

Sieci neuronowe typu GAN i GPT-2, słowa zużyte i kreatywność, czyli literacki *second-hand*

Inez Okulska

ORCID 0000-0002-1452-9840

„Wszyscy piszą, tylko nikt nie czyta”, „jak chcesz dobrze pisać, najpierw musisz poznać klasykę, zobaczyć, co stworzyli inni” – te lub podobne wypowiedzi towarzyszą entuzjastom pisarstwa od dawna. I kiedy mówię „dawna”, mam na myśli naprawdę długie wieki, bo kiedy jeszcze przekład literacki nie wyemancypował się do osobnego gatunku, służył nierzadko za podstawę twórczej działalności. Klasykami można nazwać choćby Jana Kochanowskiego, Łukasza Górnickiego czy Adama Mickiewicza, którzy natchnieni lekturą innych siadali tworzyć swoje utwory, mniej lub bardziej dosłownie czerpiąc z obcych, już istniejących dzieł¹.

Jak się ma więc do tego kreatywność?

Bez wątpienia poznajemy istniejące dzieła, by się zainspirować, to znaczy nasiąknąć wzorcami, gotowymi elementami, choćby najdrobniejszej miary, i z nich układać nowe. Poznajemy istniejące dzieła, by w euforii nie odkryć dawno odkrytego. I wreszcie poznajemy istniejące dzieła w konkretnym kontekście krytycznym, umiejscowione w trendach i prądach po to, by nabrać smaku i odróżnić dzieło od gnioła, kreatywność od kreatury. W tego typu rozważaniach zawsze pojawia się pełne nadziei pytanie o możliwość stworzenia czegoś absolutnie nowego, o definicje i granice owej nowości oraz, siłą rzeczy, o relacje z innymi dziełami.

W dobie sztuki rozszerzonej o nowe i jeszcze nowsze media intertekstualność Kristewej zastępuje dziś interpoetyka, gdzie cytatem może być nie tylko tekst, lecz także struktura, mechanizm, algorytm. Ale nawet na poziomie „zwykłego” ludzkiego pisarstwa, którego tworzywem jest język, jego słowa, frazy, składnie, Bachtin rozwiął wszelką nadzieję: niemal każde słowo w dziele jest swoją własną inkarnacją, sumą kontekstów i znaczeń, jakie zyskało wcześniej, w innych dziełach, na innych ustach. Arcydzieło literatury jest więc najczęściej wybitną mieszanką elementów starych, by nie rzec wręcz „wyświechtanych”:

¹ O czym pisało już wielu, ale ostatnio m.in. Małgorzata Łukasiewicz w swojej książce. Małgorzata Łukasiewicz, *Pięć razy o przekładzie* (Kraków: Karakter, 2018).

Każda wypowiedź związana jest z „jednym językiem” (z siłami i tendencjami dośrodkowymi) i jednocześnie ze społeczną i historyczną różnorodnością (siłami odśrodkowymi, rozwarstwiającymi język). Jest to język dnia, epoki, grupy społecznej, gatunku, kierunku itd. [...] Wszelkie konkretne słowo (wypowiedź) znajduje przedmiot, ku któremu jest skierowane, zawsze, by tak rzec, już omówiony, przedyskutowany, oceniony, ukryty w oddalającej go mgle lub, przeciwnie, przybliżony światłem wcześniej wypowiedzianych o nim słów. Przedmiot ten oplątały i przeniknęły wspólne myśli, poglądy, cudze oceny, akcenty. Słowo zwrócone ku owemu przedmiotowi włącza się w to dialogowo zaktywizowane i poruszone środowisko cudzych słów, ocen i akcentów, wplątuje się w ich skomplikowane stosunki wzajemne, upodabnia się do jednych, kontrastuje z innymi, z innymi jeszcze – przecina².

