

**Henryk Hollender**

Stowarzyszenie EBIB

henryk.hollender@gmail.com

ORCID: <https://orcid.org/0000-0002-1152-0752>

## Sztuczna inteligencja i prawdziwe biblioteki: szok przyszłości i co dalej?

**Streszczenie:** W artykule rozważamy sytuację kilku obszarów praktyki bibliotecznej, podatnych na przejmowanie przez narzędzia sztucznej inteligencji. Odrzucamy jednak wizję głębokiej zmiany z tym związanej, nie wierząc w jej zasięg i trwałość, ponieważ nowe technologie nie eliminują na ogół starych ani nie dają poczucia, że trzeba się na nie w całości „przestawić”. Nie potrafimy też dać wiary prognozom, że taka zmiana mogłaby być świadoma w sensie społecznym i prowadzona w sposób umożliwiający osiągnięcie głębokich skutków. Uważamy, że postęp w bibliotekach, komunikacji społecznej i obiegu informacji naukowej nie był w ostatnich dekadach wynikiem spójnej polityki, lecz zależał od wielu przypadkowych czynników. Zmiany na ogół prowadziły donikąd, a absolwenci szkół wszystkich szczebli nie są przygotowani, by odnosić korzyści z nowych możliwości orientacji w świecie i przyswajania wiedzy. Wymieniamy jednak kilka obszarów, w których obecność sztucznej inteligencji w bibliotekach przynosi już wyniki, pozwalamy sobie też na przypuszczenia, jakie jeszcze mogą być jej kierunki rozwoju. Udana wdrożenia mogą znacząco odmienić tryb posługiwania się bibliotekami i zmienić ich charakter w większym stopniu niż komputeryzacja rozpoczęta u schyłku ubiegłego stulecia.

**Słowa kluczowe:** digitalizacja, etyka komunikacyjna, generative pre-trained transformer, jednostka bibliograficzna, polityka naukowa, sieci neuronowe, „szok przyszłości”, specyfika kulturowa, spuścizna tekstowa

### Wyrównania i nawroty

Kiedy wydaje się nam, że następuje oto nowa epoka, warto przypomnieć sobie bardzo ongiś wpływowego futurologa i badacza rewolucji przemysłowych, Alvina Tofflera (1928–2016). „Wyrównania” i „nawroty”, czyli współistnienie zjawisk zakwalifikowanych jako przynależne do odmiennych epok, to najbardziej bodaj dramatyczny wątek jego słynnego *Szoku przyszłości* (Toffler, 1974), książki opublikowanej pierwotnie w 1970 r., skwitowany kompletnym pominięciem w hasłach przedmiotowych nadawanych książce przez biblioteki świata. Faktycznie szoki przyszłości są trudne do pojęcia. Każda zmiana jest i jej nie ma, gdzieś niby nastąpiła, a tuż obok wszystko zdaje się temu przeczyć, tu przyspiesza, ale w jakiejś nieoczekiwanej postaci, a tam, gdzie miała wyznaczać cel i dostarczać wzorca, raptem zwalnia. Po czym okazuje się, że wszystko to już jakby kiedyś było, choć z drugiej strony zasadnicze cechy tej nadchodzącej wreszcie przyszłości wyglądają nam na odległe od spodziewanych.

W świecie tekstów sztuczna inteligencja (*artificial intelligence*, AI) z pewnością dostarczy podobnych niespodzianek i zawodów; jest przecież wszechobecna, ale krucha i zbyt skomplikowana, by jednoznacznie uznać jej nadejście. Chociaż do przyjęcia jest również interpretacja Jacka Dukaja, według której inteligencją „sztuczną” jest wszelka zdolność do wykonywania operacji symbolicznych poza umysłem człowieka, czyli zjawisko liczące sobie ok. 200 tys. lat. Dziś chodzi nam raczej o suwerenną sztuczną inteligencję – operacje wykonywane nie tylko poza umysłem, nie tylko bez umysłu, który by zainicjowane operacje był w stanie rozumieć, ale także bez umysłu, który byłby w stanie je inicjować i kontrolować (Dukaj, 2023). Faktycznie dopiero ta faza wydaje się przynosić zmianę nad zmianami, odbierając maszynie zasadniczą różnicę, jaka dzieliła ją od człowieka, i czyniąc nieprzewidywalnymi

skutki jej działania – tak jak w końcu nieprzewidywalne były dla nas zawsze skutki procesów myślowych innych ludzi.

Nie wkroczy nam zatem AI (dalej: sztuczna inteligencja, SI) do kultury i dziejów jako jakaś nowa maszyna parowa. Ta zresztą już od początku działała w sferze wyrównań i nawrotów, pojawiała się i znikła, a tymczasem koło wodne modernizowano aż do lat 20. XX wieku. Sto lat później maszyny parowe nadal robią puf-puf, są atrakcjami turystycznymi jako zabytki, a choć silnik parowy nie bardzo ma się jak zużyć, w Chinach albo w Stanach Zjednoczonych można jeszcze kupić sobie nowy. Nowszy od parowego jest silnik elektryczny, który rywalizuje z nowym silnikiem wodorowym – jak na razie zwycięsko – czego nie można powiedzieć, sądząc po wynikach sprzedaży aut elektrycznych, o relacji „elektryków” z silnikami spalinowymi. Wyrównania, nawroty, residua i współwystępowanie zjawisk sprzecznych to codzienność wielkich przemian.

