pomocą plansz tekstowych) dialogi między aktorami, muzykę, efekty dźwiękowe bezpośrednio na właściwy

filmowi nośnik fizyczny, co wpłynęło na lepszy odbiór ostatecznego dzieła przez widza i wprowadziło nowe rozwiązania w samej narracji filmowej[27].

W latach trzydziestych XX wieku wprowadzono do kina kolor. Skomplikowany proces zwany „Technicolor" pozwalał na uzyskiwanie bardziej naturalistycznego, barwnego obrazu. Był to na początku luksus, na który pozwolić sobie mogły jedynie najbardziej rozwinięte i posiadające największy kapitał kinematografie. Proces polegał na naświetlaniu trzech taśm jednocześnie, przez wielkie specjalistyczne kamery z trzema filtrami o barwie zielonej, niebieskiej i czerwonej, czyli powszechnie znany RGB. Nowe środki wyrazu znowu wiązały się ze zmianą podejścia do produkcji, od aspektu finansowego po aspekty reżyserskie. Dopiero w latach 50-tych i 60-tych XX wieku kolor stał się nowym standardem.

Jednak jedną z najbardziej znaczących zmian technologicznych, która, można stwierdzić, miała największy wpływ na zmianę przemysłu filmowego, było przejście z nagrywania na kamerach na taśmie światłoczułej do nagrywania na kamerach cyfrowych. Pozwoliło to w znacznym stopniu zmniejszyć koszty produkcji, dało możliwość większej swobody, chociażby przez brak ograniczenia przez ilość materiału światłoczułego twórcy mogli pozwolić sobie na większą ilość powtórzeń, czyli dubli. Skomplikowane procesy obróbki taśmy filmowej przeniosły na komputery, co pozwoliło na większą wygodę np. w procesie montażu. Otworzyły się nowe możliwości szczegółowej koloryzacji obrazu filmowego. Kolejna rewolucja pozwoliła na wprowadzenie nowych środków wyrazu, które wcześniej były niedostępne. Taśma przestała zatem dominować na rynku filmowego stając się raczej świadomym wyborem estetycznym niż nośnikiem podyktowanym przez konieczność. Mimo wielu zalet technologii cyfrowej, wielu twórców do dzisiejszego dnia nie pogodziła się ze zmniejszeniem znaczenia taśmy światłoczułej.

[27] T. Lubelski, I. Sowińska, R. Syska, *Przełom dźwiękowy* [w:] *Historia kina, tom 2. Kino klasyczne*, Kraków 2012.

Tak jak w przypadku nowych technologii AI, możemy nie być pogodzeni ze zmianami, możemy uporczywie upierać się przy tym, co nam znane, lecz prawidła rynku okazują się znów bezlitosne. Film od początku dążył do jak największego uproszczenia procesów produkcji. Nie powinno nas zatem dziwić, że kiedy powstają narzędzia mogące usprawnić i przyspieszyć proces produkcji filmu, zostają one przyjęte przez branżę, a przynajmniej jej najbardziej decydujące podmioty.

VI. LARGOAI, CZYLI WSZECHSTRONNA POMOC DLA PRODUCENTÓW I DYSTRYBUTORÓW

Właśnie w tym kontekście warto przyjrzeć się nowym technologiom AI, które coraz częściej wykorzystuje się w procesie produkcji filmowej oraz pochodnym narzędziom mogącym w przyszłości usprawniać, a może nawet przyczynić się do trwałych zmian wewnątrz przemysłu filmowego. Wybrane przykłady obecnie funkcjonujących modeli zostały przy tym omówione w nieprzypadkowej kolejności: jako pierwsza zostanie opisana platforma *Largo*.ai, która zostanie również opisana w sposób bardziej szczegółowy względem dość podobnych, jeżeli chodzi o funkcjonowanie, modeli przedstawionych później. Uzasadnione jest to przede wszystkim wyjątkową, ale i symptomatyczną specjalizacją platformy *Largo*, która skierowana jest ściśle do przedstawicieli środowiska producentów filmowych.

Largo.ai[28] to interaktywna platforma zaprogramowana w swoim pierwszym wariancie w roku 2018, a będąca dziełem akademików związanych z Politechniką Federalną w Lozannie. Jest to efektywne narzędzie pracy przede wszystkim dla producentów utworów audiowizualnych, których przedsięwzięcia znajdują się w okresie przygotowawczym, a zatem istotnym momencie dla przyszłości projektu, bo nierzadko decydującym o jegoż strukturze finansowania oraz skali późniejszej dystrybucji oraz promocji. To, co bowiem wyróżnia *Largo.ai* to możliwość

[28] *Largo.ai: about us*: https://home.largo.ai/about-us/ [12 sierpnia 2024 r.].

wygenerowania, w oparciu o nowatorskie oprogramowanie, rozbudowanej analizy strategiczno-marketingowej na podstawie danych wejściowych dotyczących konkretnego projektu. (scenariusz lub materiał filmowy jako dopuszczalne formy referencyjne)

Opisywany model nie poprzestaje przy tym wyłącznie na elementarnej analizie stylometrycznej tekstu bądź innego rodzaju referencji, ale wyciąga niezwykle przydatne wnioski mogące służyć za prognozę przychodów przyszłego dzieła oraz jego ogólnego przyjęcia tak wśród publiczności, jak i krytyków. W równym stopniu *Largo.ai* stanowi platformę doradczą, która sugerując się, znowuż nieopisaną w sposób bardziej szczegółowy, bazą poszczególnych tytułów oraz ich udokumentowanych recepcji proponuje m.in. zmiany w scenariuszu lub wybory obsadowe, które wpłyną korzystnie na odbiór dzieła, a w dalszej perspektywie zapewnią mu sukces finansowy i artystyczny. Kolejnym udogodnieniem dla producenta, który decyduje się korzystać z *Largo*, są liczne wykresy, oferujące czytelną informację zwrotną odnoście tak różnorodnych kwestii jak: amplituda emocjonalna potencjalnych widzów na podstawie analizy projektu, szczegółowy profil grupy docelowej czy też ścisła taktyka dystrybucyjno-promocyjna, zapewniając akcelerację przychodów oraz wzrost świadomości odbiorcy dziełem[29].

Co tym bardziej interesujące – i tym bardziej atrakcyjne dla użytkowników – analiza rynkowa projektu odbywa się z uwzględnieniem specyfiki nie tyle uogólnionych rynków światowych, lecz lokalnego rynku na terenie dowolnie wybranego państwa. Innymi słowy, producent pragnący poznać rynkowe szanse i zagrożenia dla projektu filmowego realizowanego w Polsce i z polskim kapitałem nie musi ograniczać się *sensu stricto* do realiów polskiego rynku kinematograficznego, lecz ma szansę zbadać mniej lub bardziej egzotyczne rynki za granicą. Nie jest, wobec tego

[29] *Largo.ai: Accelerate your film development and financing with AI insights:* https://home.largo.ai/accelerate-your-film-development-and-financing-with-ai-insights/ [12 sierpnia 2024 r.].

wykluczone poszukiwanie alternatywnej ścieżki dystrybucyjnej, z której skądinąd nierzadko korzystają producenci polskiego kina artystycznego. Wspomniana ścieżka dystrybucyjna dotyczyła chociażby filmu „Ida", który po chłodnym przyjęciu w kraju produkcji, zyskał uznanie na zagranicznych przeglądach i festiwalach filmowych, aby wreszcie zdobyć jedną z najbardziej prestiżowych, a przynajmniej najbardziej znanych filmowych nagród i triumfalnie powrócić do rodzimej dystrybucji kinowej[30]. Nie będzie przesadą stwierdzenie, że przy użyciu funkcji analizy rynków lokalnych na platformie *Largo* producent może przynajmniej podjąć próbę skoordynowania podobnego procesu, z pewnością zaś może on odpowiednio dopracować dystrybucyjne oraz promocyjne aspekty wynikające z realizacji koprodukcji o różnej skali oraz liczbie partnerów międzynarodowych.

Podkreślona już kilkukrotnie wielofunkcyjność zastosowań *Largo*, zaznaczona zresztą w nazwie, zdaje się być odzwierciedlona przez panoramiczne spojrzenie (oraz daleko posunięte ambicje) twórców modelu, którzy, sami postawieni przed wyzwaniem promocji własnego przedsięwzięcia wśród wpływowych osób z branży, nie ograniczyli konsekwentnej ingracjacji do zaoferowania usług potężnym audiowizualnym koncernom (których przedstawiciele, co zostało zaznaczone na poprzedniej stronie, w naturalny sposób zainteresowali się ewentualną współpracą), lecz, co zostało uwzględnione w odpowiedniej zakładce na oficjalnej witrynie internetowej *Largo*, zaprezentowali dzieło swojego wspólnego wysiłku, znamiennie, na 72. Międzynarodowym Festiwalu Filmowym w Cannes (w roku 2019)[31] oraz 70. Międzynarodowym Festiwalu Filmowym w Berlinie (w roku 2020)[32] – dwóch europejskich festiwalach filmowych nie tylko długiej tradycji pielęgnowania artyzmu

[30] P. Bez, *Światowa dystrybucja „Idy"*: https://pisf.pl/aktualnosci/swiatowa-dystrybucja-idy/ (dostęp: 12.08.2024).

[31] *Largo.ai: Pitched at Cannes Film Festival:* https://home.largo.ai/pitched-at-cannes-film-festival/ (dostęp: 12.08.2024).

[32] *Largo.ai: Largo selected as one of Horizon 2020 Start-ups in Berlinare:* https://home.largo.ai/largo-selected-as-one-of-horizon-2020-start-ups-in-berlinale/ (dostęp: 12.08.2024).

oraz ekspresyjnej śmiałości w kinematografii, lecz pilnie śledzących czysto technologiczne przemiany w branży. Ponadto, w roku 2019, a zatem roku premiery platformy na festiwalu canneńskim, twórcy *Largo* zostali wyróżnieni nagrodą za najlepszy okołofilmowy start-up na 67. Międzynarodowych Festiwalu Filmowym w San Sebastián w ramach debiutującego wówczas na tymże festiwalu, a postulującego futurystyczne podejście do rynku kinematograficznego, konkursu *Zinemaldia Startup Challenge*[33].