Jeśli materia, z której mamy tworzyć, jest ograniczona i pozostaje wciąż w użyciu, wówczas okazuje się, że „kreatywność nie jest czynnością absolutną, lecz względną. Bywamy kreatywni wewnątrz naszej kultury i ramy odniesień”³ – jak pisał w swoich rozważaniach o cyfrowej sztuce znany matematyk Marcus de Sautoy. Na ograniczenie kreatywności związane z zanurzeniem w konkretnym zestawie kulturowych odniesień zwracał już w latach 70. uwagę Stanley Fish już w latach siedemdziesiątych, mówiąc o wspólnotach interpretacyjnych⁴, które można objaśnić następująco:

jesteśmy tym, czym nas uczyni szkoła, grupa rówieśnicza, lokalne autorytety, podsunięte nam lektury, poglądy cieszące się uznaniem wśród wpływowych przedstawicieli naszego zawodu, itp. Zatem protokoły naszych indywidualnych strategii interpretacyjnych zawsze pisane są ręką wspólnoty interpretacyjnej, z którą się identyfikujemy⁵.

Mowa tu wprawdzie wyłącznie o interpretacji, a więc o odbiorze, ale to samo tyczy się procesu kreacji, bo tworząc, nieustannie odnosimy się do już nauczonych (i oczekiwanych) poetyk. W zasadzie kreatywna jest głównie relacja, stanowisko, jakie się wobec tych poetyk przyjmuje – aprobata, kontra, redukcja, kontaminacja, transformacja etc., czyli właśnie te wszystkie „skomplikowane stosunki wzajemne”, o których pisał Bachtin.

Sieci neuronowe typu GAN (Generative Adversarial Networks, generatywna sieć przeciwstawna)⁶ to idealna cyfrowa wykładnia kreatywności rozumianej jako twórcze poruszanie się wyłącznie po konkretnym obszarze wzorców i oczekiwań. Architektura tego typu sieci zakłada istnienie twórczego duetu, z którego jedna sieć niczym wewnętrzne dziecko, wolne od jakiegokolwiek ramy predefiniowanych założeń, podejmuje się tworzenia. Akt ten oznacza kilkustopniowe filtrowanie i układanie przypadkowych elementów (elementów przypadkowego wektora z), w efekcie których powstaje dzieło. Swoją drogą ciekawe, co na tak czystą formę powiedziałby sam Witkacy? Niestety, czystość ta jest jedynie chwilowa i dla odbiorcy ostatecznego efektu działania sieci niedostępna, bo, jak już wspomniałam, GAN to duet. Generator tworzy, ale Dyskryminator

² Michaił Bachtin, *Problemy literatury i estetyki*, tłum. Wincenty Grajewski (Warszawa: Czytelnik, 1982): 102–103.

³ „Creativity is not an absolute but a relative activity. We are creative within our culture and frame of reference”. Marcus du Sautoy, *The creativity code. How AI is learning to write, paint and think* (London: Fourth Estate, 2019), 13.

⁴ Stanley Fish, *Interpretacja, retoryka, polityka*, red., przedm. Andrzej Szahaj, wstęp Richard Rorty, tłum. A. Szahaj (Kraków: Universitas 2002): 63.

⁵ Leszek Drong, „Od konwencjonalizmu do normatywizmu. Kilka uwag o ewolucji poglądów teoretyczno-literackich Stanleya Fisha”, *Er(r)go: Teoria – Literatura – Kultura*, vol. 12, nr 1 (2006): 25–37.

⁶ Więcej o architekturze tych sieci można przeczytać w: Rohith Gandhi, “Generative Adversarial Networks – Explained”, <https://towardsdatascience.com/generative-adversarial-networks-explained-34472718707a> (dostęp: 8.05.2019).

krytykuje, i to z mocą, o której niejeden krytyk w afekcie mógłby pomarzyć – informacja zwrotna od bliźniaczego cenzora ma natychmiastowy wpływ na dynamiczny proces doskonalenia pierwotnego dzieła. Procedura ta powtarzana jest tak długo, aż krytyk będzie kontent.

