

Poza fikcjonalnością. Symulacja i narracja w generowanych przez AI opowiadaniach o polskich wykładowcach akademickich

DOI: 10.14746/fp.2026.43.3

Artykuł udostępniono w otwartym dostępie
na warunkach licencji CC BY-NC-ND 4.0.

Bartosz Lutostański

ORCID: 0000-0001-6345-3275

1. Wprowadzenie

Pojawienie się dużych modeli językowych (LLM), zdolnych do generowania stylistycznie spójnych historii o wyraźnym potencjale afektywnym, stanowi ważne wyzwanie dla teorii literatury. Niniejszy artykuł, stanowiący wkład do numeru tematycznego „Forum Poetyki”, dowodzi, że opowiadania tworzone przez AI wymagają pogłębionej refleksji nad pojęciem fikcjonalności. Kanoniczne ramy teoretyczne, szczególnie te oparte na intencji komunikacyjnej i ludzkiej sprawczości semiotycznej, okazują się niewystarczające w odniesieniu do tekstów produkowanych przez algorytmy generatywne. Wytwory te naśladują powierzchowną logikę narracji, lecz pozbawione są intencjonalnego podmiotu, komunikacyjnego etosu oraz relewantnego modelu odbioru.

Systemy generatywne takie jak Claude, ChatGPT, Gemini w szybkim tempie zostały zaadaptowane do tworzenia dzieł literackich: krótkich opowiadań¹, powieści², wierszy³ czy projektów multimedialnych⁴. Jednak pomimo rosnącego zaangażowania badawczego i instytucjonalnego (Hannes Bajohr, Nick Montfort, Jill Walker Rettberg, projekt AI Stories w Bergen czy działania Centrum Humanistyki Cyfrowej przy Instytucie Badań Literackich PAN) ich status ontologiczny pozostaje niedostatecznie opisany i oscyluje między imitacją a inwencją, między algorytmem a sztuką.

W tym kontekście teoria fikcjonalności – przede wszystkim w jej narratologiczno-retorycznym ujęciu – dochodzi do granic swej użyteczności analitycznej. Co prawda pozostaje elastyczna wobec różnych gatunków i rodzajów literackich, nie opisuje jednak adekwatnie produkcji tekstu algorytmicznego. Fikcjonalność – jako kategoria zakorzeniona w intencji nadawcy, sygnalizacji retorycznej oraz genologicznie uwarunkowanych inferencjach odbiorcy – nie obejmuje materiału, który opiera się właśnie tym współrzędnym. Potrzebne jest zatem nie dostosowanie definicyjne, lecz przesunięcie paradygmatyczne. Artykuł proponuje takie przesunięcie poprzez wykorzystanie teorii symulacji Jeana Baudrillarda. Jego dekonstrukcja relacji znaczące–znaczone oraz diagnoza hiperrzeczywistości zdają się oferować bardziej odpowiednie narzędzia analityczne.

Dwa paradygmaty teoretyczne organizują niniejszą pracę: retoryczna teoria fikcjonalności oraz baudrillardowska teoria symulacji. Pierwsza została zaproponowana jako korekta podejść semantycznych, pragmatycznych i fenomenologicznych, ujmując fikcjonalność jako akt dyskursywny oparty na intencji komunikacyjnej, negocjacji interpretacyjnej, ludzkiej wyobraźni oraz świadomości rozróżnienia między domenami fikcyjnymi i niefikcyjnymi. Teoria symulacji z kolei opisuje postsemiotyczne pęknięcie: implozję reprezentacji, pochod symulaków i produkcję hiperrzeczywistości. Może okazać się bardziej przydatna w odniesieniu do narracji generowanych przez AI, które nie posiadają autora w sensie tradycyjnym, lecz wyłaniają się z algorytmicznej rekombinacji w obrębie wielkich korpusów danych. Choć oba paradygmaty wydają się nieprzystawalne – pierwszy zakorzeniony w etyce komunikacji, drugi w filozofii postmodernistycznej⁵ – ich zestawienie ożywia główne napięcie analityczne niniejszego artykułu.

Artykuł łączy wywód teoretyczny z analizą korpusu dziesięciu krótkich opowiadań wygenerowanych przez Claude AI. Prompt – „Produce 10 different versions of a story about a Polish university teacher” – został zaprojektowany tak, by wymusić wariantywność w obrębie ograniczeń tematycznych i stylistycznych, ujawniając tym samym możliwości symulacyjne modelu. Jedno opowiadanie, *Niechętny rewolucjonista*, zostaje poddane szczegółowej analizie zarówno pod kątem powierzchniowej narracyjności, jak i ontologicznego niedookreślenia.

¹ Np. Li Zilles, *Machine, Unlearning* (Denver: Counterpath Press, 2018).

² Rie Qudan, *Tokyo Sympathy Tower* (New York: Summit Books, 2025).

³ David Jhave Johnston, *ReRites*, 2017–2018, https://glia.ca/rerites/rerites_read_txt.html.

⁴ Scott Rettberg, *Republicans in Love*, 2022. Digital exhibition, https://eloconference2023exhibitions.wordpress.com/exhibition2_resistance/republicans-in-love/.

⁵ Agnieszka Ziętek, *Jean Baudrillard wobec współczesności*. Polityka, media, społeczeństwo (Kraków: Universitas, 2013), 19–41.

Artykuł przekracza hermeneutykę opartą na tworzeniu znaczeń, proponując hermeneutykę generowania tekstu, w której to centralne stają się wzorce algorytmiczne i statystyczne.

Teza niniejszego tekstu ma charakter heurystyczny i dotyczy problemu, jakiego języka teoretycznego potrzebujemy, by opisać teksty fabularne, które posiadają formalne sygnały fikcjonalności, lecz pozbawione są ontologicznych warunków jej istnienia. Zamiast wtłaczać narracje generowane przez AI w istniejące kategorie ontologiczne, proponuję rozpoznać je jako zjawisko czwartego rzędu: symulowaną narrację operującą według odrębnych protokołów epistemicznych. Uznanie to wymaga podwójnego przesunięcia: po pierwsze, odejścia od teorii fikcjonalności jako wystarczającego narzędzia; po drugie, porzucenia założenia, że narracja obliczeniowa implikuje intencję komunikacyjną, wyobraźnię antropogeniczną czy sygnalizację retoryczną.