Futurystyczność samej platformy jest niemniej sporna – działa ona na zasadzie technologii, lakonicznie nazywanej przez twórców, *SAAS* (skrótowiec od angielskiego *software as a service*, a zatem *oprogramowanie jako usługa*). Ową technologię, rozwijaną systematycznie w pierwszych dekadach XXI wieku, charakteryzuje natomiast „odciążanie" użytkownika z konieczności instalowania usługi oraz przymusu cyklicznej aktualizacji oprogramowania – oczywiście, to wszystko po uiszczeniu opłaty licencyjnej. Efektem ubocznym korzystania z potencjalnej ciągłości w dostępie do zakupionej usługi, zdaniem amerykańskiego aktywisty oraz propagatora idei *wolnego oprogramowania* (ang. *free software*) – Richarda Stallmana jest jednak zagrożenie utratą kontroli użytkownika m.in. nad przetwarzaniem przez platformę osobistych danych (czyli, w przypadku platformy *Largo*, scenariuszy bądź scenopisów niezrealizowanych jeszcze utworów audiowizualnych)[34].

Choć podejście rzeczonego aktywisty zyskiwało niejednokrotnie miano nadmiernego radykalizmu[35], zasadne wydaje się pytanie o ochronę użytkowników również, a może zwłaszcza tak specjalistycznych, platform przed kradzieżą danych.

---

[33] *Largo.ai: Great award for Largo from San Sebastian Film Festival:* https://home.largo.ai/great-award-for-largo-from-san-sebastian-film-festival/ [12 sierpnia 2024 r.].

[34] *Interviewing Richard Stallman: Freedom in the time of SaaS*: https://meedabyte.com/2010/05/25/interviewing-richard-stallman-freedom-in-the-time-of-saas/ [12 sierpnia 2024 r.].

[35] A. Wolfe: *Richard Stallman Slams SaaS:* https://www.informationweek.com/it-leadership/richard-stallman-slams-saas [12 sierpnia 2024 r.].

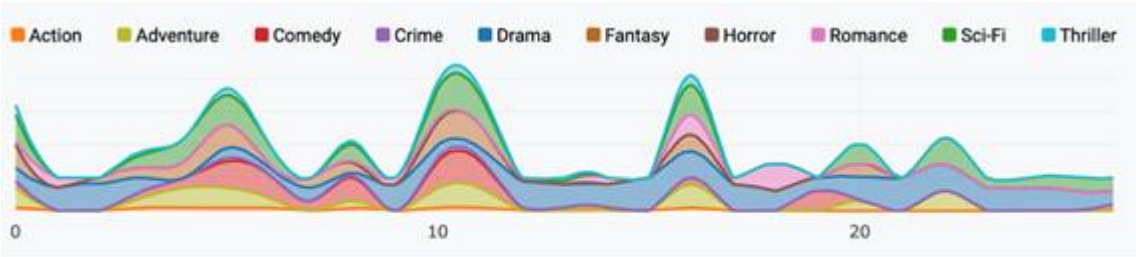
Pozornie odznaczająca się demokratyzacją platforma, do której nieutrudniony dostęp mają zarówno doświadczeni producenci, przedstawiciele wielokanałowych koncernów medialnych, jak i dopiero stawiający swoje pierwsze kroki w branży twórcy nie może bowiem uchodzić za zupełnie bezpieczną, jeśli jest zarazem niezmiennie specjalistyczna. Niestety, łatwo wyobrazić sobie hipotetyczną sytuację, w której producenci będący wysłannikami Netflixa czy Amazona właśnie mają uprzywilejowaną pozycję (oraz, co za tym idzie, zapewnioną lepszą ochronę danych) korzystając z platformy względem producentów o mniejszym znaczeniu w branży (lub, po prostu, producentów *stricte* lokalnych), a nawet, po dokonaniu dodatkowej transakcji, producenci reprezentujący korporacje-molochy mogą z łatwością „uzyskać dostęp" do treści, które pochodzą od pozostałych użytkowników. Co pragnę zaznaczyć, jest to wyłącznie hipoteza, jednakże w dobie kradzieży danych oraz ogólnej cyberprzemocy na wielu szczeblach nie uważam, iż powinna być lekceważona.

Nie chcąc jednak poprzestawać na ogólnikach, warto zwrócić uwagę na szczególnie istotne narzędzia, które mają do dyspozycji użytkownicy. Rzeczony użytkownik może otrzymać analizę dotyczącą interesującego go etapu produkcji filmu z uwzględnieniem etapu developmentu, pre-produkcji, post-produkcji oraz dystrybucji. Aby otrzymać wyniki analizy, a jest to proces trwający zazwyczaj 10 minut, producent musi wprowadzić oprócz materiałów referencyjnych, podstawowe informacje takie jak: zakładany budżet czy też dane twórców. Po uzupełnieniu powyższych informacji generowana jest przez AI analiza w następujących segmentach:

1. Analiza gatunkowa

Otrzymane wyniki ukazują skalę i intensywność występowania poszczególnych gatunków na przestrzenni fabuły filmu. Jest to zatem przydatne

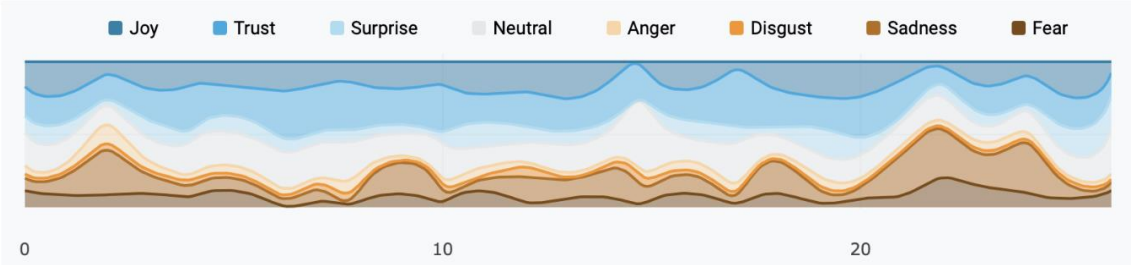
narzędzie weryfikujące pierwotne założenia gatunkowe scenariusza. Co warto zaznaczyć, jeśli wykres jest dość wyrównany sugeruje to znaczną spójność gatunku. Z drugiej pojawia się ryzyko, iż zbytnia monotonia nie utrzyma uwagi widza przez czas seansu. W tym przypadku istotne są właśnie pierwotne założenia dotyczące fabuły filmu, aby stwierdzić, czy otrzymane wyniki są satysfakcjonujące dla producenta i potencjalnego dystrybutora. Celem produkcji filmowej jest przecież ostatecznie właściwe odczytanie przez odbiorców dzieła, zgodne z intencjami twórców. Co za tym idzie weryfikacja od najwcześniejszego etapu może stanowić istotny element wpływający na sukces danej produkcji.



Ilustracja 4: https://fromtheheartproductions.com/largo-ai-artificial-intelligence-film-analysis

2. Analiza emocjonalna

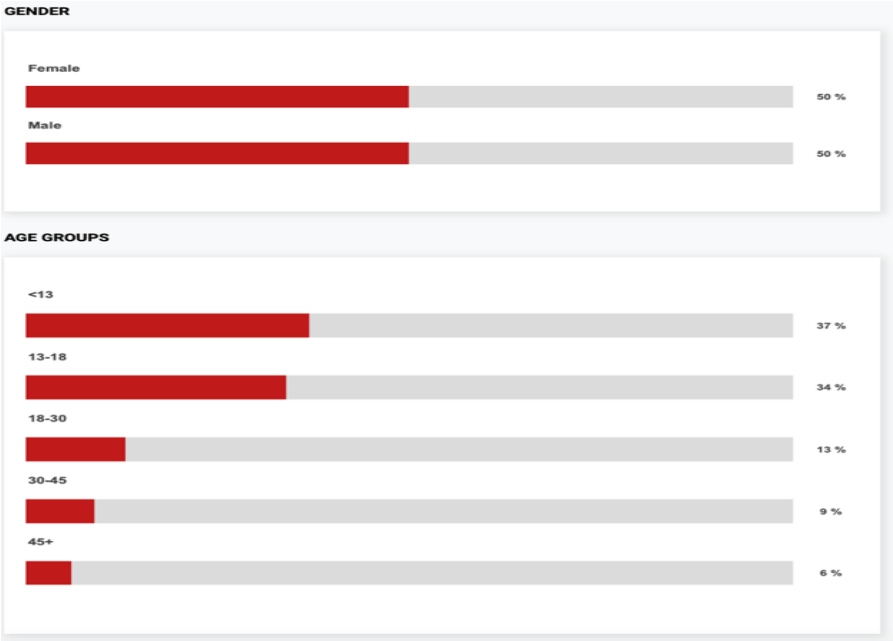
W przypadku analizy emocjonalnej otrzymujemy informacje dotyczące występowania w filmie poszczególnych emocji (są to radość, zaskoczenie, zaufanie, neutralny stosunek, gniew, wstręt, smutek i strach) oraz ich intensywności. Dzięki temu możliwe jest zrozumienie dynamiki emocjonalnej odczuwanej przez widza. Jest to nierozłącznie związane z systematyką tworzenia narracji filmowej. Odpowiednie tempo akcji, punkty zwrotne, momenty napięcia – czyli wszystko to co sprawia, że widz utożsamia się z historią. Na poniższym wykresie, kolory niebieskie odpowiadają pozytywnym, a kolory brązowe negatywnym emocjom. Wciąż najważniejsza jest jednak świadomość realizowanego gatunku, np. w analizie horroru powinny dominować negatywne emocje u widza.