Takie ujęcie procesu twórczego zahacza aż o dwie newralgiczne kwestie: po pierwsze, twórca jest niewolnikiem swojego krytyka. Po drugie, krytyk (a ostatecznie więc również twórca) jest niewolnikiem istniejących, oczekiwanych schematów, ponieważ cenzorska sieć swoją ekspertyzę opiera na „wiedzy” zdobytej ze zbioru uczącego, czyli milionów próbek istniejących „dzieł” (mogą to być teksty, zdjęcia lub muzyka czy wideo, w zależności od tego, w czym sieć ma się specjalizować). A nieustająca walka, jaka toczy się pomiędzy obiema sieciami polega na tym, że Generator próbuje stworzyć coś nowego, co będzie jednak na tyle przekonujące, że zdoła oszukać krytyka, który uzna je za „rzeczywiste”, czyli wysoce prawdopodobne. Dyskryminator *vel* inżynier Mamoń akceptuje więc wyłącznie te dzieła, które wpisują się w znane mu wzorce. Czyli dzieła, które są niczym innym jak centonem, utworem złożonym w całości z kreatywnie dobranych i ze sobą połączonych cytatów, elementów, które już gdzieś kiedyś istniały. Rozdzielczość jest tu wprawdzie nieco inna, bo cechy wyróżnione w wektorze z (te „cytaty” z nauczonych dzieł) są dla człowieka niemal niedostępne (logika podziału i wyodrębniania odbiega od naturalnej ludzkiej kategoryzacji na słowa, przedmioty, kolory czy kształty), ale założenie strukturalne jest takie samo.

Proces kreacji oparty wyłącznie na coraz to nowym zestawianiu elementów już istniejących i odwoływaniu się wyłącznie do ograniczonego zbioru oczekiwanych wyników może być oczywiście na dłuższą metę raczej fatalny w skutkach ze względu na niemal wykładniczy rozwój „chowu wsobnego”. Jeśli sieci służące do generowania, redakcji (takie jak zewnętrzne Grammarly, wbudowane funkcje w edytorach tekstu) i tłumaczenia tekstów są uczone na pewnych korpusach tekstów spełniających normy poprawności i określonego stylu, oraz są coraz chętniej używane przez ludzkich autorów, mogą przyczynić się do tworzenia coraz większej liczby „znormalizowanych” stylistycznie tekstów. Ich publikacja będzie natomiast zwiększała prawdopodobieństwa znalezienia się ich w korpusie danych uczących dla kolejnych sieci, i tak krok po kroku procent uczestnictwa owej normy w puli inspiracji będzie rósł, przekładając się na zubożenie zakresu stylistycznej czy leksykalnej różnorodności, tworząc upiorną wizję zapętlenia Bachtinowskiej dialogiczności⁷.

Na szczęście jednak generacja tekstów oparta na tym mechanizmie nie jest jedyną formą tekstowej ekspresji ani maszyny, ani tym bardziej człowieka. Kognitywistka Margaret Boden wyróżniła trzy typy kreatywności⁸:

1. zgłębiająca (*exploratory*), umożliwiająca badanie istniejących u(tworów) w poszukiwaniu wariantów granicznych, które pozwolą jednocześnie nie złamać przyjętych zasad;
2. zestawiająca (*combinatory*), pozwalająca łączyć ze sobą elementy dotychczas uznawane za absolutnie do siebie nieprzystające;

⁷ Co trochę nawiązuje m.in. do modernistycznej idei „języka literackiego” jako poprawnego, normatywnego, momentami przezroczystego w obawie przed sięgnięciem bruku. Taka postawa widoczna była w długiej tradycji przekładów podnoszących rejestr i wygładzających oryginalną chropowatość dzieł klasycznych, które wchodziły do obiegu inspiracji w takiej niepotrzebnie zubożonej, znormalizowanej formie.

⁸ Margaret A. Boden, *The Creative Mind: Myths and Mechanisms* (London: Routledge, 2004): xxx.

