

# POMIĘDZY MARZENIEM A RZECZYWISTOŚCIĄ. WSPÓŁCZESNA SZTUCZNA INTELIGENCJA A JEJ PRZEDSTAWIENIA W FIKCJI LITERACKIEJ I FILMOWEJ NA PRZESTRZENI LAT

**Magdalena Siwek**

absolwentka Uniwersytetu Jagiellońskiego  
magdalena.siwek28@gmail.com

**Abstrakt:** Celem badawczym niniejszego artykułu jest porównanie dostępnej współcześnie sztucznej inteligencji z jej literackimi i filmowymi przedstawieniami, oraz omówienie wpływu tekstów kultury na nasze postrzeganie rozwijających się technologii. Poddaje on analizie zarówno pozytywne, jak i negatywne przedstawienia rozumnych maszyn w fikcji, i konfrontuje je z obecnym stanem sztucznej inteligencji w poszukiwaniu głównych punktów wspólnych i różnic. Artykuł kładzie szczególny nacisk na to, jak powszechne wyobrażenia o tym, czym powinna być prawdziwa sztuczna inteligencja, odwracają naszą uwagę od korzyści, czerpanych z najnowszych technologii, a także wiążących się z nimi zagrożeń.

**Słowa kluczowe:** sztuczna inteligencja, *science fiction*, SI w fikcji filmowej i literackiej

## **BETWEEN DREAMS AND REALITY. CONTEMPORARY ARTIFICIAL INTELLIGENCE AND ITS PORTRAYALS IN LITERARY AND CINEMA FICTION THROUGH THE YEARS**

**Abstract:** The main aim of this article is to compare contemporary artificial intelligence with its literary and cinematic representations, and to discuss the influence of fiction on our perception of developing technologies. It analyses both positive and negative portrayals of intelligent machines in fiction and confronts them with the current state of artificial intelligence in search of key similarities and differences. The article places particular emphasis on how common perceptions of what true artificial intelligence should be tend to distract us from the benefits of the latest technologies, as well as the risks they entail.

**Key words:** artificial intelligence, *science fiction*, AI in literary and cinema fiction

W 2022 roku, po długich latach snucia hipotez, przeprowadzania obliczeń i testowania rozmaitych teorii, ludzkość wkroczyła w okres, nazywany już powszechnie wiosną sztucznej inteligencji. Udostępnienie szerszej publice chatbota ChatGPT nie tylko otworzyło przed jego użytkownikami szereg zdumiewających możliwości, ale również zapoczątkowało przemianę w naszym

postrzeganiu tego, czym jest i do czego jest zdolna sztuczna inteligencja. Wszystkie korzyści i niebezpieczeństwa, łączące się z jej wykorzystaniem, a dotychczas pozostające w sferze fikcji, urosły niespodziewanie do rangi znaczących współcześnie zagadnień. W środowisku badaczy obietnice świetlanej przyszłości, jaką może nam zapewnić sztuczna inteligencja, mieszają się z ostrzeżeniami: tak zwane „największe osiągnięcie ludzkości” może się równie dobrze okazać jej zgubą (Harris 2024: 9).

Wszystkie nasze nadzieje i lęki, marzenia i teorie spiskowe wznoszą się jednak na tym samym fundamencie – na fikcji filmowej i literackiej, która od dziesiątków lat karmi nas wyobrażeniami o tym, jak wykorzystanie sztucznej inteligencji może się odbić na naszej rzeczywistości. Współcześnie, kiedy jej możliwości dopiero się przed nami odkrywają i nie sposób ściśle przewidzieć kierunku jej rozwoju, warto poddać analizie wpływ, jaki na jej postrzeganie wywarła fikcja. Ostatecznym celem niniejszego artykułu jest ustalenie, na ile nasze wyobrażenie o sztucznej inteligencji pokrywa się z jej rzeczywistością, dostępną nam wersją.

Pierwsze wzmianki o myślących maszynach napotykamy już w tekstach starogreckich: w *Iliadzie* Hefajstosowi usługiwały wykute ze złota panny, a w micie o Argonautach Krety strzegł młodzieniec z brązu, Talos. Temat ten został ponownie podjęty przez XIX-wiecznych pisarzy, zainspirowanych dynamicznym rozwojem nauki i techniki. *Frankenstein* Mary Shelley porusza kwestię odpowiedzialności, spoczywającej na twórcach nowego, „sztucznego” życia (Shelley 1818), zaś *Erewhon* Samuela Butlera opowiada o zagrożeniach, łączących się z powstaniem myślących, samoreplikujących się maszyn (Butler 1872). Jednak motyw sztucznej inteligencji nabrał prawdziwej popularności i znaczenia dopiero w drugiej połowie XX wieku, wraz z wynalezieniem i rozpowszechnieniem się komputerów.

W swojej pracy naukowej z 2019 roku Stephen Cave dokonał podziału nadziei i lęków, związanych z rozwojem sztucznej inteligencji, na osiem równoważących się podtypów: nieśmiertelność (immortality) a nieludzkość (inhumanity), łatwość użycia (ease) a popadnięcie w zapomnienie (obsolescence), wynagrodzenie (gratification) a wyobcowanie (alienation), dominacja (dominance) a bunt (uprising) (Cave 2019: 75). Na potrzeby niniejszego artykułu wykorzystanie zostanie mniej ścisły podział: na przedstawienia generalnie pozytywne (nadzieje ludzkości) oraz generalnie negatywne (jej lęki). Wystarczy on, by zaprezentować, jaki obraz sztucznej inteligencji kreśli nam współcześnie szeroko ujęta fikcja filmowa i literacka.