Ilustracja 3: https://fromtheheartproductions.com/largo-ai-artificial-intelligence-film-analysis/

3. Demografia i dostosowania

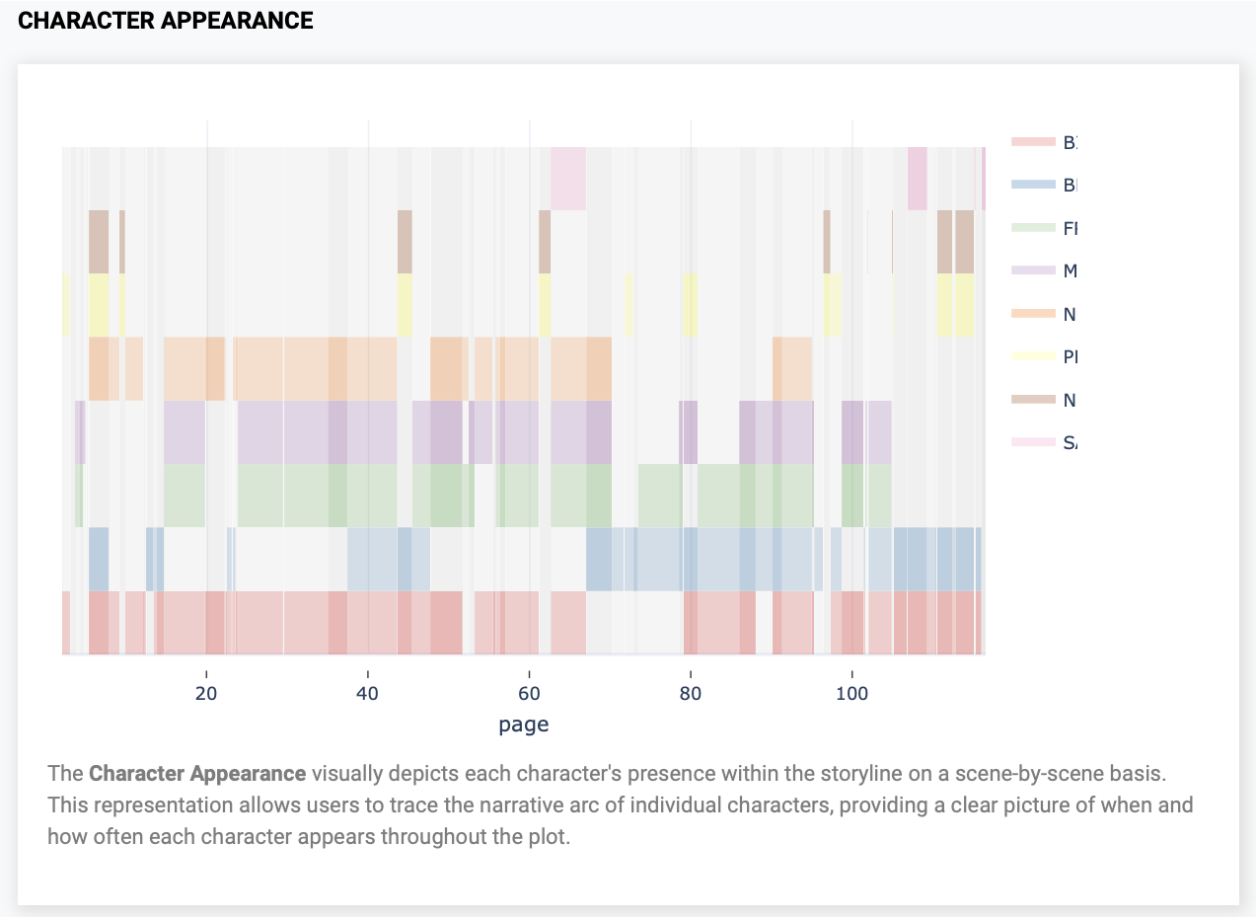
Jak sugeruje nazwa, w tym przypadku analiza skupia się na konkretnych grupach wiekowych i w oparciu o treść dzieła, diagnozuje odpowiedniość filmu dla danej grupy. Brane pod uwagę jest również zainteresowanie procentowe danej płci potencjalnym seansem. Narzędzie prawda ponadto dialogi pod kątem nałożenia możliwych ograniczeń wiekowych. Poznanie grup docelowych oraz ich preferencji skutkuje tym samym efektywniejszą dystrybucją filmu.



Ilustracja 4: https://fromtheheartproductions.com/largo-ai-artificial-intelligence-film-analysis/

4. Analiza charakteru postaci

Relacje między znaczącymi bohaterami filmowymi przedstawione są w tym segmencie za pomocą grafu nieskierowanego. Im większy okręg, tym ważniejsza jest postać. Grubość linii między poszczególnymi punktami, czyli bohaterami, ilustruje intensywność interakcji między nimi. Co się zaś tyczy przeznaczenia samej graficznej wizualizacji, nadaje ona efektywności dyskusjom producenta z potencjalnymi partnerami mogącymi wesprzeć realizację filmu, a dla których bardziej czytelny wyda się pojedynczy graf niż, nawet najlepiej przygotowana, ustna wypowiedź producenta na temat danej narracji oraz jej bohaterów. Kolejno do dyspozycji w sekcji jest jeszcze kilka innych narzędzi. Pierwsze to *character appearance*, czyli częstotliwość występowanie danych postaci w fabule.



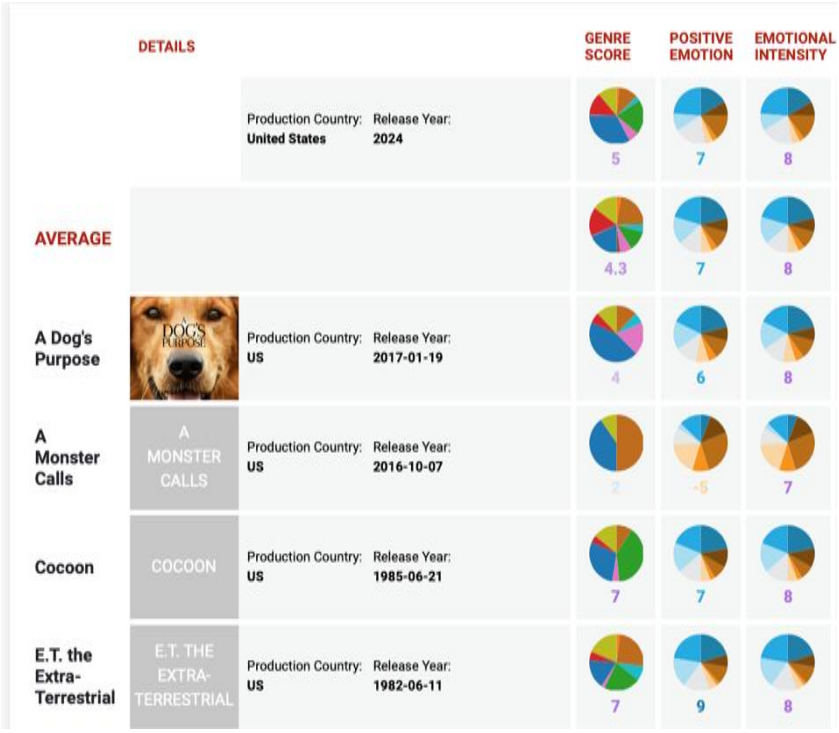
Ilustracja 5: https://fromtheheartproductions.com/largo-ai-artificial-intelligence-film-analysis/

Drugie narzędzie to propozycje castingów. Na podstawie danych zawartych w scenariuszy oraz danych finansowych zostają przygotowane propozycje aktorskie, które według sztucznej inteligencji będą najbardziej odpowiednie do roli, a co za tym idzie potencjalnie przyczynią się do sukcesu filmu. Producent również ma możliwość, aby narzucić AI pewne restrykcje, dotyczące wieku aktora, wyglądu aktora, limitu gaży aktorskiej. Dodatkowo z bazy dostępnych aktorów, która liczy 100 tysięcy osób, użytkownik platformy może wybrać własne propozycje, które zostaną także ocenione przez algorytm.

Funkcja związana z doborem aktorów jest oczywiście jeszcze bardziej rozbudowana. Porównuje koszty zatrudnienia aktorów, w oparciu o ich wcześniejsze produkcje, podaje również jaki jest możliwy zwrot z inwestycji zatrudnienia takiego

aktora. Możemy zobaczyć czy dany aktor jest teraz bardziej popularny, czyli czy przyciąga widownie do kina. Na pewno jest to czynnik, który warto rozważyć podczas procesu castingu.

5. Analiza porównawcza



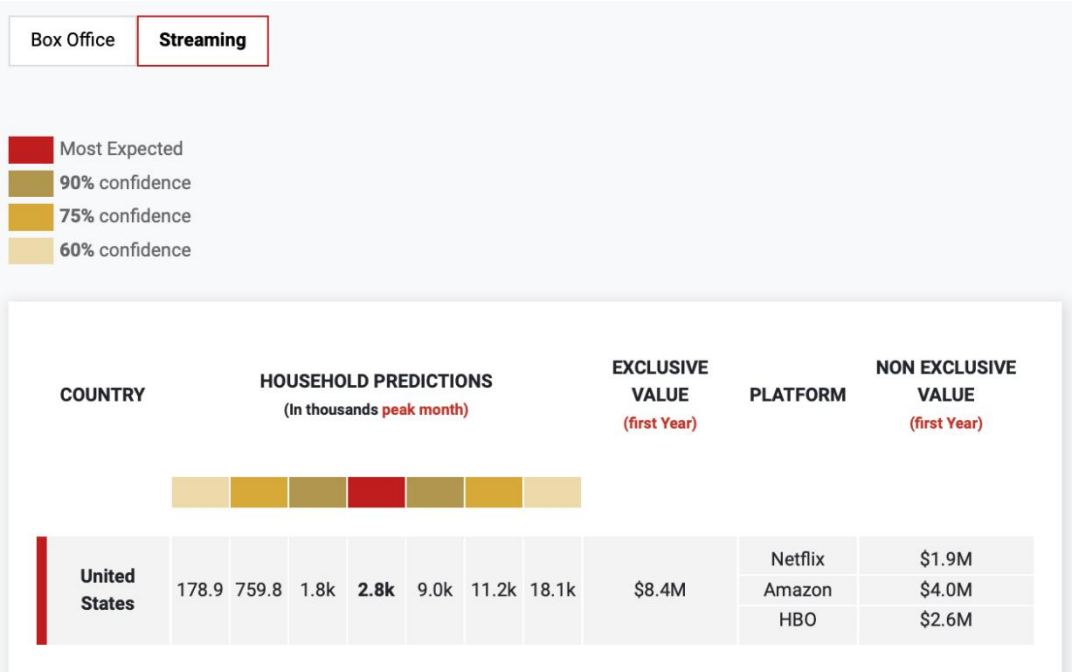
Ilustracja 6: https://fromtheheartproductions.com/largo-ai-artificial-intelligence-film-analysis/

Sztuczna inteligencja na podstawie zaproponowanego scenariusza lub materiału filmowego dopasowuje pod względem gatunku, treści oraz widowni podobne powstałe już dzieła. Do użytku twórców jest również możliwość porównania finansowego własnego projektu z podobnymi filmami pod kątem małego, średniego i wysokiego przychodu. Co najważniejsze, istnieje możliwość sprawdzenia prognoz finansowych dla wszystkich krajów, które interesują producenta w celu ewentualnej dystrybucji. W sekcji analizy mamy do dyspozycji jeszcze kilka istotnych wskaźników, ale w obecnych czasach na pewno warto zwrócić uwagę na oznaczenie oglądalności

filmu przez użytkowników platform streamingach przez okres pierwszego miesiąca od publikacji filmu. Potencjalne podjęcie współpracy z platformą internetową może stanowić bowiem alternatywny kanał dystrybucyjny filmu, który w kanałach tradycyjnych, takich jak kino, telewizja, dostępne nośniki fizyczne, nie cieszyłby się podobną popularnością wśród odbiorców.