3. przeobrażająca (*transformational*), czyli taka, której efektem są prawdziwe przełomy i kiedy zmianie ulegają elementy uważane wcześniej za niepodważalne – materiał, narzędzie, styl odbioru.

Ten ostatni typ zdaje się mieszanką zestawienia i eksploracji przy jednoczesnym przekroczeniu granic i złamaniu zasad. Klucz do sukcesu twórczego zasilanego tego typu kreatywnym postępowaniem to jednak, jak podkreśla du Sautoy⁹, przygotowanie na porażkę, bo prawdziwy przełom jest najczęściej ostatnim etapem długiej ścieżki kompletnie chybionych pomysłów. Emocjonalny związek z dziełem, niewymierny koszt procesu twórczego u człowieka, nierzadko jednak umożliwia takie zimne, niemal wyrachowane podejście do własnej porażki, każące ją traktować wyłącznie jako źródło bezpośredniej informacji o kierunku, który należy obrać w kolejnej iteracji. W takim podejściu specjalizują się natomiast adaptacyjne systemy sztucznej inteligencji.

Ponieważ uczenie sieci neuronowej na potrzeby przetwarzania języka naturalnego to najczęściej niezwykle czasochłonny proces o ogromnej złożoności obliczeniowej (nieradko niewykonalny na zwykłym komputerze osobistym), ale także ze względu na bezpośrednie korzyści wynikające z poszerzonego zbioru uczącego, w sieci dostępne są modele wstępnie wytrenowane (tzw. *pre-trained models*). Pod intrygującymi nazwami (m.in. Transformer, BERT, ELMo czy Flair) kryją się modele, którym już stworzono odpowiednią architekturę, dobrano parametry i podano ogromne korpusy tekstu (np. aktualny zrzut z całej dostępnej Wikipedii). Niektóre z nich zostały zaprojektowane z myślą o konkretnym zadaniu (np. analiza sentymentu tekstu, czyli określenie czy wypowiedź ma wydźwięk pozytywny, neutralny czy negatywny), inne, takie jak hybryda Google’a BERT, są wielozadaniowe.

Szczególnie w tym ostatnim przypadku ciekawe jest to, że zbiory uczące przeznaczone do rozwiązywania jednego problemu mogą być wykorzystane do innego, a pamiętając, że problem podpowiadania słów, tłumaczenia czy generowania całych tekstów wykorzystuje wiedzę na temat kontekstu i możliwych połączeń wydestylowanych z milionów próbek istniejących tekstów, okaże się, że cytaty z jednego typu tekstu są z powodzeniem wykorzystywane do przetwarzania innego typu tekstu. Czyli, idąc tropem metafory *stricte* literackiej, to tak, jak autor piszący fragment prozy, wykorzystujący słowa, wyrażenia, relacje i struktury wywiedzione z poezji, dramatu czy innych gatunków – co oczywiście jest naturalnym mechanizmem kreacji we współczesnej literaturze, w której mówi się o „nasłuchiwanie” idiolektów, mowy „ulicy”, o nawiązaniach do poetyki nowych mediów i tym podobnych. Niejeden autor i jego krytyk intuicyjnie już dawno więc wyczuli, że taka interpoetyka, otwarty zapraszający gest twórcy, przybliżają nas do dobrej literatury, obiecują dobrą jakość tekstu. A sztuczna inteligencja swoim działaniem tylko dowiodła wymiernych zalet owego otwarcia na inne – modele trenowane na niespecyficznych zestawach danych (tzn. tematycznie czy stylistycznie odpowiadających danemu zadaniu) generują widocznie lepsze wyniki. Transfer struktur odbywać się może zresztą nie tylko na poziomie tekstu, na którym się uczy tworzyć, ale też na poziomie architektury samej sieci – zapożyczany jest więc nie tylko materiał¹⁰.

⁹ Du Sautoy, *The creativity code*, 39.