W dodatku rewolucje techniczne mają żywot pisany przez historię gospodarczą, ta zaś ma zawsze wymiar polityczny. Zasadniczym narzędziem realizowania polityki jest organizacja i finansowanie, a doświadczenia Polski w tym zakresie raczej utrudniają wiarę w następowanie szybkich i znaczących zmian. Zmianom potrzebna jest wspólna wiedza o tym, co ma właściwie nastąpić i na co przeznaczyć fundusze. Czy pamiętamy jeszcze projekt SYNAT, czyli rzekome „zadanie badawcze” realizowane za ponad 60 mln złotych w latach 2010–2013? Tytułem (celem?) zadania było *Utworzenie uniwersalnej, otwartej, repozytoryjnej platformy hostingowej i komunikacyjnej dla sieciowych zasobów wiedzy dla nauki, edukacji i otwartego społeczeństwa wiedzy* (por. np. Biblioteka Jagiellońska, Projekty [2010], [SYNAT](#)<sup>1</sup>) i od samego początku nie było wiadomo, czy platforma ma ni mniej ni więcej tylko powstać, czy może chodzi raczej o przeprowadzenie badań, które pozwolą określić warunki, jakie spełnić ma jej projekt? Poza tym przedsięwzięciem termin „repozytoryjna platforma hostingowa” wcale zresztą nie występował. Kto z Państwa wie, jakie platformy są repozytoryjne, a jakie nie? Teraz znikła też strona internetowa SYNAT-u, a pod adresem <https://www.synat.pl/> znajdujemy inne podmioty i inne sprawy. Nie chcemy tu powiedzieć, że całe te pieniądze wyrzucono w błoto – w wielu instytucjach, które jako beneficjenci SYNAT-u wciąż jeszcze w internecie figurują, byli ludzie wiedzący, czego im potrzeba. Sięgając po finansowanie Narodowego Centrum Badań i Rozwoju (NCBiR) pod szyldem SYNAT-u, zdolni byli do posunięcia tego i owego do przodu. Niżej podpisany też wziął jakieś pieniądze „z SYNAT-u” jako pracownik pewnej uczelni, a nawet uważa, że nie na darmo. Ale podsumowania nie było, a sprawą rozliczenia zajmą się być może stosowne instytucje badające dorobek NCBiR, choć chyba nie, bo pewnie nikt nie będzie miał ochoty cofać się aż tak daleko w czasie.

Tymczasem jednak projekty utkane z frazesów, ale pozbawione celu, harmonogramu, podsumowania i rozliczenia nie znikły od czasu SYNAT-u, raczej się namnażają. Na naszych oczach następuje terminacja projektu Omnis 2, mającego przecież stworzyć w Polsce nową architekturę informacji bibliograficznej; platforma traci możliwość dalszej rozbudowy, zapewne dlatego, że nigdzie nie opisano profesjonalnie jej docelowej, wielce jakoby nowatorskiej postaci (Franke, 2024). Nie ma żadnego powodu, by przypuszczać, że SI, bez względu na swoją moc i perspektywy, nie będzie rozmienniana na drobne w wielu przedsięwzięciach niedostatecznie przemyślanych i niepodlegających pełnowartościowemu prowadzeniu. Niewiele zaś nieudanych innowacji pozostaje bez wpływu na stan zastany – raczej wywołuje regres i powrót oportunistycznie zaadaptowanych dawniejszych rozwiązań, a nawet powsta-

---

<sup>1</sup> Wszystkie odesłania do stron internetowych przedstawiają wersję aktualną w dn. 22.04.2024 r.

wanie nieoczekiwanych alternatyw. Wyrazistych przykładów takiej kontrproduktywności dostarczają np. sankcje gospodarcze, takie jak te skierowane z wiadomych powodów przeciwko Rosji. Rzecz czeka jeszcze zresztą na naukowe, czyli wnikliwie wyjaśnienia, bo nie stanowi ich bodaj żaden z tekstów dokumentujących samo zjawisko (por. np. Krueger, 2024).

## Tekstowe dziedzictwo

Rozważając zjawiska dostępu do tekstów i możliwości jego modernizacji, musimy liczyć się z tym, że dziedzictwo utrwalonej pamięci przeszło przez dwa zasadnicze przekształcenia.

Po pierwsze, dla każdej niemal umownej jednostki tekstu wytworzono dokument pochodny, umożliwiający poddanie dokumentu pierwotnego operacjom logicznym bez naruszania jego właściwości i miejsca w zbiorze. I nie tylko operacjom logicznym. Manipulując metadanymi, możemy dokument pierwotny dajmy na to wypożyczyć, postawić w innym miejscu na półce, adnotować jego tekst itp. Wszelkie katalogi, bibliografie, inwentarze i indeksy pozwalają nam poznawać zbiory bez poddawania ich przekształceniom – zarządzamy nimi umownie, nieinwazyjnie, symbolicznie. Mając metadane, można obmyślić algorytmy wielkich porządków i przekształceń.

Po drugie, wytworzyliśmy cyfrową kopię światowego korpusu tekstów, ale oryginały nie przestają istnieć, podczas gdy my zajmujemy się teraz właściwie tylko sektorem zdigitalizowanym, nie do końca świadomi, jak ma się on do wielkiego zasobu wciąż pozostającego poza cyfryzacją. Jasne, że wielkie akcje digitalizacyjne mogą jeszcze wiele przywrócić naszej pamięci, niewykluczone jednak, że spora część tego residuum ulega trwałemu zapomnieniu. O różnicach w źródłowym badaniu oryginału i wtórника, jako problemie godnym głębszego namysłu, wspomniano przy innej okazji (Hollender, 2021, s. 412–414).

Stworzyliśmy zatem dwa światy, alternatywne w stosunku do istniejącego i narastającego zasobu „oryginałów”: metadanych i wtórników. Mówimy o dziedzictwie tekstualnym, ale to jest skrót myślowy i uproszczenie, ponieważ rzeczywistość tekstów pozbawionych wersji cyfrowej, w tym także takich, dla których nie mamy metadanych, pozostaje na ogół niewykorzystana. Nic z nią nie robimy i zapewne nie wniknie w nią żadna sztuczna inteligencja. Istnieje ciemna strona rzeczywistości pisma (i obrazu), nie wiemy nawet, jak jest wielka, choć paradoksalnie obcujemy z nią w sposób autentyczny, czasem pewnie i głęboki na co dzień: została wchłonięta przez kulturę. Nie bardzo jednak jest jak planować jej operacjonalizację. Dawniejsze bibliografie wyświadczały nam uprzejmość, informując o swoim zakresie i zasięgach; nowe narzędzia organizacji wiedzy raczej tego nie czynią. W latach 90. przejmowano się retrokonwersją katalogów. Nie przerywając jej, na wszelki wypadek tworzone niekiedy elektroniczne wtórники katalogów kartkowych, z elektroniczną symulacją operacji wyszukiwania, dokonywanej na kartkach, wypełnionych pismem drukowanym, maszynowym, ręcznym. Które z tych działań zdołano zakończyć (z „żywą” kontynuacją w obszarze zbiorów przyrastających)? Tego już raczej nie wiemy i niezbyt to kogokolwiek obchodzi. Zbiór książek, katalog książek, katalog kartkowy online, katalog online – jakie są wzajemne relacje tych zbiorów, ogłasza dziś już mało która ksiąźnica, a jej pracownikom<sup>2</sup> nie zwracano tym raczej głowy w ramach kierunkowych zajęć na studiach informacyjnych.