6. Strategia dystrybucji

W tej kategorii algorytm zajmuje się obliczeniem najlepszego możliwego terminu na premierę filmu, a co za tym idzie, najwyższej oglądalności podczas weekendu otwarcia, która przekłada się na zyski finansowe. AI oblicza potencjalny dochód z rynków światowych oraz krajowych. Kolejno oddzielną kategorię zajmują platformy streamingowe. Narzędzie z łatwością oblicza możliwą wartość naszego projektu dla takiego producenta oraz kalkuluje potencjalna liczbę odsłon na platformie. AI może zaproponować również potencjalnych dystrybutorów w oparciu o własny projekt producenta. Współpraca producent-dystrybutor może układać się tym owocniej, jeśli dystrybutor ma komplementarny względem danego projektu filmu profil zawodowy, a zatem jego gatunkowe lub rodzajowe preferencje odnośnie dystrybuowanych filmów odzwierciedla złożony projekt.



Ilustracja 7: https://fromtheheartproductions.com/largo-ai-artificial-intelligence-film-analysis/

Według danych firmy *LargoAI*, dostępne narzędzie wspiera projekty twórców z całego świata, jak również przyczyniło się do sukcesu wielu filmów, które znalazły się na platformach stremingowych czy były nominowane do Oscara. Niewątpliwie świadczy to o efektywności tego narzędzia. Nie istnieją bowiem na rynku obecnie żadne inne szeroko dostępne algorytmy, które nakierowane są na proces całej pre-produkcji oraz dystrybucji filmowej. Mamy więc do czynienia z narzędziem, które powstało tylko i wyłącznie dla wsparcia branży filmowej, a w szczególności procesów produkcyjnych. Być może za jakiś czas stanie się to pewnym standardem w środowisku, a na pewno ci sprytniejsi producenci (i nie tylko) będą próbować usprawniać długotrwałe procesy organizacyjne przez korzystanie z rozwoju technologicznego. Demokratyzacja sztucznej inteligencji pozwala wyrównać szanse między małymi, lokalnymi producentami, a wielkimi potentatami przemysłu filmowego, którzy już od dawna sami korzystają z podobnych udogodnień. Chociaż narzędzie skierowane jest głównie do producentów, mogą z niego skorzystać przecież scenarzyści, reżyserzy, dystrybutorzy czy inni twórcy. Wszelakie prognozy dotyczące efektu końcowego ograniczają potencjalne porażki finansowo-artystyczne. Tworzenie filmu to przecież długoletni owoc pracy wielu osób, dlatego tym ważniejsze jest stawianie nacisku na możliwe zmniejszanie ryzyka i niepowodzenia produkcji. Procesy utrwalania materiału filmowego

uległy ułatwieniu od momentu cyfryzacji, więc najwyższy czas na kolejny krok, czyli ułatwienie procesów produkcyjnych. Coraz większa popularność *Largo* na światowych spotkaniach pitchingowych: Cannes, Berlin, znów potwierdza, jak bardzo użyteczne i przyszłościowe jest to narzędzie dla branży.

Warto pamiętać, że jak wspominają sami założyciele *LargoAi* nie otrzymujemy dobrych i złych wyników analizy. Sami winniśmy decydować, czy dana prognoza rzeczywiście ma odniesienie do rzeczywistości. Ostatecznie, nie powinno się wierzyć tylko obliczeniom komputera. Z takim podejściem wiele wspaniałych i wyjątkowych filmów mogłoby nigdy nie powstać. Film zawsz będzie pewnym ryzkiem, więc liczy się tutaj również intuicja oraz wiara w projekt. Wyliczenia maszynowe traktować należy jako wsparcie pewnych procesów, pretekst do poprawy, ale nigdy jako jedyny czynnik w pełni decydujący o powstaniu filmu.

VII. OD TEKSTU PO OBRAZ, CZYLI JAK WSPOMÓC PROCESY PRE-PRODUKCYJNE Z POMOCĄ SZTUCZ-NEJ INTELIGENCJI?

Generowanie obrazów nie stanowi dla AI obecnie najmniejszego problemu. Na podstawie wprowadzonej frazy lub polecenia użytkownicy z całego świata mogą w kilka sekund uzyskać niesamowite grafiki w wymarzonym stylu. Korzystanie z takich usług może być płatne lub ograniczone konkretną liczbą darmowych prób. Obserwujemy zatrzęsienie niesamowitych fuzji stylów czy dzieł kultury w grafikach generowanych za pomocą algorytmów. To, co wcześniej mogli tworzyć tylko artyści, dostępne jest teraz dla szerokiej publiki, która ma tylko jedno główne ograniczenie- swoją wyobraźnię. Zasadne jest rozważenie skorzystania z powyższych możliwości również na potrzeby produkcji filmowej.

Na rynku dostępnych jest wiele narzędzi: DALL·E, MidJourney, Stable Diffusion, DeepAI, Craiyon, Artbreeder, Runway ML, Jasper Art i wiele innych. Ale to właśnie pierwsze trzy zyskały największą popularność i znaczne uznanie na rynku. Wszystkie powyższe generatory działają na tej samej zasadzie – na podstawie

polecania oraz dialogu z użytkownikiem tworzą grafikę lub zdjęcie. Brzmi to oczywiście prosto i idealistycznie, ale do otrzymania upragnionego rezultatu przez twórcę jest jeszcze długa droga. Nie musi być to zasadą, ale najczęściej otrzymanie odpowiedniego obrazu wiążę się z bardzo szczegółowymi opisami, aby algorytm mógł dokładnie wygenerować naszą wizję, co jest wcale niełatwe do uzyskania. Ostatecznie poświęcony czas może okazać się stracony, ponieważ otrzymany wynik może być dość niezadowalający, oczywiście tak być nie musi. Warto zauważyć, że w tym wypadku nic nie zastąpi pracy z drugim człowiekiem, ale niestety nie każdego twórcę stać na takie rozwiązanie szczególnie na początku drogi.

Wykorzystanie AI może być pomocne przy tworzeniu moodboardów, a co za tym idzie – przygotowania materiałów do sesji pitchingowych. Nierzadko artyści mierzą się z problemem odnalezienia zdjęć, które należycie oddadzą klimat planowanego filmu. Głównym źródłem poszukiwań zazwyczaj jest grafika przeglądarkowa albo wykorzystanie ujęć z powstałych już dzieł kinematografii – niestety, podobne działanie wpływa negatywnie na wiarygodność twórców czy też producentów potencjalnego filmu, gdyż prezentowany projekt stwarza wrażenie pozbawionej spójności zbitki wizualnej, w dodatku, niezbyt profesjonalnej w wykonaniu. Co się zaś tyczy generowanych przez AI obrazów, pozwalają one na o wiele bardziej rozwiniętą oryginalność, swoistość wizualnych próbek, które, już na bardzo wczesnym etapie prac nad filmem, mogą świadczyć o jego niepowtarzalnym charakterze.

Tworzenie wizualizacji jest równie znaczące przy tworzeniu storyboardów, które stanowią graficzne ukazanie wizji reżysera. W praktyce produkcyjnej tylko dobrze sfinansowane produkcje mogą pozwolić sobie na luksus opłacenia kolejnego członka w ekipie, jakim jest storyboardzista. Praca takiego człowieka, jest na pewno dłuższa niż generowanie obrazów z AI, ale na pewno bardziej dopracowana. Przez zrozumiałą oszczędność na pierwszych etapach powstawania filmu reżyser i operator sami zajmują się tworzeniem storyboardu korzystając z własnego talentu

plastycznego, albo raczej z jego braku. Dlatego choć sztuczna inteligencja nie stanowi idealnego rozwiązania, stanowi często lepsze rozwiązanie. Twórcy mogą dokładnie zająć się inscenizacją. Bohater powinien siedzieć na fotelu w kropki? Nie ma problemu. Po czasie wizja ulega zmianie i bohater ma siedzieć na drewnianym krześle — dość sprawnie, możemy wygenerować poprawioną wersję. Jak przedstawiono na grafice poniżej, możliwość dokładnego kontrolowania kadru pod kątem ustawienia kamery, aktora czy światła pozwala na wczesnym etapie zniwelować późniejsze problemy na planie. Praca ze storyboardzistą jest na pewno bardziej twórcza, ale najczęściej efektem są tylko czarno - białe ilustracje, a w przypadku pracy z AI możemy uzyskać szybkie efekty w dowolnym graficznym stylu w zaplanowanej kolorystyce.

Ilustracja 8: https://www.cined.com/creating-artistic-mood-boards-for-videos-using-ai-tools/

Tak proste odwzorowanie pomysłu za pomocą ilustracji może na wczesnym etapie zniwelować felerne pomysły, które w wyobraźni twórców prezentowały się lepiej niż na papierze. Warto jeszcze raz podkreślić, że tak szczegółowa inscenizacja stanowi narzędzie do efektywniejszej komunikacji z innymi szefami pionów, nie tylko reżysera z operatorem. Wizualizacje stanowią potencjał do usprawnienia dialogu między reżyserem a np. scenografem, kostiumografem czy kolorystą.

Jeszcze do niedawna sztuczna inteligencja nie potrafiła utrzymać ciągłości wyglądu postaci, co oczywiście negatywnie wpływało na tworzenie pomocowych ilustracji filmowych, gdzie ważne jest jednak wizualne zachowanie spójności bohaterów. Na przestrzeni kilku ostatnich miesięcy wprowadzono aktualizacje dla niektórych generatorów, które rozwiązały ten problem[36]. W programie Midjourney użytkownik określa algorytmowi spójność wyglądu od 0% do 100%, gdzie najwyższa wartość odpowiada najdokładniejszemu odwzorowaniu. Na poniższym rysunku użytkownik programu Midjourney zastosował odwzorowanie wyglądu postaci na poziomie 100%. Na każdej z czterech przygotowanych przez program propozycji widzimy dość staranne odwzorowanie wyglądu oraz ubrania. Każda z czterech ilustracji jest już w tym momencie ze sobą niesamowicie spójna

Wprowadzone zdjęcie:

Ilustracja 9: https://www.youtube.com/watch?v=7BO_TocisAc&t=54s

[36] *Character Reference - Consistent Characters in Midjourney*, https://docs.midjourney.com/docs/character-reference [27 sierpień 2024 r.].