¹⁰O zastosowaniu rozwiązań uczenia maszynowego z obszaru analizy obrazu do automatycznej analizy tekstu pisali: Marius Pompescu i Radu Tudor Ionescu, *Knowledge Transfer between Computer Vision and Text Mining*, Bucharest: Springer, 2015, a o kwestii transferu metod wewnątrz różnych zadań przetwarzania języka naturalnego Lili Mou i in., „How Transferable are neural networks in NLP applications”, *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (2016): 478–489.

Model GPT-2 opracowany przez OpenAI miał zachwycać jakością swoich poczynań, które koncentrowały się głównie na generowaniu tekstu będącego kontynuacją zadanego fragmentu (*text completion*), tłumaczeniu maszynowym oraz czytaniu ze zrozumieniem kontekstu (*question answering*). To ostatnie zadanie polega głównie na parsowaniu tekstu, ekstrakcji informacji oraz wnioskowaniu i dla naszych rozważań nie będzie kluczowe, natomiast zarówno generowanie tekstu, jak i generowanie tłumaczeń to zadania oparte przede wszystkim na inteligentnym i twórczym żonglowaniu miliardami cytatów, na ich fragmentowaniu i kreatywnej rekompozycji.

GPT-2 musi najpierw zidentyfikować poetykę zadanego fragmentu, po czym spośród „prze czytanych” (nauczonych wcześniej) tekstów znaleźć możliwie ekwiwalentne poetyki, z których skompiluje nowy utwór odpowiadający stylem i treścią pierwszym akapitom napisanym przez człowieka. A model ten, trzeba przyznać, jest niezwykle „oczytany”, bowiem korzysta z aż 40 GB danych tekstowych pochodzących łącznie z 45 miliardów stron (i podstron) opublikowanych w Internecie. I mowa tu o wszelkich gatunkach tekstowych, od wpisów na blogach, artykułów na portalach przez komentarze, przepisy kulinarne, opinie konsumentów czy instrukcje obsługi – choć z wyjątkiem Wikipedii, która w całości została ze zbioru usunięta (jako zbyt chętnie używana przez inne modele). Działanie pełnego modelu (którego twórcy strzegą, bo – jak się buńczucznie chwala – jest tak dobry, że mógłby zostać wykorzystany do niepożądanych celów, jak choćby tworzenie niezwykle przekonujących fake newsów) naprawdę robi wrażenie.

Rzeczywisty wpis¹¹ zamieszczony na profilu GitHub, społecznościowej platformy wymiany plików i projektów oraz publikacji kodu, tutaj pochodzący z pliku pomocy komentującego zamieszczoną konfigurację środowiska programistycznego, został wczytany do modelu. Zadanie polegało na możliwie realistycznym dokończeniu tego tekstu. Poniżej prezentuję efekt działania algorytmu (na niebiesko tekst oryginalny stworzony przez człowieka, na czarno dopisany przez sztuczną inteligencję):

In Cygwin on Windows I found that I could not set the permissions of my `/.ssh/` folder to be 0600, as is required for ssh to allow you to use keys. The symptom I had was that I no matter what I did, it always modified the owner and the group, making these files read-only. I tried just starting ssh and using a program like `passwd` which modifies the owner/group to 0644 permissions, but it didn't work. It was annoying but not nearly as bad as it could be, just annoying to have to fiddle with permissions. So I decided to do this as the reverse: For each `/.ssh` folder in your Cygwin directory, set the permission of each file and directory to 2644. Then start ssh with: `ssh-add /.ssh`¹².

Twórcy modelu chwala się, że jest on w stanie kontynuować tekst w wybranym stylu i rzeczywiście, gdyby nie kolor tekstu, całość spokojnie zdałaby test Turinga, a miejsce spojenia zupełnie umknęło ludzkiemu oku i uchu. Algorytm bez trudu, jak widać „wczuł się” w autora, szafując dalej jego osobowym ‘ja’; sięgnął również po słowa „annoying” oraz „nearly”, które pojawiają się w oryginalnym dokumencie i prawdopodobnie uzyskały priorytetowe wagi jako godne wykorzystania markery stylu autora. Do stworzenia tego tekstu jednak nie wystarczyła pojedyncza

¹¹<https://github.com/lelandbatey/configDebDev/blob/master/helpFiles.md>.