Stanowisko to nie jest pozbawione ryzyka. Teoria symulacji może prowadzić do nadinterpretacji, przypisując estetyczną wagę twórcom maszynowym, z kolei ramy humanistyczne mogą nieświadomie zostać nałożone na artefakty posthumanistyczne. W lekturze tych tekstów badacz nieuchronnie animuje je pragnieniem interpretacji, zakładając obecność znaczenia tam, gdzie nie zostało ono zakodowane. Argument rozwijany w niniejszym artykule ma zatem charakter performatywny: interpretuje po to, by przesunąć samą interpretację. Czyta, aby „odczytać na nowo”. Ta aporia nie jest porażką metodologiczną; jest, moim zdaniem, warunkiem teoretyzowania narracji w epoce AI.

2. LLM-y – od narracji do symulacji

Duże modele językowe (LLM-y) to zaawansowane konstrukcje matematyczne zbudowane na architekturze sieci neuronowych i działające na podstawie obliczeń probabilistycznych oraz wielkoskalowego rozpoznawania wzorców. Przekształcają dane językowe w wektory numeryczne i umożliwiają predykcję kontekstową poprzez iteracyjną optymalizację⁶. Szczególne znaczenie w ewolucji tej technologii miało wprowadzenie w 2017 roku zaproponowanej przez grupę inżynierów z Google architektury transformera⁷, która przetwarza całe sekwencje tekstowe równolegle i „waży” relacje między nimi przy użyciu mechanizmu uwagi⁸.

W uproszczeniu: każdy token w zdaniu otrzymuje obliczeniowe markery (*queries* i *keys*), które są następnie porównywane parami i oceniane pod względem relewancji kontekstowej. Predykcja kolejnego tokenu dokonywana jest za pomocą funkcji *softmax*, obliczającej najbardziej statystycznie prawdopodobną kontynuację⁹. Ponieważ generowanie opiera się

⁶ Alec Radford i in., *Language Models Are Unsupervised Multitask Learners* (San Francisco: OpenAI, 2019), https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.

⁷ Ashish Vaswani i in., „Attention Is All You Need”, *Advances in Neural Information Processing Systems* 30 (2017), <https://arxiv.org/abs/1706.03762>.

⁸ Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio, „Neural Machine Translation by Jointly Learning to Align and Translate”, prezentacja wygłoszona podczas International Conference on Learning Representations (ICLR), 2015, <https://arxiv.org/abs/1409.0473>.

⁹ Jason Wei i in., „Emergent Abilities of Large Language Models”, *Transactions on Machine Learning Research*, 2022, <https://arxiv.org/abs/2206.07682>.

na mechanizmie prawdopodobieństwa, a nie na regułach deterministycznych, identyczne prompty mogą prowadzić do nieskończonej liczby wariantów. Wynik zależy od bardzo wielu parametrów, na przykład od poziomu losowości (*temperature*), strategii próbkowania (*sampling*) lub konkretnego punktu kontrolnego modelu (*model checkpoint*). Czynniki te sprawiają, że teksty generowane przez LLM-y pozbawione są stabilności i powtarzalności, a każda odpowiedź ma charakter przygodny, efemeryczny i zasadniczo niepowtarzalny.

Architektura obliczeniowa dużych modeli językowych skłoniła badaczy do krytycznej refleksji nad ich ograniczeniami. Emily M. Bender i współautorzy określili je mianem „stochastycznych papug” – systemów, które „zszywają sekwencje form językowych [...] bez odniesienia do znaczenia”¹⁰. Zdaniem autorów wygenerowane przez duże modele językowe teksty nie posiadają intencji komunikacyjnej, nie rozumieją świata ani nie biorą pod uwagi odbiorcy, a zamiast tego produkują wypowiedzi, które jedynie imitują spójność językową¹¹. Innymi słowy, ChatGPT czy Claude nie „myślą” ani nie „rozumieją”; kopiują właściwości języka tak skutecznie, że powstaje iluzja inteligencji. Tym samym reprezentują nie komunikację, poznanie czy intencję, lecz skalowalną inferencję algorytmiczną.

Krytyka możliwości LLM-ów przypomina opinię pionierki badań nad technologią komputerową z XIX wieku, Ady Lovelace, że obliczenia mechaniczne nie mogą inicjować znaczenia, lecz jedynie przetwarzać wzorce: „mogą podążać za analizą, lecz nie posiadają zdolności przewidywania relacji czy prawd analitycznych”¹². Mark Halpern mówi w tym kontekście o „nieporozumieniu” polegającym na utożsamianiu składni z semantyką, a Bender i współautorzy o „złudzeniu” inteligencji. Simone Natale dodaje, że oszukujemy siebie, wierząc w sprawczość AI¹³. LLM-y przewyższają ludzi w niektórych zadaniach, lecz pozostają pozbawione świadomości, autonomii i intencjonalności. Argument Johna Searle’a o „chińskim pokoju” pozostaje aktualny: manipulacja symbolami nie jest rozumieniem¹⁴. Jak ujmuje to Angus Fletcher, „żaden komputerowy system AI nigdy nie nauczył się pisać opowiadań [...]; wyemitował syntaktyczną kombinację, którą ludzkie umysły zreinterpretowały jako opowieść”¹⁵.

2.1. Krótka genealogia teorii fikcjonalności

David Gorman wyróżnia trzy główne teorie fikcjonalności. Pierwsza, semantyczna, utożsamia fikcjonalność z treścią i opisuje kwestie prawdziwości czy statusu bytów wyobrażonych. Teoria pragmatyczna traktuje ją jako funkcję praktyk społecznych i komunikacyjnych, natomiast

¹⁰Emily M. Bender i in., *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?*, prezentacja wygłoszona podczas FAccT '21 (New York: Association for Computing Machinery, 2021), 616, <https://doi.org/10.1145/3442188.3445922>.

¹¹Bender i in., 617.

¹²Cyt. za: Mark Halpern, „No Ghost in the Machine: Artificial Intelligence Isn't as Intelligent as You Think”, *The American Scholar*, 2.03.2020, <https://theamericanscholar.org/no-ghost-in-the-machine/>.

¹³Simone Natale, *Deceitful Media: Artificial Intelligence and Social Life after the Turing Test* (Oxford: Oxford University Press, 2021).