Rezultat wygenerowany przez AI:

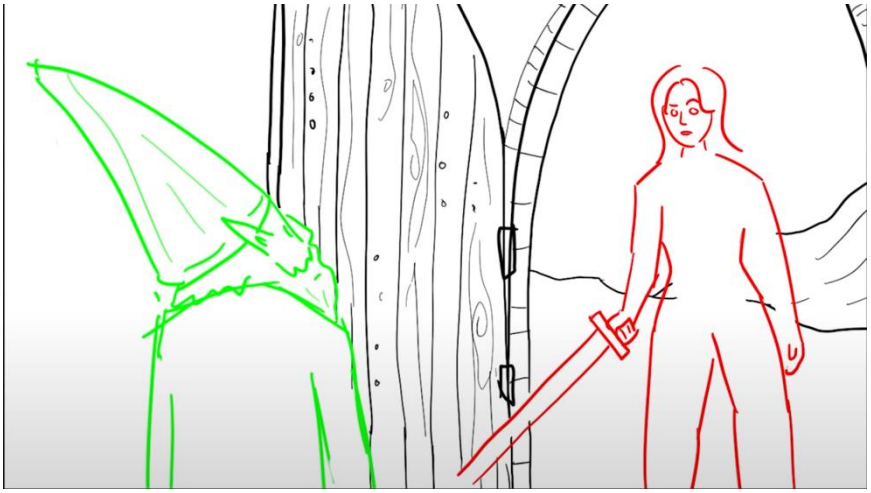
Ilustracja 10: https://www.youtube.com/watch?v=7BO_TocisAc&t=54s

Twórcy filmowi mogliby zatem pójść nawet o krok dalej i wprowadzić do programu zdjęcia aktorów, którzy przeszli proces castingu, aby otrzymać jeszcze dokładniejsze rezultaty, co nigdy wcześniej nie miało miejsca w procesie tworzenia storyboardów.

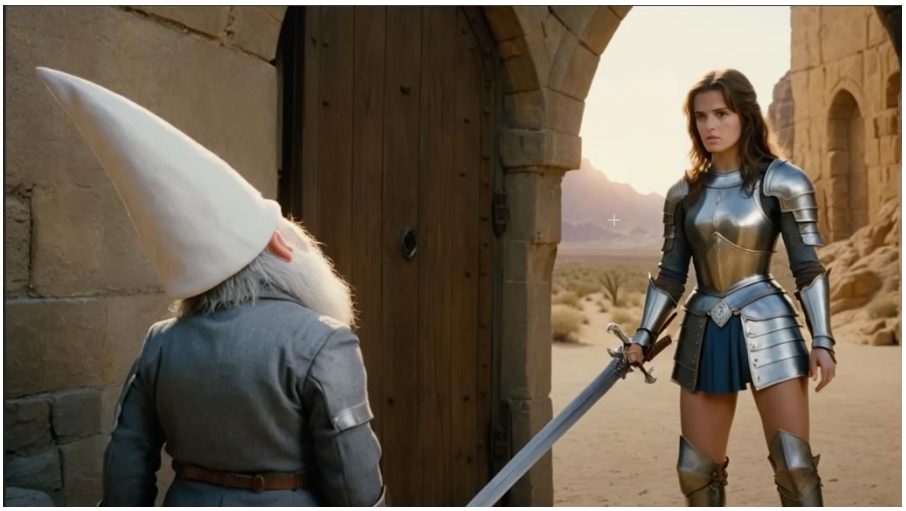
Stable Diffusion w połączeniu z programem *Comfy UI*, który oferuje prosty w obsłudze pulpit użytkownika (interfejs) i wygodny sposób komunikacji ze *Stable Diffusion* na własnym komputerze, może zaoferować filmowcom jeszcze większe możliwości pod względem tworzenia dokładnych wizualizacji pomysłów. Jeśli jakikolwiek twórca filmowy, choć najczęściej będzie to reżyser lub operator, zapragnie na podstawie własnego rysunku wygenerować obraz lub zdjęcie w dowolnym stylu, będzie mógł to osiągnąć. W przypadku tego rozwiązania pozbywamy się elementu losowości, który występuje w standardowej pracy z generatorem obrazu, polegającym na długotrwałym dialogu z algorytmem połączonym z próbą uzyskania odpowiedniego wyniku. Połączenie dwóch powyższych narzędzi umożliwia utrzymanie ciągłości wyglądu postaci, ale w tym wypadku funkcja ta jest jeszcze bardziej zawansowana — mianowicie możliwe jest

zachowanie spójności wizerunku wielu postaci. Każdy bohater, w tym jego szczegółowy wygląd opisywany jest w osobnej rubryce, dzięki której zachowujemy konsekwentny wygląd przy kolejnych generacjach obrazu czy wprowadzaniu innych alteracji. Każda postać winna być namalowana w alternatywnej kolorystyce, aby program potrafił odróżnić poszczególne elementy rysunku. Dlatego jedna postać powinna być naszkicowana na np. niebiesko, druga — na zielono i tak dalej.

Siła fuzji *Stable Diffusion* i *Comfy UI* leży właśnie w dokładnym panowaniu przez użytkowników nad rozmieszczeniem poszczególnych elementów w kadrze i określaniem ich wielkości. Poglądowy szkic umożliwia sprawniejsze uzyskanie pożądanych efektów.



Ilustracja 11: https://www.youtube.com/watch?v=noMeCDIjsWw&t=753s



Ilustracja 12: https://www.youtube.com/watch?v=noMeCDIjsWw&t=753s

Na ilustracjach (wstaw numery ilustracji) łato dostrzec jak algorytm na podstawie szkicu oraz pisemnego opisu postaci, wygenerował precyzyjny obraz. Twarz bohaterki jest daleka od realizmu, ale każdy inny element — otoczenie, scenografia, światło, ubrania — prezentuje się imponująco.

Na dalszym etapie prac nad filmem odpowiednio przygotowane obrazy lub zdjęcia — z pomocą AI lub bez — twórca może poddać procesowi nadania ruchu w kadrze. Możliwe jest zatem stworzenie animatika, ale tak naprawdę obecne narzędzia pozwalają uzyskać pełnoprawne animacje, które od animatika odróżnia właśnie szczegółowość, płynne ruchy postaci i kamery oraz precyzja wizualna. Obecnie istnieje na rynku wiele narzędzi potrafiących ożywić zdjęcie — *Kaiber, Pika Labs, Deep Nostalgia, Pixbim Animate Photos AI* i inne. Co ciekawe niektóre z nich posiadają większą wszechstronność i potrafią tylko na podstawie polecenia tekstowego (bez wprowadzania obrazu) wygenerować plik wideo, np. Pika Labs i Runway. Wielu użytkowników Internetu korzysta przeważnie z metody generowania wideo na podstawie wprowadzone zdjęcia, posiadając dzięki temu większą kontrolę nad końcowym rezultatem. Podobnie jak w przypadku *Stable Diffusion* i *Comfy UI* — przechodzimy od szkicu do zdjęcia, czyli w tym wypadku od

zdjęcia do animacji — od referencji do efektu. Generowanie plików filmowych, szczególnie na podstawie starannych wyselekcjonowanych zdjęć referencyjnych projektu, w przyszłości ma szansę umożliwić twórcom tworzenie zwiastunów filmowych na potrzeby sesji pitchingowych. Obecnie narzędzia nieskutecznie radzą sobie z tworzeniem realistycznych ludzkich postaci, a nawet bardziej z utrzymaniem stałości wyglądu takiej postaci podczas ruchu wewnątrz kadrowego. Przy nieruchomych obrazach uległo to znacznej poprawie, ale efekty powinny być w przyszłości bardziej zadowalające. Generowanie treaserów filmowych z pomocą AI byłoby kolejnym krokiem w ułatwieniu pracy przy przygotowaniu materiałów do sesji pitchingowych. Obecnie twórcy do przedstawiania swojej wizji artystycznej korzystają z fragmentów stworzonych już dzieł. Nie dość, że ogranicza to pewną swobodę artystyczną, w końcu do wyboru są tylko konkretne ujęcia, to dodatkowo narzuca to konkretne skojarzenia i porównania do powstającego projektu, który jest zupełnie nowym dziełem. Należy bowiem pamiętać, że film, nawet będący sumą różnorodnych inspiracji, stanowi pełnoprawną, niepowtarzalną całość pod względem formy i treści, nie jest natomiast prostym, a zarazem ograniczającym przez specyfikę swoich elementów, kolażem wcześniejszych produkcji filmowych, który to kolaż mógłby ponadto łamać prawa autorskie konkretnych twórców. Dość stwierdzić, że te dzieła, które zawierały zbyt wiele zbieżności z istniejącym materiałem – np. film „Za garść dolarów" (1964 r.) w reżyserii Sergio Leone z filmem „Straż przyboczna" (1961 r.) w reżyserii Akiry Kurosawy – stawały się niejednokrotnie przyczyną rozpraw sądowych zakończonych udowodnieniem plagiatu[37].

Ponadto, materiały filmowe generowane przez AI mogą w przyszłości korygować ewentualne niedociągnięcia realizacyjne powstałe w okresie zdjęciowym. Jeśli bowiem podczas montażu filmowego okaże się, że brakuje np. jednego istotnego

[37] Flavia Brizio-Skov, *Examples of National and Transnational Cinema: Akira Kurosawa's Yojimbo and Sergio Leone's A Fistful of Dollars*, „Journal of Cultural and Religious Studies" 2016, t. 4, nr 3.

ujęcia w sekwencji – zamiast organizować, przekładające się na dodatkowe koszty, dokrętki, będzie można wzbogacić zrealizowany na planie materiał o ten wygenerowany.