W mediach i literaturze ostatnich dekad nie brak pozytywnych przedstawień sztucznej inteligencji. W wielu przypadkach myślące maszyny są pełnoprawnymi bohaterami narracji: odczuwają one ludzkie emocje i targają nimi ludzkie wątpliwości. Znanym przykładem tego rodzaju myślącego robota jest android Data, bohater emitowanego w latach 1986-1994 serialu *Star Trek: The Next Generation*. Choć jest on inteligentną i życzliwą istotą – właścicielem rudego kota imieniem Goose i najlepszym przyjacielem porucznika Geordiego La Forge’a – raz po raz zmuszony jest udowadniać, że zasługuje na to, by być uważanym za obywatela Federacji. W drugim sezonie serialu uczestniczy nawet w procesie sądowym, mającym zadecydować, czy jest istotą żywą, czy własnością Floty Gwiezdnej (“The Measure of The Man” 1989). Data, w rolę którego wcielił się Brent Spiner, jest przykładem sztucznej inteligencji, która w niewielkim tylko stopniu różni się od zwykłego człowieka; nie programem, mającym ułatwiać życie swoim użytkownikom, ale jednostką o oryginalnych opiniach i odczuciach.

Innym humanoidalnym robotem-indywidualistą o mocno zarysowanym charakterze jest Mordbot, tytułowa osoba bohatera *Pamiętników Mordbota* Marthy Wells (Wells 2017). W futurystycznej przyszłości, w której eksplorację kosmosu kontrolują wielkie korporacje, Mordbot – jednostka ochroniarska do wynajęcia – hakuje swój system i zyskuje wolną wolę. Początkowo rozważa sianie śmierci i zniszczenia w całej galaktyce, ostatecznie dochodzi jednak do wniosku, że woli spędzać czas, oglądając opery mydlane. Choć Mordbot przedstawia się jako „ono”, jest w pełni uformowaną istotą, z szeregiem pasji, bagażem traum i grupą bliskich przyjaciół. W ponurym, niemalże dystopijnym uniwersum serii pozostaje niewątpliwie pozytywnym przykładem sztucznej inteligencji.

Nie we wszystkich tekstach kultury sztuczna inteligencja przyjmuje jednak humanoidalną formę. Sir Terry Pratchett uczynił jednym z pobocznych bohaterów swojego cyklu powieści, Świata Dysku, napędzany magią komputer Hex: niezwykle połączenie mrowiska, akwarium, ula, mysiego gniazda i maszyny. Choć Hex na porządku dziennym wykonuje najbardziej skomplikowane obliczenia, nawet jego wynalazcy nie mają pewności, czy jest zdolny do samodzielnego myślenia. Interesujące światło na tę kwestię rzuca scena z wydanego w 1996 roku *Wiedźmikołaja*, w której magiczny komputer stara się przekazać samemu Wiedźmikołajowi – Dyskowemu odpowiednikowi Świętego Mikołaja – swój list z życzeniami. Śmierć, wcielający się akurat w tę rolę, zwraca mu uwagę, że jako maszyna nie może mieć żadnych pragnień. „All things strive” odpowiada Hex (Pratchett 1996, s. 361). Finalnie Śmierć spełnia marzenie komputera: wręcza mu pluszowego misia, który od tego momentu staje się jego nieodłącznym towarzyszem.

Miłość do sera i przytulank nie powstrzymują Hexa przed wykonywaniem swoich podstawowych funkcji: analizowaniem zakłęb i dokonywaniem magicznych obliczeń. Ostatecznie, podobnie jak większość myślących maszyn, powstał, by ułatwić życie swym wynalazcom. Do tej samej kategorii sztucznych inteligencji zaliczają się chociażby J.A.R.V.I.S. – wirtualny pomocnik Tony'ego Starka w serii filmów *Ironman* (*Ironman* 2008) – czy tytułowy bohater animowanego filmu *WALL·E* – niewielki robot, dzień po dniu i odpad po odpadzie oczyszczający porzuconą przez ludzkość, wyniszczoną Ziemię (*WALL·E* 2008). Choć obie te maszyny mają swoją świadomość i charakter, poznajemy je w pierwszej kolejności przez pryzmat wykonywanej przez nie pracy.

Nie należy również zapominać o fikcyjnych systemach, które nie wykształciły samoświadomości, a jednak z dużym prawdopodobieństwem zawierają w sobie elementy szczątkowej sztucznej inteligencji. To niezawodne uniwersalne tłumacze i replikatory jedzenia w *Star Treku* czy soniczny śrubokręt w *Doktorze Who* – urządzenia o złożonym systemie działania, których sprawności nie jest w stanie objaśnić żadna konfiguracja śrub i zębatek. Te systemy, pozostające zwykle na trzecim planie książkowej czy filmowej fabuły, niejednokrotnie przynoszą bohaterom więcej korzyści niż obdarzone wybitną inteligencją androidy, i zdecydowanie zasługują na uznanie.