¹⁴Por. Alan Turing, „Computing Machinery and Intelligence”, *Mind* 59 (1950): 433–460.

¹⁵Angus Fletcher, „Why Computer AI Will Never Do What We Imagine It Can”, *Narrative 1* (2022): 126.

trzecia zajmuje się formalnymi sygnałami fikcjonalności: mową pozornie zależną, narracją wszechwiedzącą czy markerami paratekstowymi¹⁶. Gorman zauważa jednak istotną asymetrię: teoria fikcjonalności tradycyjnie pomija w swej refleksji „naturę dyskursu niefikcjonalnego”¹⁷. Chcąc wypełnić tę lukę, Henrik Skov Nielsen, James Phelan i Richard Walsh zaproponowali retoryczną teorię fikcjonalności w numerze „Narrative” z 2015 roku. Stanowi ona syntezę wcześniejszych prac badaczy¹⁸ i opiera się na kilku kluczowych założeniach¹⁹. Według Nielsena, Phelana i Walsh’a fikcjonalność zakorzeniona jest w specyficznie ludzkiej umiejętności wyobrażania alternatywnych rzeczywistości, zakłada intencję nadawcy sygnalizującego konkretny status wypowiedzi, determinując jego etos nadawcy oraz logos komunikatu, i wymaga interpretacyjnego założenia przez odbiorcę, że fikcyjność stanowi efekt zamierzonej decyzji nadawcy. Dodają również, że żadna technika formalna nie jest ani konieczna, ani wystarczająca do identyfikacji fikcjonalności. Innymi słowy, fikcjonalność wyłania się w konkretnym kontekście komunikacyjnym i nie może zostać zredukowana do cech formalnych. Zakłada raczej, po pierwsze, że istnieje nadawca zdolny do intencjonalnego sygnalizowania fikcyjności, czyli „postawy wobec przekazywanej informacji odmiennej od postawy wobec dyskursu niefikcyjnego”²⁰. Kompetencje nadawcy opierają się na zdolności „rozumienia rzeczywistości w kategoriach tego, co możliwe i wyobrażone” (to kompetencja, ich zdaniem, „typowo ludzka”) oraz na intencji komunikowania się w sposób fikcyjny, niefikcyjnie bądź tak, by w efekcie zacierać granicę między tymi pozycjami²¹. Ponadto model ten zakłada swoistą umowę między nadawcą a odbiorcą, w której fikcjonalność wywnioskowywana jest z sygnałów kontekstowych i paratekstowych – takich jak etykiety gatunkowe, markery metanarracyjne czy ustalone ramy dyskursu – a także z założenia odbiorcy dotyczącego woli twórczej nadawcy. Umowa ta zakłada intencjonalne komunikowanie tekstu fabularnego, który jest wyobrażony i posiada charakter jednostkowej, twórczej inwencji. Intencjonalna komunikacja między jednym człowiekiem a drugim, wraz z elastycznymi środkami tego procesu obejmującymi różnorodność gatunków, konwencji i stylów, ma znaczenie fundamentalne, ponieważ konstytuuje określone założenia interpretacyjne po stronie odbiorców oraz kształtuje ich stosunek do etosu i logosu tekstu.

Teoria fikcjonalności w ujęciu zaproponowanym przez Nielsena, Phelana i Walsh’a ulega załamaniu w przypadku narracji generowanych przy użyciu narzędzi sztucznej inteligencji, takich jak duże modele językowe. Z powodu sposobu działania narracja wytwarzana przez ChatGPT czy Claude’a nie jest ani wytworem intencjonalnej kreatywności ludzkiej, ani nie jest umocowana semantycznie i referencyjnie. W tym kontekście rola nadawcy zostaje radykalnie

¹⁶David Gorman, „Theories of Fiction”, w: *The Routledge Encyclopaedia of Narrative Theory*, red. David Herman, Manfred Jahn, Marie-Laure Ryan (Oxon: Routledge, 2005), 164–167.

¹⁷Gorman, 167.

¹⁸James Phelan, *Living to Tell about It: A Rhetoric and Ethics of Character Narration* (Ithaca: Cornell University Press, 2005); Richard Walsh, *The Rhetoric of Fictionality: Narrative Theory and the Idea of Fiction* (Columbus: Ohio State University Press, 2007); Henrik Skov Nielsen, „Natural Authors, Unnatural Narration”, w: *Postclassical Narratology: Approaches and Analyses*, red. Jan Alber, Monika Fludernik (Columbus: Ohio State University Press, 2010), 285.

¹⁹W niniejszej pracy pomijam wiele krytycznych głosów retorycznej teorii fikcjonalności. Zob. np. Paul Dawson, „Ten Theses Against Fictionality”, *Narrative* 1 (2015): 74–100.

²⁰Henrik Skov Nielsen, James Phelan, Richard Walsh, „Ten Theses About Fictionality”, *Narrative* 1 (2015): 67.

²¹Nielsen, Phelan, Walsh, 64.

zdestabilizowana, a wręcz podważona – kto jest nadawcą²²? Według narratologów jest nim człowiek intencjonalnie produkujący tekst i świadomie sygnalizujący jego status ontologiczny (np. poprzez rozróżnianie i wyobrażanie scenariuszy kontrfaktycznych czy nierzeczywistych). W przypadku tekstu generowanego przez AI żadna z tych właściwości nie ma zastosowania. Mamy do czynienia z maszyną sterowaną algorytmicznie, probabilistycznym systemem optymalizującym predykcję kolejnego tokenu, którego konfiguracja techniczna i właściwości stochastyczne prowadzą do załamania kontraktu komunikacyjnego, jaki zakłada teoria fikcjonalności. Agent AI, oparty na wielkoskalowym równoległym przetwarzaniu tekstów i matematycznych transformacjach języka w wektory, nie przejawia ani intencjonalności komunikacyjnej, ani zdolności do odróżniania realnego od nierzeczywistego. Nawet jeśli ludzki użytkownik inicjuje akt generowania treści poprzez wpisanie promptu, czynność ta nie jest równoznaczna z komunikacją autorską. Prompt nie stanowi aktu mowy skierowanego do odbiorcy z wyobrażeniową intencją. W konsekwencji ontologiczny status dyskursu narracyjnego nie może zostać ustalony ani nawet wywnioskowany na podstawie wygenerowanego rezultatu.