---

<sup>2</sup> Rzeczowniki takie jak ten używane są w tym tekście w rodzaju męskim, ponieważ autor jest zdania, że są one zdolne do wyrażania roli czy kompetencji każdej osoby bez definiowania jej płci, które zazwyczaj jest zbędne. W tym właśnie

Tak więc tofflerowskie mechanizmy nierównomierności i niejednoznaczności zmian występują także w świecie tekstów. Współistnieją epoki, typy instytucji i zbiorów, praktyki działań na tekstach i na ich opisach. Nie ma co liczyć, że jakiś program komputerowy obsłuży cały ten świat w sposób jednolity i przewidywalny, bez względu na ścisłość projektowania. Z definicji, nie obsłużą go programy generatywne „wyuczalne” (*pretrained*, uprzednio szkolone). Dla przykładu: pokolenie przemocowych działaczy może sobie tutaj w Polsce dysponować językiem, jaki wykształciła w nim szkoła i kultura kraju „demokracji ludowej”, może też nadążać za wielkimi innowatorami języka literackiego i mową mediów. Jednak dostępny dla niego serwis, taki jak ChatGPT, był ćwiczony na innym zasobie tekstów, toteż niekoniecznie rozumiemy, co też on nam podsuwa, nie mówiąc już o tym, że on może z łatwością nie zrozumieć nas, gdy zlecamy mu zadania. Chociaż na razie zlecamy mu proste i tych rozbieżności na ogół nie widać. Jest to rozmijanie się tego samego rodzaju, co to, które występuje, gdy czat odmawia podzielenia się z nami wiedzą istniejącą od niedawna, usprawiedliwiając to faktem, że zaktualizowano go w określonym roku, poza który nie jest w stanie sam z siebie wyjść. W czasie, z którego mogą pochodzić poszukiwane przez nas dane, „nie chodził już do szkoły”, nie rozwijał się, klepie więc teraz to, czego go nauczono wcześniej. Zjawisko to i inne – jakby ludzkie – ograniczenia czata są trudniejsze do wykrycia i mniej oczywiste dla jego statystycznej klienteli, która nie rozumie konwencji i nie ma żadnej wiedzy o tym, że szukając zrozumienia można się wyłożyć na czymś takim jak różnice kulturowe.

## Groźba plagiatu, nadzieja oryginalności

Moc algorytmów tworzących plagiaty z wielkich i starannie zmiksowanych zbiorów danych nie prowadzi w potocznym odbiorze do fałszerstwa, ponieważ nie widać tu już podmiotów, którym by coś konkretnego ukradziono<sup>3</sup>. Jeśli zadanie polega na stworzeniu kopii konkretnego istniejącego obiektu (namaluj mi Kate Moss w słońcu na tarasie w Tangerze itp. – otóż jest taka fotografia Alberta Watsona), to dylematy prawno-etyczne mogą powrócić, ale nie są istotne, bo harce amatorów z nowymi zabawkami to margines całej sprawy. Clou programu i esencja wiązanych z SI nadziei to tworzenie nowych, oryginalnych wypowiedzi, istniejących po to, żeby nie były podobne do wcześniej znanych, ale żeby spełniały oczekiwania zamawiających, na takich istniejących obiektach wychowanych. Wygląda – wrócimy do tego za chwilę – że to się da zrobić. Nigdy nie zanegowano zasadniczych koncepcji leżących u podstaw prawa autorskiego: dzieło oryginalne tworzy osoba fizyczna. Wytwór zlecony sztucznej inteligencji nie jest naszym tworem. Nawet jeśli od nas pochodziłaby koncepcja, instrukcja i nadzór autorski, wymyka on się prawnautorskim ocenom.

---

rodzaju się przyjęły, ale nie oznacza to już dyskryminacji, choć bez wątplenia leżała ona u źródeł takiego, a nie innego kształtowania się języka.

<sup>3</sup> Por. komentowaną przez wielu uwagę Noama Chomsky'ego: *The human mind is not, like ChatGPT and its like, a lumbering statistical engine for pattern matching, gorging on hundreds of terabytes of data and extrapolating the most likely conversational response or most probable answer to a scientific question. On the contrary, the human mind is a surprisingly efficient and even elegant system that operates with small amounts of information; it seeks not to infer brute correlations among data points but to create explanations. Let's stop calling it Artificial Intelligence and call it what it is: Plagiarism Software. It doesn't create anything, just copies existing works from artists and alters them sufficiently to escape copyright laws. It's the largest theft of property since Native American lands by European settlers (Chomsky, 2023)*. Jest interesujące, że ChatGPT zgrabnie na to odpowiedział: *Chomsky's perspective might give the impression that AI aims to replace human intelligence or creativity. However, the primary goal of AI, especially in systems like ChatGPT, is not to supplant human intellect but to augment it. AI serves as a powerful tool that can handle vast amounts of data, providing insights and support to human endeavors. It's a collaborative force, not a competing one (Cronin, 2024)*.

Ponieważ w posługiwaniu się SI jako podwykonawcą nie można wskazać działania skierowanego przeciwko autorowi czy jego prawom, pragnienie posiadania bezpłatnego i kompletnie bezpiecznego wyrobnika wydaje się powszechne. Tu potrzebna byłaby jednak refleksja nad postawami wobec różnych komputerowych asystentów i możliwościami dokonania w skali globalnej czegoś w rodzaju nowego podziału pracy. Wydaje się, że te nadzieje są wielkie, choć nieartykułowane tak wyraziście, jak w dawniejszych komentarzach do różnych cywilizacyjnych cudów związanych z masową produkcją żywności czy środków produkcji, dostępnością podróżowania, pokonywaniem chorób, zdalnym przekazie treści i emocji, oddzieleniem płciowości od odpowiedzialności za potomstwo itp. Nowe techniki informacyjne nie odgrywały nigdy porównywalnej roli, raczej kreując nowe potrzeby, niż zaspokajając te, które odczuwano jeszcze przed ich nadejściem. To zmienia się chyba od niedawna i zmienia się na wielką skalę. Rozmaite udogodnienia przyjmują się szybko i bezrefleksyjnie z poparciem rynku, także w sferach, w których ich nie oczekiwano. Wraz z nimi – inne środki, niekoniecznie przynoszące bezpośrednie skutki w sferze materialnej. A więc wszelkie postacie i następstwa „telegrafu bez drutu”, programowanie, roboty, zaawansowane algorytmy obliczeniowe, internet, usługi zdalne... Otoczeni urządzeniami nie bardzo widzimy, czym różnią się od siebie; traktujemy je jak równorzędne gadżety, a intuicja podpowiada nam, że możemy ich używać jednocześnie. Relacyjna baza danych, wyszukująca ściśle zdefiniowane obiekty w ograniczonym, choćby i wielkim zbiorze? Oprogramowanie wykorzystujące hiperłącze do wyszukiwania stron WWW? Wszystko to jest jedno i to samo i ma nam służyć! Takie jest, jak się zdaje, przeciętne postrzeganie nowych technologii i naszej ludzkiej wolności, by się nimi posługiwać. Plagiat jako zagrożenie? Oryginalność jako wymóg? Kultura masowa wie swoje i nie chce w to wnikać.