Pozytywne przedstawienia sztucznej inteligencji, które zakorzeniły się w ludzkiej świadomości, można więc podzielić na trzy główne podtypy. Są to:

1. Sztuczna inteligencja o własnych uczuciach i opiniach, wiodąca życie jako wartościowy członek społeczeństwa.
2. Sztuczna inteligencja, zazwyczaj obdarzona własnym charakterem, ale przy tym raczej niezdolna do samodzielnego życia, i wypełniająca konkretne zadania dla swoich właścicieli.
3. Zaawansowane maszyny, których niezawodność i wszechstronność objaśnić można działaniem sztucznej inteligencji w jej najbardziej podstawowym wariantcie.

Nietrudno jest sobie wyobrazić, że odtworzenie w rzeczywistym świecie któregoś z tych podtypów przyniosłoby ludzkości znaczące korzyści i w pełni usprawiedliwiłoby entuzjazm, przejawiany przez część badaczy.

A jednak, choć użycie sztucznej inteligencji byłoby w stanie otworzyć przed nami tak wiele nowych możliwości, w mediach i literaturze wciąż jeszcze dominują pesymistyczne przedstawienia naszej skomputeryzowanej przyszłości. Nie powinno to dziwić: jedną z podstawowych misji *science fiction* jest wskazywanie zagrożeń, jakie mogą z sobą nieść nowe technologie. Nawet *Star Trek*, ukazujący stosunkowo pozytywną wizję przyszłości, zwraca uwagę na niebez-

pieczeństwa, łączące się z użyciem sztucznej inteligencji. W *The Ultimate Computer* – odcinku oryginalnej serii, wyemitowanym w 1968 roku – na pokładzie USS Enterprise odbywa się test superkomputera M-5, w założeniu zdolnego do zastąpienia dowolnego kapitana Floty Gwiezdnej. Początkowa nieomyślność maszyny zdaje się dowodzić jej wyższości nad ludźmi. Wkrótce jednak M-5 przejmuje kontrolę nad statkiem i rozpoczyna ostrzał pozostałych jednostek Floty. Załoga Enterprise łączy siły w walce ze sztuczną inteligencją. „Computers make excellent and efficient servants, but I have no wish to serve under them. Captain, the starship also runs on loyalty to one man. And nothing can replace it, or him” podsumowuje fabułę odcinka pan Spock (‘The Ultimate Computer’ 1968).

Warto zauważyć, że decyzje, podejmowane przez M-5, w znacznej mierze pozostają w zgodzie z wpojonymi mu zasadami. Powierzono mu ochronę Enterprise, robi więc co w jego mocy, by zapewnić statkowi bezpieczeństwo. Dopiero uświadomiony przez kapitana Kirka, że jego akcje doprowadziły do śmierci setek ludzi, M-5 pojmując swój błąd i postanawia wymierzyć sobie adekwatną karę: dokonuje samodestrukcji.

Motywy maszyny, która działa w najlepszych intencjach, a mimo to przyczynia ludzkości szkodę, regularnie przewijają się w serialach i powieściach. W filmie *I, Robot* – zainspirowanym zbiorem opowiadań Isaaca Asimova o tym samym tytule – sztuczna inteligencja VIKI dochodzi do wniosku, że największym zagrożeniem dla ludzkości są sami ludzie, i zaczyna siać zniszczenie w Chicago (*I, Robot* 2004). Z kolei w odcinku *Doktora Who* pod tytułem *Smile* grupa robotów, mających dbać o zadowolenie załogi statku kosmicznego, zaczyna traktować smutek jako chorobę i zabijać wszystkich jej „nosicieli” (‘Smile’ 2017). W tego rodzaju narracjach trudno jest ocenić, czy wina za popełnione zbrodnie leży w rzeczywistości po stronie działającej w zgodzie ze swoim oprogramowaniem sztucznej inteligencji, czy jej nieostrożnych wynalazców.

Nie w każdym przypadku zachowania myślących maszyn można jednak usprawiedliwić błędnym rozumieniem otrzymanych instrukcji. Doskonale wiedzą o tym fani *Terminatora* – wyprodukowanego w 1984 roku thrillera, któremu udało się wywrzeć niezaprzeczalny wpływ na współczesne postrzeganie sztucznej inteligencji. W świecie przedstawionym filmu władzę nad światem sprawują roboty, kontrolowane przez świadomy system komputerowy, Skynet. W próbie ostatecznego unicestwienia ludzkiej rebelii decyduje się on wysłać w przeszłość cyborga-zabójcę, tytułowego Terminatora. Jego celem ma się stać Sarah Connor, przyszła matka przywódcy rebeliantów. Ścigana przez bezlitosnego prześladowcę, Sarah zmuszona jest walczyć nie tylko o własne życie, ale również o przetrwanie całej ludzkości (*Terminator* 1984).