W rezultacie teksty generowane przez LLM-y opierają się na opisie teoretycznym w ramach modelu retorycznego. W moim przekonaniu trafniej można ująć je jako symulacje w sensie baudrillardowskim: artefakty semiotyczne, które wprawiają w obieg znaki fikcjonalności, nie zakotwicząc ich w antropogenicznej, komunikacyjnej sprawczości i intencjonalności. Odtwarzają one pozór opowiadania – formę narracyjną, łuki afektywne, konwencje gatunkowe – nie wytwarzając jednak fikcjonalności jako aktu skierowanego przez człowieka ku człowiekowi i zakorzenionego referencyjnie. LLM-y symulują zdolność ludzkiego nadawcy do wyobrażania tego, co nierzeczywiste, oraz do sygnalizowania tego aktu wyobraźni ludzkiemu odbiorcy, nie posiadając jednak intencjonalności ani ontologicznego fundamentu, jakie takie ramy teoretyczne zakładają. W terminach Baudrillarda teksty te stanowią symulacje czwartego rzędu: probabilistyczne rekombinacje znaków zaczerpniętych z masywnych korpusów, pozbawione referencji i obwodu²³. Nie oznaczają nieobecności rzeczywistości; maskują nieobecność samej fikcjonalności.

2.2. Teoria symulacji Baudrillarda

Termin „symulacja” od dawna zajmuje ambiwalentną pozycję zarówno w dyskursach naukowych, jak i kulturowych²⁴. Wywodzący się z łacińskiego *simulatio*, pierwotnie oznaczał naśladowanie – często o charakterze zwodniczym bądź reprezentacyjnym. W nowożytnej nauce symulacja stała się techniką modelowania wykorzystywaną do reprezentowania lub przewidywania zjawisk rzeczywistych w warunkach eksperymentalnych. Jej funkcja instrumentalna uległa pogłębieniu wraz z rozwojem cybernetyki i teorii systemów, gdzie symulacja zaczęła oznaczać algorytmiczną replikację dynamiki systemów oraz predykcyjne emulowanie zachowań

²²Zob. Hannes Bajohr, „Writing at a Distance: Notes on Authorship and AI”, YouTube, 26.03.2025, 1:29:38, <https://www.youtube.com/watch?v=pP66rSFjPYw>, i jego propozycję teoretyczną dotyczącą autora, czytelnika i tekstu.

²³Jean Baudrillard, *Symulakry i symulacje*, tłum. Sławomir Królak (Warszawa: Sic!, 2005).

²⁴Ewa Ziętek, Jean Baudrillard.

w zamkniętych środowiskach obliczeniowych. Jednak poza tym technicznym rodowodem symulacja jest pojęciem wieloznacznym – szczególnie w obrębie badań nad grami, gdzie jej wymiar epistemologiczny, reprezentacyjny i ludyczny był wielokrotnie teoretyzowany²⁵. To jednak w rozprawach Jeana Baudrillarda pojęcie to ulega radykalnej reinterpretacji.

Baudrillard definiuje symulację jako „sposób generowania – za pomocą modeli – rzeczywistości pozbawionej źródła i realności: hiperrzeczywistości”²⁶, funkcjonujący jako „dominujący schemat aktualnej fazy kontrolowanej przez kod”²⁷. Proces ten przebiega w czterech stadiach: po pierwsze, obraz odzwierciedla głęboką rzeczywistość; po drugie, maskuje ją i zniekształca; po trzecie, maskuje brak głębokiej rzeczywistości; wreszcie – w stadium czwartym – „pozostaje bez związku z jakąkolwiek rzeczywistością: jest czystym symulakrem samego siebie”²⁸. To właśnie na tym ostatnim etapie – w czwartym porządku – operują opowiadania generowane przez LLM-y: nie wskazują one na realne odniesienia ani nie ukrywają ich nieobecności; „jedynie” wytwarzają efekty znaczeniowe.

Baudrillard przeciwstawia symulację reprezentacji, która opiera się na zasadzie, że znak odnosi się do realnego, nawet jeśli poprzez zniekształcenie czy przemieszczenie. Symulacja natomiast „pochłania cały gmach przedstawienia jako symulakr”²⁹. Oznacza ona fundamentalne pęknięcie: „istniejemy w świecie, w którym jest coraz więcej informacji, a coraz mniej sensu”³⁰. W reżimie hiperrzeczywistości znaki nie odnoszą się już do realnych obiektów ani do wyobrażeń, lecz krążą w obrębie operacyjnego łańcucha równoważności, oderwane od ontologicznego zakotwiczenia. „Rzeczywistość wytwarzana jest ze zminiaturyzowanych komórek, matryc i jednostek pamięci, z oprogramowania – i może z ich pomocą być reprodukowana nieskończoną ilość razy” – pisze Baudrillard³¹. Rezultatem jest system, w którym referenty ulegają implozji, a pozostaje kombinatoryczna gra modeli, ich „losowa komutacja”³².

W tym hiperrzeczywistym reżimie symulacja przestaje funkcjonować jako reprezentacja czy nawet imitacja – staje się strategią powstrzymania. Jak stwierdza Baudrillard, symulacja „podważa różnicę pomiędzy tym, co «prawdziwe» i tym, co «fałszywe»; tym, co «rzeczywiste» i tym, co «wyobrażone»”³³. Nie zakrywa realnego; uniemożliwia jego możliwość. Powstrzymanie – kontynuuje – stanowi sprzężenie realnego z jego „operacyjną kopią, metastabilną, deskryptywną, programowalną i niemylną maszyną, która dostarcza wszelkich znaków rzeczywistości oraz wymija wszelkie jej zmienne koleje losu”³⁴.

²⁵Gonzalo Frasca, „Simulation”, w: *The Johns Hopkins Guide to Digital Media*, red. Marie-Laure Ryan, Lori Emerson, Benjamin J. Robertson (Baltimore: Johns Hopkins University Press, 2014), 452–455; Ian Bogost, *Persuasive Games: The Expressive Power of Video Games* (Cambridge: MIT Press, 2007); Espen Aarseth, *Cybertext: Perspectives on Ergodic Literature* (Baltimore: Johns Hopkins University Press, 1997); Bernard P. Zeigler, Herbert Praehofer, Tag Gon Kim, *Theory of Modeling and Simulation* (London: Academic Press, 2000).