¹² Alec Radford i in., „Language Models are Unsupervised Multitask Learners”, raport techniczny, <https://d4mucfpksywv.cloudfront.net/better-language-models/language-models.pdf> (dostęp: 20.07.2019).

próbka danego pisarstwa ani też wiele dzieł danego autora, lecz na ów syntetyczny utwór złożyły się setki tysięcy wpisów innych osób, najpewniej również pochodzące z platformy GitHub oraz absolutnie niezbędnej kopalni wiedzy w trybie „samopomocy programistycznej”, czyli StackOverflow. To platforma, na której praktycy odpowiadają sobie nawzajem na pytania o konkretne problemy, nierzadko opisując własne doświadczenia w sposób nieformalny i emocjonalnie nacechowany (w końcu mowa o tych chwilach, gdy po wielu godzinach prób program się zaczyna i klawiatury fruną na skrzydłach frustracji).

GPT-2 świetnie też radzi sobie z fikcją, generując tekst, którego styl i pomysłowość mogłyby zawstydzić nie tylko grafomanów:

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English. The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science. [...] Dr. Jorge Perez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. [...] Dr. Perez believes that the unicorns may have originated in Argentina, where the animals were believed to be descendants of a lost race of people who lived there before the arrival of humans in those parts of South America. [...] However, Perez also pointed out that it is likely that the only way of knowing for sure if unicorns are indeed the descendants of a lost alien race is through DNA. "But they seem to be able to communicate in English quite well, which I believe is a sign of evolution, or at least a change in social organization," said the scientist¹³.

Jest mowa zależna, niezależna, są bohaterowie, są relacje między nimi, autorytety i trudne słowa, jest magia i tajemnica, słowem świat przedstawiony barwny niczym tęcza w krainie jednoroźców. W ostatnim zdaniu pojawia się również niezły substytut humoru: konstrukcyjnie przekonująca puenta, oparta na zaskakującym, nieoczywistym (a więc kreatywnym) zestawieniu rozwinięcia mowy u jednoroźców ze społeczną organizacją, które to zestawienie zostaje przemycone pod płaszczykiem przyczynowo-skutkowego wniosku, jakże pasującego do eksperckiego cytatu archetypicznego naukowca. Sieć najwyraźniej przebiła jednoroźce, bo nie tylko przemówiła perfekcyjną angielszczyzną, lecz również brytyjskim humorem.

Ten sam model został również użyty do wygenerowania tekstu mającego być kontynuacją pierwszego zdania z powieści Orwella *Rok 1984*¹⁴. Sieci miało ono wystarczyć do „wycucia” subtelnie futurystycznego stylu:

It was a bright cold day in April, and the clocks were striking thirteen.

I was in my car on my way to a new job in Seattle. I put the gas in, put the key in, and then I let it run. I just imagined what the day would be like. A hundred years from now. In 2045, I was a teacher in some school in a poor part of rural China. I started with Chinese history and history of science.

¹³Radford.

¹⁴„New AI fake text generator may be too dangerous to release, say creators”, *The Guardian*, artykuł on-line z 14.02.2019, <https://www.theguardian.com/technology/2019/feb/14/elon-musk-backed-ai-writes-convincing-news-fiction> (dostęp: 20.07.2019).

Ivana Kajić, doktorantka w dziedzinie informatyki z Uniwersytetu Waterloo w Ontario, postanowiła natomiast przetestować oficjalnie dostępny, uproszczony model GPT-2 na tekstach piosenek¹⁵. Podała sieci między innymi dwie pierwsze zwrotki utworu Eda Sheerana *Shape of you*:

The club isn't the best place to find a lover
 So the bar is where I go
 Me and my friends at the table doing shots
 Drinking fast, and then we talk slow

And you come over and start up a conversation with just me
 And trust me, I'll give it a chance now
 Take my hand, stop
 Put "Van the Man" on the jukebox
 And then we start to dance
 And now I'm singing like