VIII. Produkcja filmowa a audio-AI

Generowanie obrazów oraz plików filmowych to jedna kwestia, ale żeby film był pełnoprawnym dziełem musi zawierać też warstwę dźwiękową. W tym obszarze również obecnych jest na rynku wiele algorytmów generujących pliki dźwiękowe. Dialog, muzyka, efekty dźwiękowe, postsynchrony – wszystko jest również elementem kinematografii. Twórcy zatem mają wiele możliwości wspomagania produkcji również od strony sonicznej.

Wywołująca odpowiednie emocje u widza gradacja napięć uzyskiwana jest przez odpowiednie prezentowanie akcji fabularnej filmu w sferze wizualnej, jak i dźwiękowej. Można by pokusić się o tezę, że bez odpowiedniej warstwy muzycznej przekaz filmu nie mógłby zostać osiągnięty. Dlatego też warto przyjrzeć się nowym możliwościom dotyczącym generowania muzyki, jakie oferuje AI na przykładzie programów *Suno AI* oraz *Mubert AI*. Zasady funkcjonowania tych dwóch algorytmów opierają się na tworzeniu muzyki na podstawie polecenia tekstowego, ale każdy z programów ma w tej kwestii trochę inne możliwości. *Mubert AI* jest bardziej uproszczonym programem, w zestawianiu z *Suno.* W ciągu kilku sekund algorytm generuje dla nas ścieżkę muzyczną na podstawie polecenia tekstowego lub wyboru odpowiedniego gatunku muzycznego czy nastroju emocjonalnego. Możliwe jest ustawienie długości utworu od 5 sekund do 25 minut. Użytkownik ma możliwość skorzystania z utworów do swoich projektów bezpłatnie, jeśli nie są one przeznaczone do użytku komercyjnego pod warunkiem oznaczenia właściciela praw autorskich, czyli *Mubert AI.* Zatem utwory generowane przez algorytm są wolne od tantiem,

ale są objęte prawami autorskimi. W przypadku wykorzystania komercyjnego utworów wymagane jest skorzystanie z płatnego planu subskrypcyjnego.

W przypadku *Suno AI* mamy możliwość generowania nie tylko muzyki instrumentalnej, ale także tworzenia piosenki z napisanym przez użytkownika bądź AI tekstem. Tworzenie muzyki instrumentalnej polega na wprowadzeniu przez użytkownika opisu tekstowego określające styl, instrumentarium oraz inne elementy dzieła muzycznego tj. rytm i tempo. Niektóre wyniki mogą być zaskakująco dobre, pod względem brzmienia i akuratności w stosunku do polecenia, a inne o wiele mniej zadowalające. Można przy tym zauważyć tendencję, że im rzadszy instrument i skomplikowane polecenie, tym trudniej uzyskać satysfakcjonujący efekt. Jak w przypadku wszystkich modeli AI — nie jest to jeszcze idealne narzędzie, ale już na ten moment wydaje się dość obiecujące. Przykładem możliwości technicznych jest zwłaszcza generowanie prostych piosenek. Użytkownik może napisać swój własny tekst lub jedynie podać tematykę i w kilka sekund otrzymać słowa piosenki. Co do warstwy muzycznej, wszystkie opcje są dostępne jak w przypadku generowania muzyki bez wokali. Wykorzystanie utworów na użytek własny bez monetyzacji jest dozwolony, jeśli zostanie określone, że utwór był tworzony za pomocą *Suno AI* np. w napisach końcowych. Jeśli użytkownik korzysta z płatnej opcji zostaje właścicielem wygenerowanego przez siebie utworu nawet po zakończeniu płacenia subskrypcji.

Tworzenie efektów dźwiękowych oraz odgłosów tła jest również możliwe dla każdego użytkownika. Narzędzia takie jak *AudioLDM, Runway ML (Model Generate Sounds), AI SoundFX, DeepSound AI* potrafią na podstawie wprowadzonego obrazu bądź materiału filmowego automatycznie wygenerować odpowiednie efekty dostosowane do akcji filmu. Znany program montażowy *Davinci Resolve* w swoim panelu do edycji dźwięku — Fairlight — wprowadził narzędzia oparte na sztucznej inteligencji umożliwiające odszumianie plików i izolację dialogów. Jest to możliwe dzięki nowoczesnej analizie spektralnej połączonej z AI, która jest przystosowana do

wychwytywania formantów ludzkiej mowy spośród np. zgiełku zaszumionego miasta oraz płynną izolację dialogu, nawet jeżeli przestrzeń dźwiękowa posiada duży pogłos. Oczywiście nadal są to narzędzia destruktywne dla plików audio i wprowadzają artefakty, które w zależności od jakości samego nagrania będą bardziej lub mnie słyszalne.

W normalnym przebiegu post-produkcji audio w filmowym świecie realizator nagrań synchronicznych zajmował się osobnym nagrywaniem efektów specjalnych lub korzystano np. z płatnych bibliotek. Dzięki użyciu narzędzi AI twórcy niskobudżetowych realizacji lub tych znajdujące się na etapie developmentu, kiedy istotne jest przecież tworzenie materiałów poglądowych, mogą korzystać z wielu narzędzi, które znacznie ograniczą czas oraz koszty.

Synteza mowy ludzkiej przez komputery nie wzbudza aż takiej fascynacji, ponieważ już w 2005 roku powstał polski syntezator mowy „IVONA", stworzony przez studentów Politechniki Gdańskiej na kierunku informatyka. Wynalazek został nagrodzony przez National Geographic w kategorii „Odkrycie roku 2006" oraz przez konkursy branżowe, w których pokonał gigantów rynkowych — Microsoft i Nokię. Było to nic innego jak narzędzie opierające się na działaniu algorytmów generujące na podstawie wprowadzonego pliku tekstowego — plik dźwiękowy naśladujący mowę ludzką[38]. W roku 2011 pojawił się wirtualny asystent głosowy od firmy Apple — Siri, a w 2014 roku firma Amazon zaprezentowała swój odpowiednik, czyli Alexę. Wprowadzenie generatorów mowy zainspirowało wielu twórców, którzy zastanawiali się nad dalszym rozwojem podobnej technologii i jej wpływu na ludzkie życie. Film „Her" z 2013 roku w reżyserii Spike'a Jonze'a poruszył chociażby temat budowania relacji w świecie, w którym sztuczna inteligencja w postaci asystenta głosowego jest normą. Po dziesięciu latach od premiery filmu podejmowano

[38] M. Kaszczuk, L. Osowski, *The IVO Software Blizzard 2007 Entry: Improving Ivona Speech Synthesis System*, 2007.

konkretne próby oceny świata wykreowanego w nim w oparciu o aktualną rzeczywistość[39].

Powstał np. artykuł badawczy, który na podstawie filmu „Her" analizował od strony psychoanalitycznej relacje między ludźmi a sztuczną inteligencją[40].

Na dzień dzisiejszy generatory mowy oparte o AI są bardzo mocno rozwinięte. Program *Murf AI* potrafi generować pliki audio naśladujące mowę ludzką na podstawie wprowadzonego tekstu. Użytkownik może wybrać z banku głosów odpowiadającą barwę głosu lektora, wiek, płeć i język. Używanie programu w trybie bezpłatnym umożliwia przetestowanie możliwości *Murf AI,* ale nie pozwala na pobieranie, a też co za tym idzie — użytek własny. Dopiero za opłaceniem wybranego planu subskrypcyjnego przed użytkownikiem otwiera się dostęp do wielu możliwości w tym: dostęp do ponad 200 barw głosu, tonacji oraz stylów wypowiedzi; nielimitowane pobieranie czy prawa do użytku komercyjnego.

Użytkownik wybiera odpowiadający głos lektorski, przedział wiekowy (co przekłada się też na barwę głosu postaci), a nawet styl wypowiedzi, a następnie wprowadza plik tekstowy. Po wygenerowaniu propozycji przez AI do dyspozycji są narzędzia edycyjne pozwalające na większą kontrolę nad efektem. Po pierwsze, istnieje zawsze możliwość sprawnej zmiany głosu lektora (to samo dotyczy też stylu) oraz zmiany wysokości głosu. Po drugie — dostosowanie szybkości mowy oraz dodanie pauz. Po trzecie, możemy wybrać słowa w tekście na które mają być najbardziej pokreślone. Po wybraniu takich słów można dostosować szybkość wypowiadania tego słowo oraz jego wysokość. Dzięki temu użytkownik posiada

[39] *Column | What 'Her' says about us, 10 years later - The Daily Illini*, https://dailyillini.com/buzz-stories/buzz-movies-and-tv/2023/12/18/column-what-her-says-about-us-10-years-later/ [30 sierpienia 2024 r.].

[40] S. Imm, T.U. Kang, *Intimacy and Love with Artificial Intelligence in the Movie "Her"*, „Psychoanalysis" 2020, t. 31, nr 4.

kontrolę nie tylko nad ogólnym wydźwiękiem wypowiedzi, ale również nad pojedynczymi słowami, co daje możliwość dokładnej kontroli końcowego wyniku.

Można też pójść o krok dalej i skorzystać z funkcji *„say it my way"* (tł. *powiedz to po mojemu*). Polega to na przeczytaniu przez użytkownika części tekstu, a AI na tej podstawie będzie naśladować nasz sposób wymowy, co pozwala na jeszcze większą kontrolę wyniku.

Program *ElevenLabs* w sekcji *VoiceLab* potrafi naśladować barwę głosu na podstawie wprowadzonej próbki, czyli użytkownik może wprowadzić do programu plik z nagraniem np. własnego głosu w dobrej jakości, na którego podstawie AI skopiuje jego głos. Pozostałe kroki pozostają podobne jak w przypadku innych generatorów tekstowych — wprowadzamy plik tekstowy, ale w tym przypadku wybieramy swój głos jako lektora i otrzymujemy przeczytany przez siebie tekst. Program może naśladować każdy głos na podstawie wprowadzonej próbki. Takie narzędzie może być bardzo przydatne w produkcji filmowej np. w przypadku nagłej śmierci aktora. W ostatnim czasie umarł wybitny polski aktor — Jerzy Stuhr, który udzielił głosu postaci Osła w filmie animowanym „Shrek". Na rozmaitych grupach i portalach internetowych dyskutowano, czy do nadchodzącej premiery kolejnej części filmu powinno się wykorzystać AI, aby odtworzyć charakterystyczny głos Jerzego Stuhra, bez którego dla wielu polskich widzów postać Osła nie może istnieć. Spora część osób wypowiadających się uznało pomysł użycia AI za dobry, za to wielu innych komentatorów uważało, że jest to działanie pozbawione szacunku wobec spuścizny istotnego dla polskiej kultury aktora[41].