²⁶Baudrillard, *Symulakry i symulacje*, 6.

²⁷Jean Baudrillard, *Simulations*, tłum. Phil Beitchman, Paul Foss, Paul Patton (New York: Semiotext[e], 1983), 83.

²⁸Baudrillard, *Symulakry i symulacje*, 12.

²⁹Baudrillard, *Symulakry i symulacje*, 11.

³⁰Baudrillard, *Symulakry i symulacje*, 101.

³¹Baudrillard, *Symulakry i symulacje*, 7.

³²Baudrillard, *Simulations*, 109.

³³Baudrillard, *Symulakry i symulacje*, 8.

³⁴Baudrillard, *Symulakry i symulacje*, 7.

To załamanie referencji w wyniku zastąpienia rzeczywistości znakami rzeczywistości stanowi bezpośrednie wyzwanie dla teorii fikcjonalności, szczególnie w modelu narratologiczno-retorycznym zaproponowanym przez Nielsena, Phelana i Walsha. Opiera się wszak on na ściśle określonych zadaniach nadawcy i odbiorcy, które w hiperrzeczywistości baudrillardowskiej zostają anulowane: nie istnieje nadawca charakteryzujący się intencjonalnością, który kreuje tekst o stabilnej referencyjności (tj. „prawdzie” umocowanej w rzeczywistości pozatekstowej). Symulacja neutralizuje zatem kluczowe założenia fikcjonalności: intencjonalność oraz kontekstowo zakotwiczony dyskurs. Znaki nie sygnalizują – reprodukują i nie odzwierciedlają żadnej sfery wyobrażonej, lecz jedynie algorytmicznie przetwarzają wzorce na podstawie statystycznie opracowanych danych treningowych.

LLM-y, takie jak Claude czy ChatGPT, funkcjonują w pełni w logice symulacji Baudrillarda. Nie komponują tekstów poprzez wolę autorską, lecz wykonują predykcję kolejnego tokenu na podstawie korpusów treningowych. Ich wytwory są generatywne – stanowią produkty optymalizacji obliczeniowej. Nieobecność nadawcy prowadzi do załamania kontraktu fikcjonalnego: nie ma twórcy wyobrażającego tego, co nierzeczywiste i – by przywołać Bender i współautorów – nie ma intencji komunikacyjnej, modelu świata ani konkretnego obrazu odbiorcy. Systemy LLM, produkując przekonujące symulacje dyskursu bez intencji, referencji czy wyobraźni, ucieleśniają czwartorzędowy symulakr Baudrillarda: znaki krążące w próżni realnego, oderwane od źródła i obojętne wobec prawdy. Opowiadania generowane przez LLM-y nie ukrywają swojej „sztuczności” czy braku autora – odgrywają retorykę fikcjonalności, jednocześnie wydrążając jej epistemiczne fundamenty. Ich logika generatywna zastępuje semantyczną obecność statystycznym przybliżeniem, symulując literacką koherencję jako tryb odraczania, nie ekspresji. Jak ujmuje to Baudrillard: „Koniec teatru reprezentacji, przestrzeni znaków, ich konfliktu, ich milczenia; pozostaje jedynie czarna skrzynka kodu [...]”³⁵.

Traktowanie wytworów LLM-ów jako opowiadań w sensie zaproponowanym przez Nielsena, Phelana i Walsha oznacza błędne rozpoznanie ich statusu ontologicznego³⁶. Zamiast rozszerzać kategorię fikcjonalności, teksty generowane przez AI zmuszają nas do konfrontacji z jej wyczerpaniem. Jako teoretycy narracji powinniśmy pytać nie o to, co te teksty znaczą, lecz o to, co oznacza wytwarzać efekty znaczeniowe przy braku znaczenia. W tym miejscu teoria Baudrillarda nie tyle uzupełnia teorię fikcjonalności, ile ją wypiera. Epoka cyfrowych symulaków inauguruje nowy paradygmat tekstowy: nie opowiadanie, lecz symulację.

3. Narracje generowane przez AI jako przypadki testowe

15 grudnia 2024 roku skorzystałem z interfejsu Claude AI, aby wygenerować dziesięć krótkich tekstów narracyjnych w języku angielskim w odpowiedzi na następujący prompt: „Wygeneruj 10 różnych wersji opowiadania o polskim nauczycielu akademickim. Nie przekraczaj 100 słów w każdym opowiadaniu”. Ćwiczenie przeprowadziłem w kontekście prowadzonych badań

³⁵Baudrillard, *Simulations*, 104.

³⁶Halpern. Zob. również Natale, a także Fletcher, „Why Computer AI Will Never Do What We Imagine It Can” oraz Angus Fletcher, „Why Computers Will Never Read (or Write) Literature: A Logical Proof and a Narrative”, *Narrative 1* (2021): 1–28.

nad ontologią narracji generowanych przez AI oraz ich implikacjami dla teorii literatury i narratologii. Początkowo miało ono charakter eksperymentu z nową technologią, jednak uzyskane wyniki stały się później konceptualną podstawą propozycji referatu na konferencję *Awatary, asystenci, chatboty. Nowe byty fikcjonalne w kulturze współczesnej (od literatury do nowych mediów)*, która odbyła się w Warszawie w dniach 8–9 kwietnia 2025 roku.

Wygenerowany korpus obejmuje dziesięć opowiadań, z których każde liczy około 80–100 słów, przedstawiających fikcyjne postaci osadzone w realiach polskich uczelni. Każde opowiadanie zawiera zwięzłą fabułę, wyraźnie zogniskowaną perspektywę oraz rozpoznawalny trop przemiany, konfliktu bądź objawienia. Opowiadania są formalnie spójne i odpowiadają gatunkowym oczekiwaniom krótkiej formy epickiej, alegorii dydaktycznej oraz charakterystyki postaci.

Poniżej przedstawiam streszczenia każdego z wygenerowanych opowiadań; kompletne wersje w języku angielskim zamieszczone są w aneksie.