Być może więc mamy poczucie mocy i sprawczości, prawnego bezpieczeństwa oraz istnienia niejasnych, ale skutecznych technik, których działania nie rozumiemy, ale które okazują się spełniać nasze życzenia. Statystycznie rzecz biorąc, nie rozważamy zazwyczaj kwestii, czemu coś w ogóle miałyby te życzenia spełniać; świat uwierzył, że jeśli coś można mieć za darmo, to mieć to trzeba. I oto nagle jest uczelnia, w uczelni zajęcia, a na zajęciach osoba – poproszona o wyszukanie w bazie dajmy na to PubMed artykułu o określonych cechach formalnych i treściowych – bez żenady powierza to zadanie urzędzeniu zwanemu na ogół przez studentów „goglami”. Traktując internet jako źródło szybkiej wiedzy, przyzwyczajona do zwracania się do „niego” z różnymi prośbami, artykułowanymi zazwyczaj w języku naturalnym, czyli bez żadnego namysłu nad jakimiś tam „terminami wyszukiwawczymi”, wpukuje natychmiast w swój telefon prośbę: „znajdź mi w bazie PubMed pierwszy artykuł w języku polskim, dotyczący możliwości laparoskopowej operacji dróg żółciowych”. Lawina ruszyła; przypomnienie, że to ona, osoba uczestnicząca w zajęciach, miała umieć przeprowadzić wyszukiwanie w bibliograficznej bazie danych, budzi zdumienie: dlaczego „google” nie mają tego zrobić za nas, skoro istnieją i skoro będą szybsze w działaniu?

I już nie dziwimy się, mając w innej grupie słuchacza czy słuchaczkę, relacjonującą z nutą podziwu możliwości wykorzystywania przez kogoś ze znajomych ChatGPT do obsługi programu Excel. *To [były] prawdopodobnie pojedyncze procedury excelowe, typu 'oblicz odchylenie standardowe' czy 'wylicz czas spłacania pożyczki przy zmiennej stopie procentowej'. Ale Excel na szczęście to coś znacznie potężniejszego* – napisał mi znawca tematu, któremu opowiedziałem na Facebooku tę historię<sup>4</sup>. Równocześnie coś jest jednak na rzeczy: czat najwyraźniej potrafi wykonywać czynności, które z pewnym wysiłkiem uczyliśmy się wykonywać sami i było to oczywistością. Teraz jest inaczej. Śledź paczkę

<sup>4</sup> P. Wimmer do H. Hollendra, Facebook, 22 lutego 2024. Dziękuję Pawłowi za fascynujące posty na temat zastosowań sztucznej inteligencji. Por. też: WIMMER, P. (2024). *Z Excelem na studia*. Warszawa: Wydaw. Libretto.



z chatbotem – doradza ostatnio firma InPost. Kilka lat temu mogło nam się wydawać, że możliwość śledzenia drogi paczki do adresata jest rzeczą ważną, jeśli ktoś otrzymuje wiele paczek (np. biblioteka kupująca książki) albo kiedy przesyłka pilnie oczekiwana, albo szczególnie wartościowa... Procedura śledzenia nie była uważana za trudną, ale teraz okazuje się, że jest wystarczającym zawracaniem głowy, by rozejrzeć się za maszynką, która nas od niej uwolni i po prostu poda wynik. To pewnie dopiero początek bardzo wygodnego życia, które najwyraźniej stanowiło marzenie ludzkości.

Toteż media społecznościowe roją się już od reklam programów i usług, w których rozmaite mniej lub bardziej zaawansowane programy robią coś nie tylko dla nas, ale i za nas. Wygląda, że zasadniczy sprzeciw wobec takich praktyk pochodzi raczej od twórców, których dzieła – językowe, plastyczne, muzyczne itp. były używane do „trenowania” programów zdolnych następnie do samoczynnego tworzenia niż od osób zainteresowanych jakością produktu końcowego oraz ogólnym niepsuciem rynku – szefów, polityków, wydawców, słuchaczy. Ludzkość jest skłonna zaakceptować twórców, posługujących się sztuczną inteligencją, wraz z ich dziełami. Liczy się efekt, a nie podejmowane działania, a wątpliwości etyczne wydają się nie na miejscu dopóki za końcowy efekt odpowiada fizyczna osoba ludzka, znana imiennie swojemu odbiorcy – przełożonemu albo nabywcy na rynku.

Inaczej rzecz ma się bodaj tylko w obszarze piśmiennictwa naukowego – praca firmowana nazwiskiem autora musi być wykonana samodzielnie; prac wykonanych niesamodzielnie nie liczy się do dorobku (Lee, 2023). ChatGPT nie może być uznany za autora czy współautora pracy naukowej – to jest jasne, ale czy byłoby jasne bez jednoznacznych zastrzeżeń w tej sprawie ze strony czołowych wydawców, jakich fala osiągnęła swój szczyt w 2023 roku? (Sample, 2023). Warto przy tym dostrzec, że rygorystyczny etyczny w świecie edytorstwa naukowego idzie w parze z ostrożnością, by nie okazywać autorom podejrzliwości i nie wchodzić im na głowy. Tak, jakby nie chodziło o to, jak artykuł powstał, ale o to, by w gronie autorów nie znalazła się nagle publicznie dostępna usługa wykonywana przez podmiot niebędący człowiekiem.

Mówiąc o powstawaniu artykułu, nie mamy na myśli przebiegu badań i przygotowania publikacji – tu jest już tylu krzemowych asystentów, że korzysta się z nich skwapliwie, za późno na jakiegokolwiek ingerencje. Wielu publicystów poświęca uwagę tej sprawie, a tytuły artykułów, takie jak *Więcej czatu* (Rotkiewicz, 2024) nie pozostawiają żadnych wątpliwości: zastosowania szkolonych programów generatywnych są liczne i łatwo je wymienić. Cytowany artykuł zasiewa jednak ziarna niepokoju, i to już w leadzie: *Badacze coraz częściej i chętniej sięgają po narzędzia sztucznej inteligencji. Im więcej sukcesów, tym więcej pytań o potencjalne zagrożenia* (Rotkiewicz, 2024).