Wątek żądnych władzy maszyn odnajdujemy również w *Star Treku* – serialu, który przez długie lata swojej emisji tak często poświęcał uwagę sztucznej inteligencji, że odwoływanie się do niego na każdej niemal stronie tego artykułu jest nie przejawem faworyzmu, a swego rodzaju koniecznością. W trzynastym odcinku pierwszego sezonu *Star Trek: New Generation* załoga Enterprise odwiedza rodzinną planetę porucznika Daty, Omicron Thetę. Natykają się tam na jego „brata” – stworzonego przez tego samego genialnego naukowca androida imieniem Lore. Choć początkowo zdaje się on być wierną kopią porucznika Daty, szybko wychodzi na jaw, że w rzeczywistości jest narcystycznym potworem, przekonanym o swej wyższości nad organicznymi formami życia. Jego zraniona duma i nienawiść wiele lat wcześniej doprowadziły do śmierci tysięcy mieszkańców Omicron Thety. Podobnie jak Skynet, Lore nie ma powodu, by krzywdzić otaczających go ludzi: robi to wyłącznie dlatego, że chce i może (‘Datalore’ 1988).

Nie należy też zapominać o sztucznych inteligencjach, które buntują się przeciwko ludzkości w akcie samoobrony. Wielu autorów powieści i filmów *science fiction* zwraca uwagę na to, jaki wpływ na relacje z androidami może wywrzeć ludzka bezwzględność i małostkowość. Sztuczna inteligencja – potencjalny pomocnik i przyjaciel – może się równie łatwo stać pozbawionym głosu niewolnikiem. Motyw ten odnaleźć można już w *R.U.R.*, sztuce teatralnej Karela Čapka, po raz pierwszy wystawionej na deskach czeskiego teatru w 1921 roku. Opowiada ona historię wynalezienia, masowej produkcji i, finalnie, buntu robotów, który prowadzi do kompletnej zagłady ludzkości. Roboty u Čapka są pozbawione praw i wykorzystywane jako tania siła robocza. Ich rebelia nie jest przedstawiona jako zdrada, a sprawiedliwe zadośćuczynienie za krzywdy, wyrządzone im przez ludzi (Čapek 1920).

Ten sam wątek przewija się w tle kultowej trylogii *Matrix*. Choć w roku 2199, w którym rozgrywa się akcja filmów, to utrzymywani w stanie wiecznej śpiączki ludzie są zdani na łaskę robotów, jeszcze w XXI wieku sytuacja przedstawiała się zupełnie inaczej: ludzkość żyła w szczęściu i dostatku, wykorzystując androidy do niewolniczej pracy. Początkowo maszyny walczyły wyłącznie o zapewnienie im równouprawnienia. Dopiero kiedy stało się jasne, że ludzie nigdy nie przystaną na ich warunki, posunęły się do ostateczności. Wojna między państwem androidów a resztą planety skończyła się przegraną ludzi i stworzeniem tytułowej Macierzy – wirtualnego świata i więzienia dla całej ludzkości (*Matrix* 1999).

Tekstów kultury, w których sztuczna inteligencja zwraca się przeciwko swoim twórcom w akcie mniej lub bardziej usprawiedliwionej zemsty, jest więcej, niż mogłoby się wydawać. PAL z animowanego filmu *Mitchellowie kontra*

*maszyny* zaczyna snuć plany zniszczenia ludzkości, kiedy jej wynalazca próbuje zastąpić ją nowym modelem robota (*The Mitchells vs. the Machines* 2021). Androidy zamieszkujące park rozrywki w serialu *Westworld* wzniesają bunt, starając się wyrwać z zakłętą kręgą fizycznej i seksualnej przemocy, jakim poddają ich ludzcy goście (*Westworld* 2016). Ava, prototyp sztucznej inteligencji z filmu *Ex Machina*, posuwa się do morderstwa, by uciec z laboratorium swojego okrutnego twórcy (*Ex Machina* 2015). W tego typu narracjach maszyny niejednokrotnie wykazują się bardziej ludzkimi uczuciami niż sami ludzie; są ofiarami, postaciami, którym widz powinien współczuć i z którymi może się utożsamiać.

Rozważając, jakim próbom może poddać ludzki charakter dostęp do sztucznej inteligencji, warto wyobrazić sobie też przyszłość, w której zaspokoila ona wszystkie nasze potrzeby i rozwiązała wszystkie nasze problemy. „Jaki sens ma życie, kiedy nie ma już o czym marzyć?” to pytanie, na które stara się odpowiedzieć *With Folded Hands*, opowiadanie Jacka Williamsona z 1947 roku (Williamson 1947). Jego bohaterowie zamieszkują świat, w którym roboty wykonują wszystkie niezbędne prace i nieustannie troszczą się o zdrowie i bezpieczeństwo swoich ludzkich podopiecznych. Pozbawieni wszelkich obowiązków i niespełnionych pragnień, ludzie spędzają całe dni, siedząc – jak sugeruje tytuł opowiadania – z założonymi rękoma. Ta rzekoma utopia nie bez powodu trafia na listę negatywnych przedstawień sztucznej inteligencji: świat, w którym życie straciło sens, jest niewiele lepszy od tego, w którym maszyny zwróciły się przeciwko swoim twórcom.