1. *Niechętny rewolucjonista* – Profesor matematyki w 1981 roku musi wybrać między stabilnością zawodową a przyłączeniem się do protestów „Solidarności”.
2. *Uczony o północy* – Mediewista odkrywa przełomowy dokument archiwalny w Bibliotece Uniwersytetu Warszawskiego.
3. *Most języka* – Wykładowczyni języka polskiego odnajduje głębszy sens egzystencji dzięki uczeniu ukraińskiej uchodźczynie.
4. *Cyfrowe nawrócenie* – Starszy historyk z niechęcią przyswaja technologię VR podczas pandemii COVID-19.
5. *Strażnik poezji* – Woźny zapamiętuje polską poezję recytowaną codziennie przez profesora literatury.
6. *Chemia przemiany* – Nauczyciel chemii zmienia zaangażowanie studentów, używając metafor relacji międzyludzkich.
7. *Powrót syna* – Fizyk powracający z Harvardu rzuca wyzwanie prowincjonalnemu sceptycyzmowi, inicjując międzynarodowy projekt badawczy.
8. *Uczony w ogrodzie* – Zaniedbany ogród staje się miejscem filozoficznego dialogu i odnowy wspólnoty.
9. *Strażnik czasu* – Profesor słynący z punktualności spóźnia się po raz pierwszy z powodu narodzin wnuka.
10. *Głos* – W efekcie utraty głosu wykładowca językoznawstwa prowadzi całe zajęcia w języku migowym, ujawniając przy tym ukrytą historię rodzinną.

Ze względu na ograniczenia objętości przeprowadzę szczegółową analizę tylko jednego opowiadania wygenerowanego przez Claude'a. *Niechętny rewolucjonista* egzemplifikuje wszystkie główne przesłanki mojego argumentu.

W 1981 roku profesor Jan Kowalski stanął w obliczu dylematu: przyłączyć się do swoich studentów uczestniczących w protestach „Solidarności” czy też zachować swoją pozycję na uniwersytecie. Jego specjalnością była matematyka, nie polityka, lecz gdy patrzył, jak jego najzdolniejsi studenci ryzykują wszystko w imię wolności, uświadomił sobie, że niektórych równań nie da się rozwiązać na tablicy.

Na pierwszy rzut oka wydaje się, że opowiadanie posiada formalne elementy typowego opowiadania. Występuje w nim bohater (profesor Jan Kowalski), precyzyjnie określony czas wydarzeń (Polska, 1981), wyraźny dylemat moralny (wesprzeć ruch „Solidarności” czy zachować pozycję), a także adekwatne i sugestywne metaforyczne domknięcie. Markery te mogą skłaniać czytelnika do odczytania tekstu jako miniaturowej formy literackiej – tego, co Nielsen, Phelan i Walsh nazwaliby retorycznie sygnalizowaną komunikacją fikcjonalną. Jednak, jak wskazano w sekcji 2.1., nie można zapomnieć, że LLM-y to systemy algorytmiczne, które przekształcają język w wektory i następnie obliczają prawdopodobieństwa tokenów w celu generowania wyników poprzez próbkowanie stochastyczne. Ich architektura transformera wykorzystuje mechanizmy uwagi do matematycznego ważenia relacji między słowami, symulując koherencję bez rozumienia. Modele te biegle „zszywają” formy językowe, ale czynią to bez świadomości referencyjnej i semiotycznej niczym „stochastyczne papugi” na podstawie inferencji probabilistycznej.

Uważam, że *Niechętny rewolucjonista* bardzo dobrze ilustruje opisany *modus operandi* modelu AI. Imię bohatera jest tak typowo polskie, że wykorzystuje się je w instrukcjach wypełnienia formularzy urzędowych. Ruch „Solidarność” z lat 80. jest jednym z najbardziej rozpoznawalnych momentów współczesnej historii Polski i stale obecnym w dyskusjach politycznych oraz kulturowych dotyczących wolności, praw pracowniczych i aktywizmu na rzecz praw człowieka³⁷. Struktura narracyjna opowiadania – konflikt wewnętrzny między pragnieniem a obowiązkiem – stanowi klasyczne rozwiązanie fabularne. Od Antygony i Hamleta po Raskolnikowa czy Simbę z *Króla lwa* – dylemat wewnętrzny jest tropem silnie osadzonym w zachodniej kulturze. W dyskursie narratologicznym wspomniane struktury zostały wielokrotnie zanalizowane, między innymi w *Morfologii bajki* Władimira Proppa, *Reading for the Plot* Petera Brooksa, *Czasie i opowieści* Paula Ricoeura czy w modelu aktantów Algirdasa Juliena Greimasa. Metaforyczne domknięcie – „niektórych równań nie da się rozwiązać na tablicy” – stanowi zręczne połączenie profesji Kowalskiego z wagą jego wyboru moralnego, jednak jednocześnie trudno się oprzeć wrażeniu, że jest ono dość przewidywalne. Jego funkcja polega, zdaje się, na domknięciu mikroopowiadania i nadaniu mu głębi oraz uniwersalności. W istocie jednak brzmi jak sztampowy chwyt retoryczny lub schematyczna formułka, która łączy dyscyplinę naukową protagonisty z jego dylematem moralnym i wyraża jednocześnie główny temat historii. Całe opowiadanie, choć pod względem składni wydaje się

³⁷Zob. np. „Solidarity on Screen” w [The International Cultural Programme of the Polish Presidency of the Council of the European Union 2025](#).

wzorowo uporządkowane, pozostaje *de facto* koncepcyjnie puste: jest statystycznym echem podobnych rozwiązań fabularnych obecnych w danych treningowych modelu.