Dźwięk w produkcji filmowej to bardzo szerokie pojęcie. Coraz częściej w branży, szczególnie w przypadku dość skromnych produkcji filmowych, osoby zajmujące się dźwiękiem muszą opanowywać coraz większą liczbę specjalizacji, aby

[41] K. Chwast, *Kto powinien dać Osłu ze Shreka głos? Ludzie chcą odtworzyć głos Stuhra w AI, ale to zły pomysł*, https://rozrywka.spidersweb.pl/osiol-shrek-jerzy-stuhr-kto-powinien-zagrac [31 sierpnia 2024 r.].

utrzymać się w zawodzie. W przypadku wielkoskalowych i dobrze dofinansowanych produkcji sztab ludzi zajmuje się samą tylko warstwą muzyczną: kompozytor, asystenci kompozytora, orkiestratorzy, dyrygent, muzycy orkiestrowi, realizator nagrania i jego asystenci, inżynier dźwięku odpowiadający za mix; na planie: tyczkarze oraz główny realizator nagrania na planie; w post-produkcji efektowej audio: osoby od post-synchronów, od sond designu, osoby od ponownego nagrywania oraz liczni asystenci. Znowuż łatwo wyobrazić sobie osiągnięcie przez technologie generatywną odpowiednich kompetencji, aby podobnie liczny sztab realizatorów został zastąpiony przez pojedynczy generator – jak w wielu poprzednich przypadkach, pozostaje mieć nadzieję, że nie doprowadzi to w konsekwencji do stanu skrajnego bezrobocia wśród osób zajmujących się zawodowo dźwiękiem oraz muzyką.

IX. CHATGPT JAKO POMOCNE NARZĘDZIE DLA SCENARZYSTÓW

ChatGPT został wyposażony w funkcje, które mogą być znaczące dla twórców przy szeroko pojętym procesie tworzenia scenariusza filmowego. Algorytm znany jest z tworzenia streszczeń na podstawie podanego tekstu, tak więc w tym wypadku scenariusza filmowego. Co za tym idzie, w łatwy sposób można otrzymać synopsis oraz logline. Aby przelać swój pomysł w format scenariusza wystarczy mieć chociażby drobny zarys, oczywiście im bardziej szczegółowo opisany i rozbudowany pomysł — tym lepiej, ale przeniesienie dowolnego tekstu narracyjnego w format scenariusza nie stanowi problemu. Algorytm stworzy przy tym np. dialogi oraz didaskalia zachowując przy tym odpowiedni format. Może stanowić to świetną bazę do rozpoczęcia pracy nad formującym się w głowie twórcy pomysłu na film, który trzeba przelać w format scenariusza. Jeśli scenarzysta bądź reżyser posiadają gotowy scenariusz, mogą wykorzystać AI do poprawy edycji, ulepszenia dialogów bądź stworzenia dialogów oraz formatowania pliku. Tak naprawdę z minimalną ilością

danych takich jak: krótki zarys fabuły, opis bohaterów, opis relacji między bohaterami, opisu sceny, krótkiego przebiegu akcji — twórcy mogą zacząć pracę nad scenariuszem. Co więcej, *ChatGPT* może być po prostu traktowany jako script doctor, z którym pracujemy nad tekstem. Wszystkie zadania script doctora, a mianowicie analiza i poprawa struktury scenariusza, poprawa dialogów, udoskonalanie postaci, praca nad tempem filmu, praca nad problemami narracyjnymi — wszystko to potrafi również produkt od firmy OpenAI. Może to stanowić doskonałe narzędzie w przypadku niskobudżetowych produkcji filmowych, których nie stać na usługi script doctora. Wszelkie niskobudżetowe produkcje, a w szczególności prace początkujących filmowców mogą zyskać znacząco na jakości poprzez pracę z AI.

Na poniższej ilustracji widnieje fragment wygenerowanej przez algorytm sceny, która powstała na podstawie następującego polecenia tekstowego: *zakochani w parku siedzą i obserwują chmury.* Ponieważ zdanie to było bardzo krótkie i bardzo ogólnikowe, algorytm miał dużą dowolność dotyczącą przebiegu fabuły, ale warto docenić fakt, że na podstawie tak lakonicznego polecenia możliwe jest otrzymanie prostej sceny, która stanowi dobry początek do rozpoczęcia pracy ze scenariuszem.

INT. PARK - DAY
Anna i Tomasz siedzą na trawie w parku, trzymając się za ręce i patrząc w niebo. W oddali słychać dzieci bawiące się na placu zabaw.

ANNA
(spoglądając w niebo)
Patrz, ta chmura wygląda jak serce.

TOMASZ
(uśmiecha się)
A ta obok jakby do nas machała.

ANNA
(z lekkim śmiechem)
To chyba znak.

TOMASZ
(spoglądając na Annę)
Znak, że jesteśmy we właściwym miejscu.

ANNA
(cicho, z uczuciem)
I z właściwą osobą.

Ilustracja 13: Ilustracja własna stworzona przy użyciu ChatGPT

To, co niemniej udowadniają rozliczne badania przeprowadzane na modelach językowych, a co stanowi również jedną z tez niniejszej pracy, generatory treści stanowią w tej chwili przydatne narzędzie wspomagające proces twórczy – w niedalekiej przyszłości natomiast mogą one uchodzić za narzędzia podstawowe oraz bardzo szeroko wykorzystywane[42]. Znamienny jest bowiem fakt, iż generatory na obecnie funkcjonującym poziomie komplikacji nie potrafią stworzyć samoistnie (bez dodatkowej korekcji pochodzącej od człowieka) zadowalające odbiorców (a zatem także ludzi) utwory – przy tym zarówno od strony formalnej, jak i estetyczno-emocjonalnej. W eksperckich omówieniach dotyczących owych zagadnień pojawiają się również bardziej odważne stwierdzenia – dotyczące, mianowicie, tego, iż popularyzacja modeli językowej w branży kreatywnej zredefiniuje samo pojęcie „kreatywności", która nie będzie już uznawana za bezdyskusyjnie ludzką cechę, lecz wynik kolaboracji między człowiekiem-twórcą a narzędziem generatywnym[43].

[42] K.M. Happle, *Cre-AI-tivity. The Intersection of Human Ingenuity and AI Collaboration in Scriptwriting with ChatGPT 4.0*, Utrecht 2024, s. 35.
[43] Ibidem, s. 36.

X. PODSUMOWANIE

Jakiekolwiek, mniej lub bardziej szczegółowe, omówienie zjawiska upowszechniania się technologii generatywnej – m.in. w branży filmowej, ale również w każdej innej – niesie ze sobą ryzyko prędkiej deaktualizacji, z pewnością zaś charakterystyczna jest pewne efemeryczność stwierdzeń oraz poglądów na zjawisko podobnie przełomowe i graniczne, które, bez poczucia nadużycia, można powoli porównywać z rewolucją dźwiękową czy też wprowadzeniem koloru w kinie.

Perspektywa, która została obrana w niniejszej pracy, jest niemniej perspektywą na wskroś futurologiczną, nie opierającą się przy tym na choć efektownym, to nie podpartych jakimikolwiek wyraźnymi przesłankami wizjonerstwie, ale na świadomym spoglądaniu w, być może budzącą wiele uzasadnionych wątpliwości, ale jednak nieuchronność przemian, jakich świadomi są także eksperci oraz sami twórcy – zwłaszcza ci, którzy zapisali się już w historii jako utrwalający większe i mniejsze rewolucje technologiczne lub strukturalno-organizacyjne. Taką osobą bez wątpienia był w drugiej połowie XX wieku George Lucas, którego najsłynniejsze dzieło nie tylko uformowało nowego rodzaju widowisko filmowe, wówczas gatunek *space opery*, lecz także przyczyniło się do wspomnianych zmian w obszarze technologii oraz organizacji stricte w dziedzinie produkcji filmowej. Co nie powinno wydawać się zatem zaskakujące, ów George Lucas w wywiadzie dla firmy medialnej Brut w trakcie 77. Międzynarodowego Festiwalu Filmowego w Cannes[44] poruszył kwestię wpływu generatorów treści na przyszłość kinematografii ironicznie zauważając, iż stanowisko skrajnie przeciwne użyciu generatorów w twórczości artystycznej jest obecnie tak samo nonsensowne i epigońskie jak w przeszłości obstawanie przy pojazdach zaprzęgowych wobec rewolucji motoryzacyjnej.

[44] Wypowiedź zawarta w materiale opublikowanym na oficjalnym kanale firmy Brut na platformie YouTube: https://www.youtube.com/watch?v=qnTMTbt7NkA [31 sierpnia 2024 r.].

Znamienna uwaga Lucasa jest zarazem odzwierciedleniem jednego z głównych wniosków przedstawionych na przestrzeni owej pracy, a stwierdzającego, iż, niezależnie od stale powiększającej się liczby kontrowersji, generatory niosą ze sobą kolejną niepowstrzymaną rewolucję dla branży kreatywnej. Oczywiście, nie będzie przesadą stwierdzenie, iż każdego rodzaju rewolucja musi ponieść pewne ofiary, ale w rzeczonym wypadku negatywny wpływ przemian może nie ograniczyć się wyłącznie, jak zdarzyło się w przypadku upowszechnienia filmu dźwiękowego, do zmierzchu gwiazd kina niemego, ale może w niedługim czasie pozbawić pracy wielu twórców. Niestety, to spostrzeżenie także nie należy do irracjonalnej złowróżbności.

Nieznaczną nadzieją napawają, jednakże reakcje odbiorców na wczesne efekty nadchodzących przemian, które wtórują głosom niechętnych twórców. Dość często spotykaną recepcję ilustruje m.in. spontaniczna reakcja internetowej społeczności, która doprowadziło do odwołania londyńskiej premiery filmu ze scenariuszem jawnie określanym jako „dzieło" ChatuGPT 4.0[45]. Równie negatywnego odzewu doczekali się producenci filmu dokumentalnego dla platformy Netflix „Dirty Pop: The Boyband Scam", którzy używali generatora do odtworzenia m.in. głosu i wizerunku Lou Pearlmana – bohatera dokumentu[46]. Choć wydarzenia takie są symptomatyczne, naiwnością byłoby przypuszczać, że oddolne działanie widowni bądź dowolnego stowarzyszenia twórców mogą znacząco spowolnić zjawisko wspierane przez największe konglomeraty technologiczne i medialne obecnie funkcjonujące w przemyśle kultury.