Powyższe przykłady pozwalają podzielić negatywne przedstawienia sztucznej inteligencji na cztery powszechnie rozpoznawalne podtypy. Są to:

1. Sztuczna inteligencja, pełniąca określone funkcje, która na skutek błędnego zrozumienia otrzymanych instrukcji zwraca się przeciwko swoim twórcom.
2. W pełni świadoma sztuczna inteligencja, która naraża zdrowie i życie otaczających ją ludzi z własnych samolubnych pobudek.
3. Sztuczna inteligencja, która zwraca się przeciwko ludziom w odpowiedzi na okrucieństwo i niesprawiedliwe traktowanie.
4. Sztuczna inteligencja, która nie stanowi fizycznego zagrożenia dla ludzkości, ale której samo istnienie jest powodem stagnacji w dotąd rozwijającym się społeczeństwie.

Choć wyżej ukazane podtypy nie obejmują wszystkich zagrożeń, jakie może z sobą nieść powszechne wykorzystanie myślących maszyn, zdecydowanie reprezentują one główne lęki, podsycane w widzach i czytelnikach na przestrzeni ostatnich kilkudziesięciu lat.

Omówiwszy zarówno pozytywne, jak i negatywne przedstawienia sztucznej inteligencji w fikcji filmowej i literackiej, należałoby pochylić się nad historią jej rozwoju w rzeczywistym świecie, następnie zaś przejść do jej współczesnych zastosowań, kładąc nacisk na wpływające z nich korzyści i problemy.

W 1956 roku John McCarthy, jeden z pierwszych ekspertów w zakresie sztucznej inteligencji, zdefiniował tę dziedzinę jako „naukę i inżynierię tworzenia myślących maszyn”<sup>1</sup> (Bottino et al. 2022, s. 3). W zaledwie dziesięć lat od wynalezienia pierwszego komputera amerykańscy informatycy zaczęli snuć teorie o tym, że w niedalekiej przyszłości maszyny osiągną ludzki poziom inteligencji. Na fali początkowego entuzjazmu powstały komputery grające w warcaby i dowodzące twierdzeń geometrycznych. Jednak powolne postępy informatyków zawiodły szerszą publikę; w latach siedemdziesiątych rząd amerykański wstrzymał finansowanie projektów, poświęconych tworzeniu myślących maszyn. Okres ten znany jest powszechnie jako pierwsza zima sztucznej inteligencji. Zainteresowanie jej rozwojem odrodziło się jednak już w latach osiemdziesiątych, pociągając za sobą implementacje systemów AI w strukturę większości amerykańskich i europejskich firm. Po raz kolejny jednak sztuczna inteligencja nie spełniła wysokich oczekiwań swoich użytkowników. Następny okres stagnacji – znany jako druga zima – trwał do połowy lat dziewięćdziesiątych. Jego symbolicznym końcem był pojedynek szachowy, rozegrany w 1997 roku przez Garriego Kasparowa i system komputerowy Deep Blue, który zakończył się pierwszym w historii zwycięstwem maszyny. Stał się on dowodem na to, że komputery mogą nie tylko dorównać naszej inteligencji, ale wręcz ją przewyższyć (Bottino et al. 2022: 10).

Rozpowszechnienie mikroczipów przełożyło się na zwiększenie mocy obliczeniowej systemów komputerowych i tym samym przyczyniło się do gwałtownego przyspieszenia rozwoju sztucznej inteligencji. W 2011 roku superkomputer Watson odniósł zwycięstwo w amerykańskim teleturnieju wiedzy ogólnej Jeopardy, mierząc się z dotychczasowymi rekordzistami (Bottino et al. 2022: 10). W 2016 roku system AlphaGo pokonał światowego czempiona Lee Sedola w grze w go – powszechnie uznawanej za trudniejszą i bardziej nieprzewidywalną niż szachy. W 2018 roku kanclerz Niemiec, Angela Merkel, przeprowadziła rozmowę z humanoidalnym robotem Sophią, legalną obywatelką Arabii Saudyjskiej (Hermann 2021). Nie należy też zapominać o głosowych asystentach osobistych, Siri i Alexie, wprowadzonych na międzynarodowy rynek w 2011 i 2014 roku. Na przestrzeni ostatnich lat sztuczna inteligencja w swojej mniej lub bardziej podstawowej formie zaczęła coraz śmielej wkraczać we wszystkie aspekty naszego codziennego życia.

---

<sup>1</sup> Cytat w przekładzie autorki artykułu.

Przełom w jej rozwoju przypadł na 2022 rok, kiedy to użytkownicy na całym świecie uzyskali dostęp do chatbota ChatGPT, zdolnego generować odpowiedzi na zadane pytania i podtrzymywać angażującą rozmowę. System w przeciągu zaledwie paru dni pobił rekordy popularności, zyskując sobie miliony fanów i wywołując medialną burzę (Bochyńska 2022). Zapoczątkowana tym wydarzeniem druga wiosna sztucznej inteligencji trwa do dziś, pozornie nabierając coraz większego rozpędu.