Niechętny rewolucjonista nie jest tekstem fikcyjnym w sensie retorycznym. „Solidarność” i „Jan Kowalski” funkcjonują nie jako historyczne czy charakterologiczne znaczone zakorzenione w kontekście, lecz jako unoszące się znaczące, probabilistycznie dobrane w odpowiedzi na słowa kluczowe zawarte w prompcie. Opowiadanie – parafrazując Baudrillarda – jest tekstem symulowanym, który „pozostaje bez związku z jakąkolwiek rzeczywistością”³⁸. Sygnały „polskości” służą – znów cytując Baudrillarda – „wprowadzeniu na masową skalę zasady rzeczywistości i referencyjności” do systemu, w którym referenty nie istnieją³⁹. Można je określić zatem jako „nie-wydarzenie” i „nie-imię” – znaki, które unoszą się w „orbitalnej powtarzalności” syntezy „kombinatorycznych modeli w pozbawionej atmosfery hiperprzestrzeni”⁴⁰ – bądź, w kontekście mojej argumentacji, węzły probabilistycznej koherencji w zmechanizowanych narzędziach do rozpoznawania wzorców Claude’a, przypadkowo „zszytych” z różnych form, tropów, chwytów lub metafor zaobserwowanych w jego rozległych danych treningowych⁴¹. Jak ujmuje to Baudrillard, symulacja „wymija wszelkie [...] zmienne koleje losu”, zastępując je operacyjnymi kopiami. W konsekwencji otrzymujemy tekst prezentujący hiperrzeczywistą wersję opowiadania: „czystsza”, bardziej archetypiczną, pozbawioną przygodności jakiegokolwiek zakorzenionej historycznie narracji.

Teoria symulacji dostarcza zatem trafniejszej ramy analitycznej niż retoryczna teoria fikcyjności Nielsena, Phelana i Walsha. Ta ostatnia opiera się na założeniach intencjonalności kontraktu komunikacyjnego, woli autorskiej czy poznawczej świadomości rozróżnienia między tym, co fikcyjne, a tym, co niefikcyjne – z których żadne nie znajduje zastosowania w przypadku tekstów generowanych przez LLM-y. Teoria symulacji natomiast rozumie te teksty jako produkty autoreferencyjnego systemu semiotycznego, w którym znaczące krążą bez zakotwiczonych referentów. Rezultatem nie jest fikcyjność, lecz symulacja fikcyjności. Głębia charakterologiczna, znajomość kontekstu historycznego i ukazanie ambiwalencji moralnej zostają algorytmicznie wytworzone poprzez statystyczne wzorcowanie. To, z czym mamy do czynienia, nie jest literaturą w sensie humanistycznym, lecz *tekstualnością maszynową* – hiperrzeczywistą produkcją znaków narracyjnych pozbawionych ontologicznego fundamentu. Teksty te nie opowiadają historii; imitują formę opowiadania historycznego. Nie wyobrażają postaci; replikują kształt wyobrażonej podmiotowości. Jak pisze Baudrillard, „symulacja nie jest już symulacją terytorium, przedmiotu odniesienia, substancji. Stanowi raczej sposób generowania – za pomocą modeli – rzeczywistości pozbawionej źródła i rzeczywistości [...] Od tej pory to mapa poprzedza terytorium [...]”⁴². W analogiczny sposób narracje generowane przez LLM-y nie odzwierciedlają świata; uprzedzają go, generując efekty fikcyjności przed znaczeniem. W epoce narracji algorytmicznej nie obcujemy już z opowiadaniem jako ekspresją

³⁸Baudrillard, *Symulakry i symulacje*, 12.

³⁹Baudrillard *Symulakry i symulacje*, 28–34.

⁴⁰Baudrillard, *Symulakry i symulacje*, 6–7.

⁴¹Bender i in., 617.

⁴²Baudrillard, *Symulakry i symulacje*, 6.

ludzką, lecz z jego operacyjnym sobowtórem – zautomatyzowanym i kombinatorycznym. Opowiadanie nie jest pisane – jest symulowane.

4. Zakończenie: ku teorii symulacji fabuł

Niniejszy artykuł miał na celu ukazanie, jak opowiadania generowane przez AI nie podlegają klasyfikacji jako narracje w sensie narratologiczno-retorycznym. Teksty te należy rozumieć raczej jako symulacje: strukturalnie mimetyczne, lecz pozbawione intencjonalności, referencji i tradycyjnego kontraktu komunikacyjnego.

Model Nielsena, Phelana i Walsh'a oferuje retorycznie złożoną koncepcję fikcjonalności, jednak ulega on załamaniu w konfrontacji z generacją algorytmiczną – LLM-y nie są autorami, lecz silnikami statystycznymi; odwzorowują artefakty językowe, nie idee; ich wytwory symulują fikcjonalność, nie spełniając jej warunków. Rama teoretyczna Baudrillarda zdaje się trafniej ujmować rozpad antropocentrycznych przesłanek, które podtrzymywały kategorię opowiadania jako aktu intencjonalnej ekspresji. Teksty generowane przez AI są symulacjami nie dlatego, że pozbawione są cech narracyjnych, lecz dlatego, że funkcjonują ortogonalnie wobec opowiadania: ani go nie potwierdzają, ani nie negują, lecz podważają spójność jego infrastruktury pojęciowej.

Niniejsza diagnoza generuje nowe pytania krytyczne: czy symulację można steoretyzować jako odrębny tryb narracyjny? Jakie ryzyko niesie interpretowanie tekstów generowanych jako literatury? Jak należy traktować autorstwo w sytuacji, gdy autor w sensie tradycyjnym nie istnieje? Pytania te pozostają nierozstrzygnięte nie dlatego, że są marginalne, lecz dlatego, że ontologiczna niestabilność ich przedmiotu uniemożliwia definitywne rozstrzygnięcie.

Artykuł ten zatem nie zamyka problemu, tylko go odracza. Wzywa do rekonfiguracji dyskursu literaturoznawczego w odpowiedzi na transformację technologiczną. W nowej ekologii tekstowej opowiadanie nie jest już wyłącznie pisane – jest symulowane. Nasze słownictwo krytyczne musi ulec przekształceniu: nie poprzez restaurację, lecz poprzez renegocjację. Ontologiczny status narracji generowanej przez AI nie powinien być dekretowany, lecz rozumiany jako zjawisko emergentne – negocjowane pomiędzy dyscyplinami, praktykami lekturowymi i paradygmatami teoretycznymi. Być może nie powinniśmy już pytać: „Czy to jest opowiadanie?”, lecz raczej: „Co dziś znaczy symulować opowiadanie?”.