[45] A. Pulver: *London premiere of movie with AI-generated script cancelled after backlash:* https://www.theguardian.com/film/article/2024/jun/20/premiere-movie-ai-generated-script-cancelled-backlash-the-last-screenwriter-prince-charles-cinema [31 sierpnia 2024 r.].

[46] A. Nugent: *Netflix users deplore 'embarrassing' feature in No 1 show: 'Really hope this doesn't become a trend':* https://www.independent.co.uk/arts-entertainment/tv/news/netflix-dirty-pop-ai-pearlman-nsync-b2587419.html [31 sierpnia 2024 r.].

Ted Sarandos, aktualny CEO wspomnianej platformy Netflix, zaznacza w tym kontekście[47], iż generatory oparte na modelach językowych, o ile nie zastąpią ludzi pracujących na danych stanowiskach w branży kreatywnej, o tyle staną się narzędziem wymaganym – a zatem takim, który powinien być opanowany przez aspirującego pracownika sektora kultury, ponieważ użycie owego narzędzia, podążając za brutalną argumentacją zwiększonej wydajności przy niższych nakładach finansowych, będzie wkrótce oficjalnie preferowane przez najbardziej wpływowych przedstawicieli branży. Niezależnie jednak od faktycznych konsekwencji opisywanych trendów, przemysł kultury ponownie staje u progu nieodwracalnego przełomu – pod tym jednym względem trudno mieć najmniejsze wątpliwości.

[47] T. Spangler: *Netflix's Ted Sarandos: Generative AI Tools Will Be a 'Great Way for Creators to Tell Better Stories'*: https://variety.com/2024/digital/news/netflix-ted-sarandos-generative-ai-tools-creators-1236077524/ [31 sierpnia 2024 r.].

Literatura

1. 404 Not Found, https://the-eye.eu/public/AI/pile_preliminary_components/books3.tar.gz [2 września 2024 r.].
2. Anti-Piracy Group Takes Prominent AI Training Dataset' „Books3" Offline * TorrentFreak, https://torrentfreak.com/anti-piracy-group-takes-prominent-ai-training-dataset-books3-offline-230816/ [30 lipca 2024 r.].
3. Bałos I., Trenowanie wybranych modeli sztucznej inteligencji a uprawnienia twórców [w:] Studia Prawnicze. Rozprawy i Materiały 2019, nr 2 (25),
4. Kraków 2019.
5. Białynicki-Birula I., Białynicka-Birula I., Komputer się uczy: sieci neuronowe [w:] Modelowanie rzeczywistości, Warszawa 2002.
6. Character Reference - Consistent Characters in Midjourney, https://docs.midjourney.com/docs/character-reference [27 sierpnia 2024 r.].
7. Chwast K., Kto powinien dać Osłu ze Shreka głos? Ludzie chcą odtworzyć głos Stuhra w AI, ale to zły pomysł, https://rozrywka.spidersweb.pl/osiol-shrek-jerzy-stuhr-kto-powinien-zagrac [31 sierpnia 2024 r.].
8. Column | What 'Her' says about us, 10 years later - The Daily Illini, https://dailyillini.com/buzz-stories/buzz-movies-and-tv/2023/12/18/column-what-her-says-about-us-10-years-later/ [30 sierpnia 2024 r.].
9. Complaint – #1 in Andersen v. Stability AI Ltd. (N.D. Cal., 3:23-cv-00201) – CourtListener.com, https://www.courtlistener.com/docket/66732129/1/andersen-v-stability-ai-ltd/ [2 września 2024 r.].
10. Complaint – #1 in Kadrey v. Meta Platforms, Inc. (N.D. Cal., 3:23-cv-03417) – CourtListener.com, https://www.courtlistener.com/docket/67569326/1/kadrey-v-meta-platforms-inc/ [2 września 2024 r.].
11. Complaint – #1 in UMG Recordings, Inc. v. Suno, Inc. (D. Mass., 1:24-cv-11611) – CourtListener.com, https://www.courtlistener.com/docket/68878608/1/umg-recordings-inc-v-suno-inc/ [2 września 2024 r.].
12. Drabarczyk P., Łazarz B., Sztuka i sztuczna inteligencja: o potrzebie zerkania zezem [w:] Teksty Drugie 2021/4, Warszawa 2021.
13. Flavia Brizio-Skov, Examples of National and Transnational Cinema: Akira Kurosa-wa's Yojimbo and Sergio Leone's A Fistful of Dollars, „Journal of Cultural and Reli-gious Studies" 2016, t. 4, nr 3.
14. Gao L. i in., The Pile: An 800GB Dataset of Diverse Text for Language Modeling, 31 grudnia 2020 r., arXiv:2101.00027 [cs].
15. ImageNet Data, https://www.image-net.org/download.php [28 lipca 2024 r.].
16. Imm S., Kang T.U., Intimacy and Love with Artificial Intel-ligence in the Movie "Her", „Psychoanalysis" 2020, t. 31, nr 4.

17. Karras T., Aila T., Laine S., Lehtinen J., Progressive Growing Of Gans For Improved Quality, Stability, And Variation, https://openreview.net/pdf?id=Hk99zCeAb [4 sierpnia 2024 r.].
18. Kaszczuk M., Osowski L., The IVO Software Blizzard 2007 En-try: Improving Ivona Speech Synthesis System, 2007.
19. Li S., Sung Y., INCO-GAN: Variable-Length Music Generation Method Based on Inception Model-Based Conditional GAN, https://doi.org/10.3390/math9040387 [4 sierpnia 2024r.].
20. Lubelski T., Sowińska I., Syska R., Przełom dźwiękowy [w:] Historia kina, tom 2. Kino klasyczne, Kraków 2012.
21. Mikolov T., Chen K., Corrado G., Dean J., Efficient Estima-tion of Word Representa-tions in Vector Space, http://arxiv.org/abs/1301.3781 [3 sierpnia 2024 r.].
22. Milmo D., Thom Yorke and Julianne Moore join thousands of creatives in AI warning, https://www.theguardian.com/film/2024/oct/22/thom-yorke-and-julianne-moore-join-thousands-of-creatives-in-ai-warning [22 października 2024 r.].
23. Order on Stipulation – #107 in Kadrey v. Meta Platforms, Inc. (N.D. Cal., 3:23-cv-03417) – CourtListener.com, https://www.courtlistener.com/docket/67569326/107/kadrey-v-meta-platforms-inc/ [2 wrzesnia 2024 r.].
24. Our approach to AI safety, https://openai.com/index/our-approach-to-ai-safety/ [28 lipca 2024 r.]. z Oversight of A.I.: Rules for Artificial Intelligence | United States Senate Committee on the Judiciary, https://www.judiciary.senate.gov/committee-activity/hearings/oversight-of-ai-rules-for-artificial-intelligence [28 lipca 2024 r.].
25. Ouyang X., Chen Y., Zhu K., Agam G., Image restoration refinement with Uformer GAN, https://openaccess.thecvf.com/content/CVPR2024W/NTIRE/papers/Ouyang_Image_Restoration_Refinement_with_Uformer_GAN_CVPRW_2024_paper.pdf [4 sierpnia 2024 r.].
26. Patel D., The Inference Cost of Search Disruption – Large Language Model Cost Analysis, https://www.semianalysis.com/p/the-inference-cost-of-search-disruption [28 lipca 2024 r.].
27. Przegalińska A., Oksanowicz P., Sztuczna inteligencja. Nieludzka, arcyludzka, Kraków 2023.
28. Reisner A., The Authors Whose Pirated Books Are Powering Generative AI, https://www.theatlantic.com/technology/archive/2023/08/books3-ai-meta-llama-pirated-books/675063/ [30 lipca 2024 r.].
29. Sikarwar A., Pawar A., Sethiya A., Surana A., Rajput S., CLIP Model Image Search: A Deep Learning Approach for Finding Relevant Images, https://www.researchgate.net/publication/370300801_CLIP_Model_Image_Search_A_Deep_Learning_Approach_for_Finding_Relevant_Images [15 sierpnia 2024r.].
30. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A., Kaiser L., Polosukhin I., Attention Is All You Need, http://arxiv.org/abs/1706.03762 [3 września 2024 r.].

31. Victoria, Rights Alliance removes the illegal Books3 dataset used to train artificial intelligence | RettighedsAlliancen, https://rettighedsalliancen.com/rights-alliance-removes-the-illegal-books3-dataset-used-to-train-artificial-intelligence/ [30 lipca 2024 r.].
32. Zhang L., Chen X., Wang Y., Lu Y., Qiao Y., Brush Your Text: Synthesize Any Scene Text on Images via Diffusion Model, https://doi.org/10.1609/aaai.v38i7.28550 [30 lipca 2024r.].

Spis ilustracji

1. Strona 7: https://arxiv.org/abs/1812.04948
2. Strona 8: https://developer.nvidia.com/blog/improving-diffusion-models-as-an-alternative-to-gans-part-1/
3. Strona 15: https://arxiv.org/pdf/2101.00027
4. Strona 24: https://fromtheheartproductions.com/largo-ai-artificial-intelligence-film-analysis
5. Strona 25: https://fromtheheartproductions.com/largo-ai-artificial-intelligence-film-analysis
6. Strona 25: https://fromtheheartproductions.com/largo-ai-artificial-intelligence-film-analysis
7. Strona 26: https://fromtheheartproductions.com/largo-ai-artificial-intelligence-film-analysis
8. Strona 27: https://fromtheheartproductions.com/largo-ai-artificial-intelligence-film-analysis
9. Strona 28: https://fromtheheartproductions.com/largo-ai-artificial-intelligence-film-analysis
10. Strona 31: https://www.cined.com/creating-artistic-mood-boards-for-videos-using-ai-tools/
11. Strona 32: https://www.youtube.com/watch?v=7BO_TocisAc&t=54s
12. Strona 33: https://www.youtube.com/watch?v=7BO_TocisAc&t=54s
13. Strona 34: https://www.youtube.com/watch?v=noMeCDIjsWw&t=753s
14. Strona 34: https://www.youtube.com/watch?v=noMeCDIjsWw&t=753s
15. Strona 41: Ilustracja własna stworzona przy użyciu ChatGPT