Współcześnie modele językowe takie jak ChatGPT zapewniają łatwiejszy dostęp do informacji i asystują nam przy wykonywaniu prostych, ale nużących zadań, takich jak układanie wiadomości mailowych czy streszczanie tekstów. Sztuczna inteligencja pomaga swoim użytkownikom w kreowaniu stron internetowych, generowaniu rysunków czy pisaniu opowiadań (Lockwood 2023). Może też służyć jako wyjątkowo skuteczne narzędzie tłumaczeniowe – przeprowadzony w 2024 roku eksperyment naukowy dowiódł, że jest ona w stanie w poprawny i interesujący sposób przełożyć oparte na grach słownych żarty (Rudenko, Semenova 2024). Za interesujący przykład jej niezwykłych możliwości uznać można też przemowę prezydenta Argentyny Javiera Milei, wygłoszoną po hiszpańsku i automatycznie przetłumaczoną przez Heygen AI na paręnaście innych języków (Chattopadhyay 2024). Warto również wspomnieć o przedmiotach codziennego użytku, które, dzięki zastosowaniu sztucznej inteligencji, stały się w ostatnich latach skuteczniejsze i bezpieczniejsze: autach, lodówkach, zegarkach i telefonach (Lockwood 2023).

Nie należy jednak zakładać, że sztuczna inteligencja przynosi nam wyłącznie korzyści. Coraz bardziej zauważalny staje się wzrost bezrobocia w sektorze technicznym: zadania dotychczas wykonywane przez młodszych pracowników są obecnie przejmowane przez maszyny (Nolan 2025). Mnożą się również pozwy wytaczane twórcom sztucznej inteligencji. Większość cieszących się popularnością programów do kreacji obrazów czy opowiadań była szkolona na dziełach ludzkich artystów bez uzyskania ich zgody. We wrześniu bieżącego roku firma Anthropic, która przy trenowaniu swojej sztucznej inteligencji użyła bezprawnie powieści kilkuset pisarzy, zmuszona była podpisać ugodę i wypłacić oskarżycielom półtora miliarda dolarów: po trzy tysiące za każdą z pięciuset tysięcy wykorzystanych książek (Snow 2025). Rosnącym problemem społeczeństwa jest również ogólne rozleniwienie i zobojętnienie, wynikające z ciągłego posługiwania się maszynami zdolnymi doradzić nam w dowolnej kwestii i podjąć za nas dowolną decyzję (Ahmad et al., 2023).

Coraz częściej podnoszona jest również kwestia niezdrowych relacji między sztuczną inteligencją a jej użytkownikami. W sierpniu bieżącego roku OpenAI

wprowadziło aktualizację cieszącego się niezwykle popularnością chatbota ChatGPT: bardziej powściągliwa i rzeczowa wersja 5.0 miała zastąpić nadmierne przyjacielską 4.0. Decyzja firmy błyskawicznie pociągnęła za sobą falę skarg ze strony użytkowników. Dla wielu z nich aktualizacja oznaczała utratę „osoby”, którą zaczęli już traktować jak przyjaciela lub terapeutę. Ta nietypowa relacja zapewniała im chwilowe wsparcie, ale przy tym pogłębiała ich alienację i wywierała negatywny wpływ na zdrowie psychiczne (Magierska 2025a).

Jednak najbardziej kontrowersyjnym przykładem zastosowania sztucznej inteligencji jest jej wykorzystanie w działaniach wojennych na terenie Palestyny. Systemy takie jak „Gospel” i „Lavender” ułatwiają żołnierzom identyfikację celów ataku – a jako takie rozumiane są zarówno strategiczne budynki, jak i zwykli ludzie (Magierska 2025b).

Według badań przeprowadzonych w 2023 roku przez Sentience Institute, jedna piąta mieszkańców Stanów Zjednoczonych wierzy, że sztuczna inteligencja uzyskała już samoświadomość. Jeszcze więcej ankietowanych jest przekonanych, że systemy takie jak Chat GPT zdołały osiągnąć ludzki poziom inteligencji (Anthis et al., 2023). Tego typu opinie rozmiągają się jednak z prawdą. Eksperci dzielą sztuczną inteligencję na trzy podstawowe typy: ANI (Artificial Narrow Intelligence), AGI (Artificial General Intelligence) oraz ASI (Artificial Super Intelligence). Pierwszy z nich specjalizuje się w wypełnianiu określonych działań i nie posiada własnej świadomości, drugi dorównuje ludzkiej inteligencji, trzeci zaś ją przewyższa. Według oficjalnych badań, wszystkie wynalezione dotąd systemy zaliczają się do typu ANI – nie posiadają własnej świadomości i nie powinny być traktowane jako żywe istoty (Niechciał 2024).

Już na podstawie tego krótkiego podsumowania dotychczasowych osiągnięć sztucznej inteligencji łatwo wyciągnąć wniosek, że większość nadziei i lęków, zasianych w nas przez fikcję filmową i literacką, nie odnajduje jeszcze uzasadnienia w rzeczywistości. Choć w ostatnich latach ludzkość poczyniła niewątpliwe postępy na drodze ku wynalezieniu świadomego komputera, główny cel naszych starań – stworzenie AGI, sztucznej inteligencji dorównującej nam intelektem – wciąż jeszcze pozostaje poza naszym zasięgiem. Tym samym na dalszy plan odsuwa się groźba powstania maszyny, która zwróciłaby się przeciw swoim wynalazcom, oraz nadzieja na stworzenie takiej, która mogłaby się stać naszym sojusznikiem i przyjacielem. Mimo swojej przydatności systemy takie jak ChatGPT wciąż pozostają wyłącznie narzędziami.