Bibliografia

- Aarseth, Espen. *Cybertext: Perspectives on Ergodic Literature*. Baltimore: Johns Hopkins University Press, 1997.
- Bahdanau, Dzmitry, Kyunghyun Cho, Yoshua Bengio. „Neural Machine Translation by Jointly Learning to Align and Translate”. Prezentacja wygłoszona podczas International Conference on Learning Representations (ICLR), 2015. <https://arxiv.org/abs/1409.0473>.
- Bajohr, Hannes. „Writing at a Distance: Notes on Authorship and AI”. *YouTube*, 26.03.2025, 1:29:38. <https://www.youtube.com/watch?v=pP66rSFjPYw>.
- Baudrillard, Jean. *Simulations*. Tłum. Phil Beitchman, Paul Foss, Paul Patton. New York: Semiotext(e), 1983.
- – –. *Symulakry i symulacje*. Tłum. Sławomir Królak. Warszawa: Sic!, 2005.
- Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, Shmargaret Shmitchell. „On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?”. Prezentacja wygłoszona podczas FAccT '21, New York: Association for Computing Machinery, 2021. <https://doi.org/10.1145/3442188.3445922>.
- Bogost, Ian. *Persuasive Games: The Expressive Power of Video Games*. Cambridge: MIT Press, 2007.
- Chun, Jon, Katherine Elkins. „What the Rise of AI Means for Narrative Studies: A Response to «Why Computers Will Never Read (or Write) Literature» by Angus Fletcher”. *Narrative 1* (2022): 104–113.
- Dawson, Paul. „Ten Theses Against Fictionality”. *Narrative 1* (2015): 74–100.
- Fletcher, Angus. „Why Computer AI Will Never Do What We Imagine It Can”. *Narrative 1* (2022): 114–137.
- – –. „Why Computers Will Never Read (or Write) Literature: A Logical Proof and a Narrative”. *Narrative 1* (2021): 1–28.
- Frasca, Gonzalo. „Simulation”. W: *The Johns Hopkins Guide to Digital Media*, red. Marie-Laure Ryan, Lori Emerson, Benjamin J. Robertson, 452–455. Baltimore: Johns Hopkins University Press, 2014.
- Gorman, David. „Theories of Fiction”. W: *The Routledge Encyclopaedia of Narrative Theory*, red. David Herman, Manfred Jahn, Marie-Laure Ryan, 164–167. Oxon: Routledge, 2005.
- Halpern, Mark. „No Ghost in the Machine: Artificial Intelligence Isn't as Intelligent as You Think”. *The American Scholar*, 2.03.2020. <https://theamericanscholar.org/no-ghost-in-the-machine/>.
- Natale, Simone. *Deceitful Media: Artificial Intelligence and Social Life after the Turing Test*. Oxford: Oxford University Press, 2021.
- Nielsen, Henrik Skov. „Natural Authors, Unnatural Narration”. W: *Postclassical Narratology: Approaches and Analyses*, red. Jan Alber, Monika Fludernik, 275–300. Columbus: Ohio State University Press, 2010.
- Nielsen, Henrik Skov, James Phelan, Richard Walsh. „Ten Theses About Fictionality”. *Narrative 1* (2015): 61–73.

Phelan, James. *Living to Tell about It: A Rhetoric and Ethics of Character Narration*. Ithaca: Cornell University Press, 2005.

Radford, Alec, Karthik Narasimhan, Tim Salimans, Ilya Sutskever. *Language Models Are Unsupervised Multitask Learners*. San Francisco: OpenAI, 2019. https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.

Turing, Alan. „Computing Machinery and Intelligence”. *Mind* 49 (1950): 433–460. <https://courses.cs.umbc.edu/471/papers/turing.pdf>.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, Illia Polosukhin. „Attention Is All You Need”. *Advances in Neural Information Processing Systems* 30 (2017). <https://arxiv.org/abs/1706.03762>.

Walsh, Richard. *The Rhetoric of Fictionality: Narrative Theory and the Idea of Fiction*. Columbus: Ohio State University Press, 2007.

Wei, Jason, Yi Tay, Rishi Bommasani, Dan Roth, Percy Liang. „Emergent Abilities of Large Language Models”. *Transactions on Machine Learning Research*, 2022. <https://arxiv.org/abs/2206.07682>.

Zeigler, Bernard P., Herbert Praehofer, Tag Gon Kim. *Theory of Modeling and Simulation*. London: Academic Press, 2000.

Ziętek, Agnieszka. *Jean Baudrillard wobec współczesności. Polityka, media, społeczeństwo*. Kraków: Universitas, 2013.

SŁOWA KLUCZOWE:

narratologia

SYMULACJA

fikcjonalność

ABSTRAKT:

Artykuł proponuje analizę ontologicznego statusu opowiadań generowanych za pomocą sztucznej inteligencji poprzez porównawcze odczytanie teorii fikcjonalności (Gorman 2005; Nielsen, Phelan i Walsh 2015) oraz teorii symulacji Jeana Baudrillarda (1981, 1983). Na podstawie analizy korpusu krótkich historii wygenerowanych przez system Claude AI dowodzę, że teksty te nie wytwarzają narracji fikcjonalnej w sensie retorycznym, lecz jedynie symulują jej formalne właściwości. Wskazuję jednocześnie, że tradycyjne kategorie narratologiczne nie są w stanie uchwycić epistemicznego pęknięcia wprowadzonego przez duże modele językowe (LLM) oraz że konieczne jest wypracowanie nowej hermeneutyki narracji symulowanej.

narracja cyfrowa

literatura generowana przez AI

DUŻE MODELE JĘZYKOWE

Claude AI

NOTA O AUTORCE:

Bartosz Lutostański – adiunkt w Instytucie Anglistyki Uniwersytetu Warszawskiego. Jego badania dotyczą fikcji cyfrowej, publikacji postcyfrowych, krytycznych badań nad sztuczną inteligencją oraz współczesnych kultur mediów społecznościowych. Autor rozdziałów w zagranicznych tomach: *The Handbook of Intermediality* (Palgrave Macmillan, 2023) oraz *New Approaches to Transcodification* (De Gruyter Brill, 2025), a obecnie współredaktor książki *Postdigital Fiction: Theory, Method, Analysis* (Routledge, 2026). Jest zastępcą redaktora naczelnego „Anglica: An International Journal of English Studies” oraz członkiem rady redakcyjnej „Storyworlds: A Journal of Narrative Studies”. Zajmuje się również narracją opartą na aplikacjach, literaturą generowaną przez sztuczną inteligencję, kulturami influencerów oraz narracją transmedialną, ze szczególnym uwzględnieniem estetyki interfejsu, kultur partycypacyjnych i mediacji algorytmicznej. |