Nikt w każdym razie nie skarciłby autora, gdyby ten napisał program zdejmujący z niego część typowo autorskich obowiązków (koncepcja i planowanie badań, analiza wyników, formułowanie wniosków itd.). W pracy podpisałby się jako współautor i jako twórca aparatury czy oprogramowania. Tyle że opracowanie aparatury dostarczającej materiałów o prawdziwie twórczym charakterze, a więc opartej na sieci neuronowej, jest odnotowane w autorskich oświadczeniach etycznych. I na tym nie koniec. *Wiodące czasopisma naukowe wymagają od autorów podpisania formularza deklarującego, że są oni odpowiedzialni za swój wkład w pracę. Ponieważ ChatGPT nie może tego zrobić, nie może być autorem*, orzeka Holden Thorp (2023), redaktor naczelny „Science”. Tę formułę powtarza się z pewną satysfakcją, bo ona bynajmniej nie zabrania udziału pozamózgowych sieci neuronowych w pracach wiodących do publikacji. Jako przykład podaje się nawet napisanie streszczenia (usłyszane w audycji radiowej od mocno lansującej się „jutronautki”), tak jakby streszczenie nie było trudną autorską ro-

botą, tym trudniejszą, że w dużych zespołach autorskich mało który z jego członków panuje nad całą linią rozumowania, stanowiącą szkielet tekstu i istotę zawartej w nim treści. Niektóre czasopisma każą podawać w artykule, czego dokonał każdy z autorów z osobna, inne nie. Ale brzytwa Thorpa precyzyjnie i wyłącznie wycina pisanie przez czata całego tekstu lub jego fragmentu, przy założeniu wynikającym z wcześniej ugruntowanych i – jak rozumiemy – praktykowanych zasad, iż w powstawaniu pracy w ogóle nie brali udziału autorzy ukryci – *ghost writers*, a zwłaszcza – twórcy nieświadomi, iż ich tekst splagiatowano.

W zaleceniach „Science” nie było najwyraźniej potrzeby, by to powtarzać; może przyjęto by to jako nietakt, niedelikatność? Najwyraźniej Thorp uważa, że gdyby czat pomagał w pisaniu, to wszyscy by się zgodzili, by było to jawne, tyle że musiałby mieć on wszystkie atrybuty i sprawności inteligentnego autora, a na dodatek dysponować wolną wolą. Takiego czata jeszcze nie ma. Brzytwa jest ostra, ale nie tnie na oślep, redaktor naczelny bardzo się pilnuje, żeby nie zakazywać za dużo, zwłaszcza tego, o czym już od dawna wiadomo, że jest zakazane i że zakazu się przestrzega.

## Tu i teraz

Jak ta brzytwa zadziała w warunkach polskich, gdzie obfitość oskarżeń i sporów o plagiat nie pozwala wierzyć, iż zasady etyki naukowej są powszechnie respektowane? I jeszcze na dodatek w warunkach, w których operacje na tekstach – nie tylko tworzenie opowieści, ale także dobór argumentów i lokowanie ich w kompozycji artykułu, sprawiają nam autorom coraz więcej trudności? Jasne, że wciąż czytamy dobre artykuły naszych kolegów i to nie tylko w najlepszych czasopismach światowych: duża część redagowanych w Polsce czasopism osiągnęła poziom indeksowania w Scopusie czy Web of Science. Mają wyliczony Impact Factor, dobrą redakcję i korektę, funkcjonalny layout, a punktacja przyznawana przez ministerstwo nauki w tym czy innym wydaniu politycznym bynajmniej nie zawsze woła o pomstę do nieba. Kto jednak wie, jak wygląda redakcyjna kuchnia? Gdzie są oświadczenia polskich czasopism na temat ich polityki korzystania z neuronowych wspomagaczy i dopalaczy?

Niżej podpisany nie uważa, by miał odpowiednią wiedzę w tym zakresie, ale niestety zdaje sobie sprawę, że posiadanie jakiegokolwiek wiedzy bywa źródłem dyskomfortu. Jest nim na pewno wspomnienie artykułu otrzymanego do publikacji w uczelnianym „publishing house” (bo tak chciały nazwać swoje wydawnictwo władze uczelni, oddając je czasowo pod opiekę zespołu uczelnianej biblioteki). Był to szkic, który autor/ka uważał/a za dojrzały do publikacji i dopytywał/a się, kiedy to wreszcie nastąpi. Nie było możliwości nawiązania rozmowy, że artykuł jest niegotowy do jakiegokolwiek redakcyjnej obróbki. Każdy, kto ociera się dziś o instytucje naukowe, może zebrać garść podobnych doświadczeń. Słyszysz się nawet, że ktoś tam zamierza wydawać czasopismo naukowe, które nie będzie miało numerów – jak na każdej „stronce” materiał będzie się pojawiał (i znikał?) w trybie nieregularnym. Wydawca nie zauważył cykliczności czasopism i nie wie, do czego jest potrzebna, słyszał tylko, że trzeba uzyskać w Bibliotece Narodowej International Standard Serial Number.

Biblioteka, nawet niepołączona z wydawnictwem, stanowi niezły punkt obserwacyjny. Kto dostaje dziś od autora prac naukowych, który ponadto prowadzi zajęcia, listę lektur dla studentów i widzi, że na tej liście dane publikacji są wolne od pomyłek? Afiliacje – kto pamięta, jaka jest oficjalna i stosowana w cytowaniach nazwa jego uczelni? Potoczna korespondencja – prawie zawsze pisana na kolanie, i zapewne z głową, ale o istnieniu i funkcjach interpunkcji to głowa już nie wie. A autorów artykułów nazywających redakcję korektą spotyka się niestety powszechnie, podobnie jak studentów

niezdolnych po przeczytaniu praktycznie czegokolwiek do przetworzenia (zwłaszcza streszczenia) czy interpretacji tekstu. Jasne, że wspominamy tutaj zjawiska znane, niekiedy ze zgrozą dyskutowane, a nawet badane, może bez zgrozy, ale zdecydowanie w poczuciu szoku przyszłości (Baron, 2023). Dyskusja nad umysłowością i alfabetyzacją osób wychowanych ze smartfonem w dziecięcej rączce dopiero się jednak w Polsce zaczyna i daleko w niej jeszcze do lęku, z jakim przed kilkadziesiąt laty przyjęto telewizję; powiedzmy raczej, że jest niemrawa, bo smartfona mają wszyscy, a przedwczesna smartfonizacja dzieci następuje powszechnie i raczej na życzenie rodziców niż przeciwko nim. Pozostaje mieć nadzieję, że rozwijając się, dyskusja ta scali się z dawniejszymi, generalnie pesymistycznymi diagnozami stanu czytelnictwa i kompetencji kulturowych oraz okoliczności politycznych, utrudniających w Polsce np. promocję czytelnictwa (por. Szot, 2023).