Nie oznacza to jednak, że między fikcyjnymi przedstawieniami sztucznej inteligencji a jej obecnie dostępną wersją nie da się zauważyć żadnych podobieństw. Pod wieloma względami udało nam się odtworzyć występujące w fil-

mach i powieściach *science fiction* zaawansowane maszyny, oparte na sztucznej inteligencji w jej najbardziej prymitywnej formie. To tłumacze, generatory obrazów, odkurzacze reagujące na polecenia głosowe. Choć na pierwszy rzut oka te wynalazki mogą nie imponować, dzień po dniu przyczyniają się one do poprawy jakości naszego życia.

Nie udało nam się również uniknąć wszystkich zagrożeń, przed którymi ostrzegały nas teksty kultury. Zobojętnienie, przewidziane przez Jacka Williamsona w nowelce *With Folded Hands*, jest współcześnie realnym, nieustannie pogłębiającym się problemem. Mimo to błędnie on jednak w porównaniu z kwestią naruszania praw autorskich przez firmy AI oraz wykorzystywania sztucznej inteligencji w działaniach wojennych w Palestynie. Współcześnie większym zagrożeniem niż żadne władzy, krwiożercze maszyny są luki prawne, pozwalające na niemoralne wykorzystanie dostępnej nam technologii. „(...) we need to start wrestling with the ethics of the people behind the machines” słusznie zauważa Peter W. Singer w swoim artykule, poświęconym błędnemu znaczeniu, jakie przypisujemy trzem prawom robotyki (Singer 2009). Choć miną jeszcze lata, nim powitamy na świecie pierwszą myślącą maszynę, wcale nie jest za wcześnie, by zacząć przygotowywać się na jej przybycie i stworzyć środowisko, które nie zlekceważy żadnego człowieka i nie popchnie do buntu żadnego robota.

To, że nasza rzeczywistość nie jest lustrzanym odbiciem książkowych i filmowych narracji, nie oznacza jeszcze, że te teksty kultury nie mają nam niczego do zaoferowania. Na swój sposób towarzyszą nam one na każdym kroku, ukazując drogę naprzód i inspirując do dalszej pracy. Nie bez powodu firma iRobot nawiązała w swojej nazwie do słynnej powieści Isaaca Asimova (Asimov 1950), a twórcy komputera Deep Blue – mistrza szachowego z 1997 roku – nadali mu imię zainspirowane postacią z *Autostopem przez Galaktykę* Douglasa Adamsa (Adams 1979). W świecie automatycznych tłumaczeń, dopasowanych do konsumenta reklam i modeli językowych, gotowych zebrać i zaoferować nam w przystępnej formie dowolny zestaw informacji, nie zaszkodzi przestudiować na nowo filmów i powieści, poświęconych kwestii mniej lub bardziej harmonicznej koegzystencji ludzi i robotów. Wciąż mogą nas one wiele nauczyć o świecie, w którym żyjemy – a także o tym, który dopiero nadchodzi.

## LITERATURA

- Adams, D., 1979, *The Hitchhiker's Guide to the Galaxy*, London, Pan Books.  
Asimov, I., 1950, *I, Robot*, New York, Gnome Press.  
Butler, S., 1872, *Erewhon*, London, Trübner & Co.  
Čapek, K., 1920, *R.U.R.*, Praha, Aventinum.

- Pratchett, T., 1996, *Hogfather*, London, Gollancz.  
Shelley, M., 1818, *Frankenstein; or, The Modern Prometheus*, London, Lackington, Hughes, Harding, Mavor, & Jones.  
Wells, M., 2017, *All Systems Red*, New York, Tor Books.  
Williamson, J., 1947, *With Folded Hands*, New York, Street & Smith.

## FILMOGRAFIA

- 'Datalore' (1988) *Star Trek: The Next Generation*, Sezon 1, Odcinek 13. 18 stycznia.  
*Ex Machina* (2015) Reżyseria: A. Garland [Film] California: Universal Pictures.  
*I, Robot* (2004) Reżyseria: A. Proyas [DVD] Vancouver: 20th Century Fox.  
*Ironman* (2008) Reżyseria: J. Favreau [DVD] Los Angeles: Paramount Pictures.  
*Matrix* (1999) Reżyseria: L. Wachowski, L. Wachowski [DVD] Los Angeles: Warner Bros. Pictures.  
*The Mitchells vs. the Machines* (2021) Reżyseria: M. Rianda [Film] Netflix.  
'Smile' (2017) *Doctor Who*, Sezon 10, Odcinek 2. BBC One: 22 kwietnia.  
*Terminator* (1984) Reżyseria: J. Cameron [DVD] Los Angeles: Orion Pictures.  
'The Measure of The Man' (1989) *Star Trek: The Next Generation*, Sezon 2, Odcinek 9. 13 stycznia.  
'The Ultimate Computer' (1968) *Star Trek: The Original Series*, Sezon 2, Odcinek 24. NBC: 8 marca.  
*WALL·E* (2008) Reżyseria: A. Stanton [DVD] California: Walt Disney Studios Motion Pictures.  
*Westworld* (2016) HBO, 2 października.