W zarysowanej scenie pojawia się zapowiedź odważnej interakcji damsko-męskiej, ale wciąż jeszcze jest szafa grająca i chęć tańca, spotkanie ciał będzie niedosłowne, z nutą irlandzkiej pacytyny. W dalszym fragmencie oryginału, owszem, już pojawiają się coraz bardziej bezpośrednie obrazy, ale w ramie asonansów i metafory. Fragment dopisany przez sieć zdaje się natomiast wprowadzać zupełnie nową poetykę, krótszej, bardziej rytmicznej frazy, z silniejszym podmiotem, i choć na próżno szukać tu rymów, to siła ekspresji (łącznie z zaskakującym „you better get dressed”, sugerującym, że sytuacja była bardziej rozwojowa niż mogliśmy początkowo przypuszczać) bliższa jest tekstom utworów spod znaku hip-hopu:

You know what I mean
 That's the way it is
 So please, I know it's weird
 But I know it's right
 Because I just tell it
 And I say "Hi, nice to meet you
 You better get dressed"

Tekstowe inspiracje wynikające ze wspomnianego szerokiego, niespecjalizowanego „oczytania” sieci znalazły zatem widoczne odzwierciedlenie w utworze. Model ten jest okrojoną wersją swojego genialnego brata i owa redukcja manifestuje się najwyraźniej w większej podatności na wzorce nauczone wcześniej przy jednocześnie mniejszej zdolności „wczucia się” w styl zadanego fragmentu, co ostatecznie owocuje jednak ciekawym mieszanym poetyk, owym kreatywnym eksplorowaniem możliwości granicznych (na krok przed wściekłym snem zielonych idei Chomsky’ego).

¹⁵Ivana Kajić, „Alternate endings: Lyrics completion using GPT-2”, <http://www.ivanakajic.me/blog/2019/03/31/alternate-lyrics> (dostęp: 15.07.2019).

Podejście generatywne, żonglowanie cytatami, niepoahamowany styl korzystania z rozmaitych leksykonów – odważni przedstawiciele OuLiPo, czyli francuskiej szkoły literatury automatycznej z lat sześćdziesiątych, najwyraźniej doczekali się całkiem godnych potomków, choć trudno raczej będzie opisać perypetie życiowe maszyn, na próżno szukać u nich porywów serca, starzanych nerwów, złych wyborów, szczęścia i tego wszystkiego, co składa się na proces twórczy.

A prawdziwe dzieło cieszy tak naprawdę potencjalnością tego procesu. Nie wiem, co czuł artysta, kiedy tworzył, ale jak miło percypując jego dzieło założyć, że w ogóle mógł coś czuć. Sztuka stworzona przez sztuczną inteligencję wciąż jeszcze jawi się więc jako „sztuczna”, bo jest z tej potencjalności odarta, słowem – zupełnie nieludzka.