Jasne, że przyjmuje się chętnie, iż są zielone wyspy na tle komunikacyjnej bylejakości społeczeństwa. Na pewno nie są „byle jacy” ani czytani i nagradzani pisarze, ani autorzy cytowanych artykułów naukowych, nie nasuwa takiego skojarzenia ogólny obraz polskich środowisk kreatywnych. Czy zatem oczekujemy, że sztuczna inteligencja, chociażby jako narzędzie akulturacji, alfabetyzacji i rozwoju, pozwoli znacząco powiększać liczbę i rolę społeczną zarówno twórców, jak i twórczych obywateli? Czy mamy przyjąć, że inteligenci niezbyt pełnowartościowi (zażartujmy: *artificial intelligentsia*) są lub będą przez sztuczną inteligencję ratowani i szkoleni? Czy jest w ogóle na to zapotrzebowanie? Czy w Polsce chodzi teraz o bardziej sprawczych, świadomych i wolnych ludzi (por. np. Janicka, 2017), czy tylko o korzystniejsze autoprezentacje, efekty i pozory?

### **Czego jeszcze nie ma (i może nie będzie)**

Skoro nauka, szeroko wykorzystując sztuczną inteligencję, odrzuca jakiegokolwiek bezpośrednie udzielenie przez nią wsparcia w budowaniu twórczej wypowiedzi, to może odnotowany jakiś czas temu brak zainteresowania bibliotek perspektywami i zastosowaniami SI (Wheatly, 2019) jest zrozumiały i nie ma co się mu dziwić? Może nieprzypadkowo nowe ustawodawstwo europejskie nie porusza kwestii powstawania tekstów naukowych?

Teksty literackie to jednak jest coś innego, więcej tu punktów styrczych i więcej wątpliwości. Europejska Rada Stowarzyszeń Tłumaczy Literatury wydała wnikliwe oświadczenie, w którym zwraca się uwagę, że utwory literackie mogą służyć do trenowania sztucznej inteligencji, na co autorzy takich utworów mogliby się nie zgadzać (Oświadczenie, 2023). *Maszyna nie jest tłumaczem, a zaledwie translatoidem. Nie tłumaczy, lecz generuje materiał tekstowy. Wykorzystanie sztucznej inteligencji prowadzi do standaryzacji przekładu, zubaża kulturę języka pisanego i język jako taki, między innymi za sprawą efektu primingu (tendencji do sugerowania się pierwszą z opcji podanych przez maszynę) i samoskażenia (pogorszenia jakości wyników, kiedy maszyna uczy się sama od siebie* (Oświadczenie, 2023). Jak się zdaje, podobna argumentacja mogłaby również dotyczyć tekstów naukowych. *Koncepcja, jakoby niektóre książki lepiej niż inne nadawały się do przetwarzania za pomocą sztucznej inteligencji, otwiera drogę dla niebezpiecznego rozróżnienia między literaturą wysoką a każdym innym tekstem* (Oświadczenie, 2023). Zgoda, ale nie ma zwyczaju kwalifikowania jako nieuczciwości posługiwanie się translatorami przy czytaniu tekstów przekładanych i nawet tłumacz literatury najwyższych lotów może posłużyć się taką pomocą, poszukując najlepszego odpowiednika jakiegoś słowa, wyrażenia, zdania, może ciągu zdań – dokładnie tak, jakby sięgał do drukowanego słownika. A jeśli tłumacz pobierze i przetłumaczy cały fragment? Rozumiemy, że nadal wolno mu to zrobić, ponieważ nikt go nie sprawdzi – liczy się wartość efektu końcowego, za którą tłumacz nie przestaje odpowiadać, z cze-



gokolwiek korzystał. Od tego jednak miejsca w Oświadczeniu tłumacze brną we wznieście okrzyki o przewagach człowieka nad maszyną, o prawie tłumacza do rozwoju duchowego, wreszcie, psując cokolwiek klasę oświadczenia (no bo pewnie nie jego skuteczność) – o prawie do wynagrodzenia. Lodowata logika uczonego, który jednym zdaniem zdyskwalifikował sieci neuronowe jako niezdolne do podpisania się, jednak jakoś lepiej nas przekonywała.

Brakuje w tej dyskusji pewnej perspektywy, która wydaje się odwracać proporcje między szlachetnym samotnym tłumaczem a maszyną wymagającą przyuczania, czyli unicestwiającą oryginalność (popętniającą plagiat). Zasygnalizujmy ten problem skrótowo, może ktoś go kompetentnie rozwinie; chciałoby się tego doczekać, bo moglibyśmy tu mieć do czynienia z niezwykłymi spotkaniami badaczy, tłumaczy i bibliotekarzy (pracowników informacji). Jest mianowicie (wspominamy o tym powyżej) wielki korpus prac niezdigitalizowanych, zdigitalizowanych niekompetentnie (zwłaszcza: bez podania warstwy tekstowej), niewydanych nigdy drukiem; są zapomniane prace i zapomniani autorzy. Są wreszcie dzieła ongiś obiegające, dziś uchodzące za niezrozumiałe. Zwykle istnieniu takowych towarzyszyły wezwania, by je „na nowo odczytać”. Może w tym celu warto je na nowo przetłumaczyć? Może wnikliwie, modyfikowane w trakcie badania przekłady automatyczne, amplifikując w kolejnych próbach różne znaczenia, ujawniłyby nowe wersje tekstu i taki np. Jan Jonston (1603–1675), jedyny właściwie polski uczyony XVII w., nie licząc Jana Heweliusza, ale wtórny i uważany za niewolnika swoich żmudnych taksonomii, bynajmniej nieczytany dziś chętnie ze względu na trudną łacinę – nie ujawniłby twórczego oblicza? W dodatku nudziarze i grafomani zostawili spuściznę większej bodaj objętości niż umysły bardziej zasługujące na znajomość ich dzieł, ale podział może być fałszywy, listy rankingowe mogą się zmieniać i pisarze słabo teraz obecni po kolejnym szoku przyszłości mogą okazać się pionierami.