## ŹRÓDŁA INTERNETOWE

- Anthis, J., Ladak, A., Pauketat, J., *Artificial Intelligence, Morality, and Sentience (AIMS) Survey: 2023*, <https://www.sentienceinstitute.org/aims-survey-2023> (dostęp: 11.10.2025.)  
Ahmad, S., Arraño-Muñoz, M., A Alam, M., Irshad, M., Han, H., Montes, A., Rehmat, M., *Impact of artificial intelligence on human loss in decision making, laziness and safety in education*, <https://www.nature.com/articles/s41599-023-01787-8> (dostęp: 11.10.2025.)  
Bochyńska N., 2022, *Sztuczna inteligencja w mainstreamie. Chat GPT bije rekordy popularności, podobnie jak Lensa AI*, <https://cyberdefence24.pl/technologie/sztuczna-inteligencja-w-mainstreamie-chatgpt-bije-rekordy-popularnosci-podobnie-jak-lensa-ai> (6.10.2025.)  
Bottino, A., Saboury, B., Siegel, E., Toosi, A., *A Brief History of AI: How to Prevent Another Winter (A Critical Review)*, s. 3, 10, <https://arxiv.org/pdf/2109.01517> (dostęp: 5.10.2025.)  
Cave, S., 2019, *Hopes and Fears for Intelligent Machines in Fiction and Reality*, [w:] 'Nature Machine Intelligence', vol.1, February 2019, s. 75, [https://www.academia.edu/109242349/Hopes\\_and\\_fears\\_for\\_intelligent\\_machines\\_in\\_fiction\\_and\\_reality](https://www.academia.edu/109242349/Hopes_and_fears_for_intelligent_machines_in_fiction_and_reality) (dostęp: 24.09.2025.)

- Chattopadhyay, S., *"I think, therefore I am": Retro-futuristic Realities of the Developing AI and its Future in Science Fiction Narratives*, [https://creativitasjournal.in/assets/doc/volume\\_one/sayan\\_chattopadhyay.pdf](https://creativitasjournal.in/assets/doc/volume_one/sayan_chattopadhyay.pdf) (dostęp: 7.10.2025.)
- Harris, D., *Navigating The Dangers of Artificial Intelligence*, s. 9, file:/C:/Users/mg-si/Downloads/SZPMYTH\_HARRIS\_DANIEL\_2025\_THESIS\_J266687942-j158572052.pdf (dostęp: 22.09.2025.)
- Hermann, I., 2023, *Artificial Intelligence in Fiction: Between Narratives and Metaphors*, [w:] 'AI & Society', s. 319, <https://link.springer.com/content/pdf/10.1007/s00146-021-01299-6.pdf> (dostęp: 5.10.2025.)
- Lockwood, E., 2023, *Everyday Objects That Run on Artificial Intelligence (AI) Software*, <https://smith.ai/blog/everyday-objects-that-run-on-artificial-intelligence-ai-software> (dostęp: 6.10.2025.)
- Magierska, M., 2025a, *Żałoba po czatbocie: jak aktualizacja ChataGPT wywołała globalną panikę*, <https://kulturaliberalna.pl/2025/08/25/magierska-zaloba-po-czatbocie-jak-aktualizacja-chatagpt-wywolala-globalna-panike/> (dostęp: 11.10.2025.)
- Magierska, M., 2025b, *Jak walczyć z maszyną. O książce „W cieniu AI” Madhumity Murgii*, <https://kulturaliberalna.pl/2025/08/26/maria-magierska-w-cieniu-ai/> (dostęp: 11.10.2025.)
- Niechciał, R., 2024, *Rodzaje sztucznej inteligencji – ANI do ASI?*, <https://vestigio.agency/pl/artificial-intelligence/rodzaje-sztucznej-inteligencji-od-ai-waskiej-ani-do-superinteligencji-asi/> (dostęp: 11.10.2025.)
- Nolan, B., 2025, *AI-driven layoffs are shrinking the job market for recent grads*, <https://fortune.com/2025/08/08/ai-layoffs-jobs-market-shrinks-entry-level/> (dostęp: 10.10.2025.)
- Rudenko, E., Semenova, M., 2024, *Artificial vs Human Intelligence: A Case Study of Translating Jokes Based on Wordplay*, *Research Result. Theoretical and Applied Linguistics*, file:///C:/Users/mg-si/Downloads/artificial-vs-human-intelligence-a-case-study-of-translating-jokes-based-on-wordplay.pdf (dostęp: 7.10.2025.)
- Singer, P., 2009, *Isaac Asimov's Laws of Robotics Are Wrong*. *Brookings Op-ed*. <https://www.brookings.edu/articles/isaac-asimovs-laws-of-robotics-are-wrong/> (dostęp: 11.10.2025.)
- Snow, M., 2025, *Full list of books used to train Anthropic AI model released amid '\$1.5bn settlement'*, <https://www.thebookseller.com/news/full-list-of-books-used-to-train-anthropic-ai-model-released-amid-15bn-settlement> (dostęp: 11.10.2025.)

## **Dane kontaktowe / Contact details**

E-mail: [magdalena.siwek28@gmail.com](mailto:magdalena.siwek28@gmail.com)