Bibliografia

- Bachtin, Michaił. *Problemy literatury i estetyki*. Przetłumaczone przez Wincenty Grajewski. Warszawa: Czytelnik, 1982.
- Boden, Margaret A. *The Creative Mind: Myths and Mechanisms*. London: Routledge, 2004.
- Drong, Leszek. „Od konwencjonalizmu do normatywizmu. Kilka uwag o ewolucji poglądów teoretyczno-literackich Stanleya Fisha”. W *Er(r)go: Teoria – Literatura – Kultura*, vol. 12, nr 1 (2006): 25–37.
- Fish, Stanley. *Interpretacja, retoryka, polityka*. Redakcja i przedmowa Andrzej Szahaj. Wstęp Richard Rorty. Przetłumaczone przez Andrzej Szahaj. Kraków: Universitas, 2002.
- Gandhi, Rohith. *Generative Adversarial Networks – Explained*, <https://towardsdatascience.com/generative-adversarial-networks-explained-34472718707a>, (dostęp: 8.05.2019).
- Kajić, Ivana. *AIternate endings: Lyrics completion using GPT-2*, <http://www.ivanakajic.me/blog/2019/03/31/aiternate-lyrics> (dostęp: 15.07.2019).
- Kazimierska, Agata. „Prawdziwe kłamstwa”. *Tygodnik Powszechny*. Artykuł on-line z 15.04.2019, <https://www.tygodnikpowszechny.pl/prawdziwe-klamstwa-158508> (dostęp: 20.07.2019).
- Łukasiewicz, Małgorzata. *Pięć razy o przekładzie*. Kraków: Karakter, 2017.
- Mou, Lili i in. “How Transferable are neural networks in NLP applications”. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (2016): 478–89.
- “New AI fake text generator may be too dangerous to release, say creators”. *The Guardian*. Artykuł on-line z 14.02.2019, <https://www.theguardian.com/technology/2019/feb/14/elon-musk-backed-ai-writes-convincing-news-fiction> (dostęp: 20.07.2019).
- Pompeanu, Marius i Radu Tudor Ionescu, *Knowledge Transfer between Computer Vision and Text Mining*, Bucharest: Springer, 2015.
- Radford, Alec i in. *Language Models are Unsupervised Multitask Learners*. Raport techniczny, <https://d4mucfpksywv.cloudfront.net/better-language-models/language-models.pdf> (dostęp: 20.07.2019).
- du Sautoy, Marcus. *The creativity code. How AI is learning to write, paint and think*. London: Fourth Estate, 2019.

SŁOWA KLUCZOWE:

krytyka literacka

SZTUCZNA INTELIGENCJA

machine learning

ABSTRAKT:

Czy kreatywność to wyłącznie domena człowieka? Czy sieć neuronowa, choćby najbardziej skomplikowanej architektury, nakarmiona materiałem stworzonym i wybranym przez człowieka może być kreatywna, a jeśli nawet, to czy jej dzieło nie będzie wobec ludzkiego wtórne? A może, jak chciał Bachtin, a za nim Kristeva, każda nasza wypowiedź i tak jest skazana na wtórność, bo taka jest natura języka? Czym jest kreatywność, co potrafi sztuczna inteligencja, do jakich refleksji krytycznoliterackich skłaniać może jej twórczość, szczególnie w kontekście relacji intertekstualnych, interpoetyckich? W artykule odpowiedzi szukam na przykładzie funkcjonowania sieci neuronowych typu GAN oraz modelu GPT-2. Oprócz fragmentów analizowanych tekstów i nawiązań do teorii literatury pojawia się również wprowadzenie do struktury i istoty omawianych rozwiązań technologicznych.

KREATYWNOŚĆ

creative writing

INTERPOETYKA

INTERTEKSTUALNOŚĆ

NOTA O AUTORZE:

Inez Okulska – doktor nauk humanistycznych w zakresie literaturoznawstwa. Po przejściu barwnej ścieżki humanistycznej (na którą składały się m.in. lingwistyka, komparatystyka literacka, kulturoznawstwo, filozofia), zakończonej podoktorskim stażem na Harvard University, podjęła studia magisterskie na kierunku automatyka i robotyka na Politechnice Warszawskiej. Obecnie doktorantka w Szkole Doktorskiej TIB Polskiej Akademii Nauk w zakresie informatyki technicznej. Metody sztucznej inteligencji, a w szczególności metody przetwarzania języka naturalnego, którymi obecnie zajmuje się naukowo w Państwowym Instytucie Badawczym NASK, doskonale łączą te odległe, zdawałoby się, dziedziny, zwłaszcza że metody te najchętniej aplikuje właśnie od analizy materiału literackiego.

Publikowała w „Przekładańcu”, „Literaturze na Świecie”, „Czasie Kultury”, „Poznańskich Studiach Polonistycznych”, „Forum Poetyki” oraz prezentowała wyniki swoich badań na kilkudziesięciu krajowych i zagranicznych konferencjach zarówno humanistycznych, jak i technicznych. |