Tego typu fantazje można kontynuować. Wyobrażalne jest żądanie nowych narzędzi oceny tekstów naukowych, które polegałyby na poszukiwaniu w nich cech czysto językowych, zbliżających je do innych tekstów, które już dowiodły swojej (merytorycznej, nie literackiej!) oryginalności i wpływowości. Nie wydaje się to mniej racjonalne od oceniania artykułu po ocenach, jakie otrzymały inne artykuły opublikowane wcześniej w tym samym czasopiśmie – a tak czynimy, oceniając i przyznając autorowi czy instytucji „punkty za czasopisma”, w których zamieszczają swoje teksty.

Równie wyobrażalne jest posłużenie się sztuczną inteligencją dla wyłonienia innych niż tradycyjne jednostek tekstu. Biblioteka czy bibliografia gromadzi, rejestruje, udostępnia i przetwarza takie jednostki, częściowo zresztą tylko w postaci metadanych. Kryterium wyłonienia jednostki to domniemany akt jej materialnego powstania i zachowanej integralności; najlepiej w tym systemie czują się książki i czasopisma, podlegające kryteriom rynku, który zarządza jednostkami towaru i wcześniej wykreował stosowne miary. Drukowane klocki czy rękopisy gorzej się czują w tym systemie, nawet z czasopismami jest mnóstwo problemów, jeśli chodzi o obsługę metadanych. Mocną i jednoznaczną jednostką, podatną na przekształcenia typu streszczanie czy wyłanianie opisu bibliograficznego, jest tylko artykuł. Czy gdyby sztuczna inteligencja rozsypała nam wszystkie domniemane jednostki tekstu i ułożyła je na nowo, na przykład według kryterium integralności treściowej, to nie wyłoniłaby się z tego możliwość odmiennej rekonstrukcji treści? Żeby tego dokonać, trzeba by pewnie przepuścić przyszłą sieć neuronową przez bardzo specyficzny Wielki Model Językowy.

A co z obrazami? Ze względu na swą jednoznaczną postać fizyczną mogą się nam wydawać kompletnymi, zamkniętymi wypowiedziami, jednoznacznie odgraniczonymi od innych, ale porządkowane na

sposób biblioteczno-bibliograficzny ujawniają, jak bardzo potrafią się od siebie różnić. Wprowadźmy w takiej np. [Europeanie](#) frazę „sala operacyjna”. Otrzymamy foty z dwóch filmów („Zazdrość i medycyna”, „Godzina szczytu”), kilka historycznych zdjęć z różnych budynków szpitalnych, a następnie pracę naukową z zakresu architektury, gdzie wspomina się o operacyjnej sali bankowej. Kto to wykozysta i w jaki sposób? Posłużenie się wyszukiwaniem zaawansowanym jest dostępne, ale byłoby wielce pracochłonne, zwłaszcza dla przygodnego użytkownika serwisu. Czy jednak nietekstowy materiał wizualny nie mógłby nam się jakoś nietekstowo sklejać ze sobą, grupować, agregować za pomocą dostępnego przecież wyszukiwania obrazem? Jasne, że tu nie wchodzi w grę wydatki, jakie ponosi np. nabywca policyjnego systemu rozpoznawania twarzy, nie będzie zatem dostępna skuteczność porównywalna z takimi systemami. Natomiast mielenie przez systemy wyszukiwawcze wąskiego wachlarza danych tekstowych obsługujących nietekstowe materiały, z czym mamy do czynienia w tej chwili, wydaje się przedłużaniem w nieskończoność prowizorki, na jaką już przestajemy zasługiwać.

Czy zatem misja [Europeany](#) nie powinna już zapowiadać odchodzenia od współczesnej, wielce ograniczonej inżynierii lingwistycznej (Mykowiecka, 2007), na rzecz posługiwania się komunikatami językowymi i obrazowymi na równi, dla grupowania i penetrowania znaczeń? A także: czy postęp ten, być może trwale zabezpieczający nas przed niedoborem informacyjnym, nie pograży nas w jeszcze większym chaosie? Jakie czynniki zapewnią mu dyscyplinę i odpowiedzialność?

## Próby, zastosowania, postępy

Na tle namnażających się i pączkujących zastosowań SI jej dostępne w tej chwili narzędzia biblioteczne wydają się marginesem, któremu nikt nie poświęca większej uwagi. Wygląda to tak, jakby sięgano po nią głównie po to, żeby uczyć posługiwania się programami, opartymi na... sztucznej inteligencji (Tanzi, 2024). Ale może i szerzej: uczyć innowacyjnych sposobów przyswajania i porządkowania wiedzy. Tyle że niektórzy autorzy są zafascynowani sztuczną inteligencją, ale już niekoniecznie bibliotekami i bibliotekarzami: w swoich badaniach, rozważaniach i ocenach perspektyw stosowania SI przez biblioteki trzymają się na dystans od tego środowiska, tak jakby nie miało ono własnych, ściśle zdefiniowanych zadań. Otrzymujemy od nich wizję bibliotekarzy jako badaczy blogów i serwisów wiki, w bliżej nieokreślony sposób „gromadzących informacje”, w nieskończoność „mapujących treści” i dokonujących streszczeń itp. (Oyelude, 2021). Zgadza się to z powszechnymi intuicjami, że funkcje biblioteki, serwisu bibliograficznego i wydawnictwa nakładają się już na siebie, ale nie wskazuje żadnych jednoznacznych nawiązań do obszarów, w których bibliotekarzom faktycznie potrzebna jest pomoc dla wytworzenia nowej jakości. Takiej, jaką wytworzyli, projektując np. mechanizmy kontekstowego i dynamicznego linkowania naukowych zasobów informacyjnych. Innowacyjność i profesjonalną trafność tamtych rozwiązań trudno będzie kiedykolwiek doścignąć.

Wskazuje się też na następujące możliwości sztucznej inteligencji w bibliotekach:

- pomoc w obsłudze użytkowników,
- rozwój narzędzi komunikacji z otoczeniem społecznym, w tym marketingu.

Obydwa te kierunki wydają się oczywistością w całym sektorze usług, nie tylko informacyjnych. Poimińmy je tutaj, jako że nie wydają się one dotyczyć istoty środowiska bibliotecznego. Możliwość szybkiego wytwarzania materiału graficznego (ilustracyjnego), jaka fascynuje wielu bibliotekarzy, może uatrakcyjniać obieg informacji, pobudzać wyobraźnię, wносить element kreatywności do informacyjnych procedur, ale czy pozwoli zmodyfikować ich zasadnicze algorytmy? Czy wyobrażalne jest np.

posłużenie się jakimś generatywnym narzędziem do wykreowania materiału ilustracyjnego w celu uzupełnienia utraconych lub zniszczonych obiektów? W ten sam sposób muzea mogłyby obok posiadanych obrazów wieszać nieistniejące, ale możliwe (albo udokumentowane historycznie, lecz materialnie nieznanne). Jesteśmy już gotowi na taką grę, skoro akceptujemy np. w komnatach Wersalu wykorzystanie pustych ram obrazów do prezentacji filmów związanych tematycznie ze zwiedzaniem wnętrza. Ponieważ zaś pewne różnice między biblioteką i muzeum są wciąż niewarte zacierania, przypomnijmy tu jeszcze niemiecki projekt Der Bilderbazar – eine Zeitmaschine, polegający na estrahowaniu z bogatego zasobu ilustracji czasopisma „Der Bazar” (od 1857 roku) elementów ilustracji. Jednostkami powstającego zbioru nie były poszczególne ilustracje, lecz obiekty zilustrowane, np. sztuki modnej odzieży. Narzędzie nauczyło się takiej analizy i zostało wykorzystane do tworzenia wystawy przedstawiającej materialne obiekty trójwymiarowe uzupełnione elementami „obrazkowymi” i metadanymi umożliwiającymi komentowanie całości lub dalsze poszukiwania (Sühl-Strohmeier i Tappenbeck, 2024, s. 225)<sup>5</sup>.

Ciekawsze wszakże wydają się zapowiedzi wspierania procesu gromadzenia i organizowania zbiorów, w tym rozwój narzędzi opracowania rzeczowego (Golub, 2021), a nawet narzędzi zarządzania zbiorami, które łączyłyby w sobie dobór i opracowanie. Donosząc o rynekowym debiucie narzędzia Over-Drive, „Library Journal” zapowiada możliwość tworzenia i kontroli zbioru, w którym teksty/dokumenty byłyby rozpoznawane przez o wiele większą niż obecnie liczbę punktów dostępu i w elastyczny sposób agregowane dla tworzenia nieograniczonej liczby podzbiorów, spełniających różne chwilowe lub stałe potrzeby. Rysuje się tu możliwość podejmowania zadań, w których wyszukiwanie i dobór tekstów, a także tworzenie nowych tekstów – zasadnicze czynności wspieranego przez bibliotekarza użytkownika systemu informacyjno-wyszukiwawczego – stają się sekwencją kroków nierozdzielnych, podlegających ścisłej kontroli, którą zapewnia szczegółowy wgląd w ich jednostkowe cechy. Prześledzenie losów tego pomysłu wymagałoby odrębnych badań; jak się zdaje, zmierzamy tu w kierunku tworzenia wirtualnej biblioteki ze wszystkiego, co dotąd powstało jako zapis, na potrzeby każdego zapytania informacyjnego, włączając wszakże do dostępnego instrumentarium także narzędzia artykulacji samego zapytania (Enis, 2021)<sup>6</sup>.

Biblioteka uzbrojona w sztuczną inteligencję staje się zatem agencją usługową organizującą zbiór informacji potrzebny dla zrealizowania konkretnego zadania. Może to być skromne lub rozległe zapytanie faktograficzne, na które może i mógłby odpowiedzieć ChatGPT, ale nie, my chcemy sami, tzn. z dedykowanym oprogramowaniem, z nakreśleniem niuansów i alternatyw, których czat jako konformista będzie unikał, dla obsłużenia problemu badawczego o niezdefiniowanej jeszcze bazie źródłowej (a samą bazę chcemy mieć w obrębie swojego warsztatu, czego czat z założenia nie daje).

Co więcej, biblioteka może zapewne profilować użytkownika i tworzyć model jego potrzeb, podsuwając mu dokumenty, o które nie prosił. Byłby to zapewne kierunek cokolwiek niebezpieczny ze względu na ryzyko siłowego zmutowania projektu badawczego lub czytelniczego. Formuła biblioteki jako ser-

<sup>5</sup> Książka, z której się o tym dowiadujemy, nowy podręcznik bibliotekarstwa naukowego, wydana w formule open access, nie jest odnotowana (czyli udostępniona) przez żaden katalog biblioteczny w Polsce spośród katalogów należących do NUKAT lub „połączonych” przez Bibliotekę Narodową [dostęp 18.04.2024]. Por. też: <https://codingdavinci.de/de/projekte/bilderbazar-eine-zeitmaschine> [dostęp 18.04.2024].

<sup>6</sup> Jak się zdaje, przydatne tu byłoby pojęcie granuli informacyjnej i obliczeń granularnych, por. Borowska (2023, zwł. s. 30 i następane). Pojęcie granulacji (uziarnienia) występuje przede wszystkim na gruncie mechaniki; w refleksji humanistycznej może ono stanowić metaforę wytwarzania według ścisłych, niepotocznych kryteriów najmniejszych jednostek sygnału czy informacji, tak by tworzyły one jednorodny zbiór.

wisu społecznościowego nie jest dokładnie tym, czego chce świat akademicki, biblioteki publiczne w końcu pewnie też nie, choć one szybciej by uległy. Ale nie ulegną, ponieważ nie otrzymają finansowania, które pozwoliłoby coś takiego uruchomić. Trudno też zresztą wyobrazić sobie budżety bibliotek uczelnianych, udzielających swoim użytkownikom nieodpłatnie wszystkich dostępnych w skali światowej zasobów, odpowiadających nie tylko ich zapytaniom, ale także wyrażonym i niewyrażonym intuicjom. Serwisy typu Academia czy ResearchGate mają przecież o wiele węższe i mniej różnorodne „zbiory” niż to, co sugerujemy tutaj, a przecież są płatne. Możliwe są jednak być może jakieś formuły pośrednie.

Tym samym skonstatować można, że biblioteka traci „swoje” zbiory, co nie znaczy, że się ich wyzbywa, po prostu w każdej chwili można poza nie wyjść, a użytkownika nie interesuje, skąd płyną udostępniane mu w jednolitej postaci teksty. Obecne wyszukiwarki zintegrowane oraz zagregowane usługi, takie jak [Polska Platforma Medyczna](#), zdają się wskazywać kierunek, w jakim być może nastąpi rozwój.

Zaś po drugiej stronie „lady” (?!) bibliotekarz też wykorzysta SI do rozwiązywania problemów. Przykładem może być organizacja zbioru pod kątem zapotrzebowania na określoną liczbęostępów. Ilu użytkowników będzie chciało równocześnie korzystać z cyfrowej wersji podręcznika medycznego? Prawdziwie funkcjonalny i równocześnie optymalny ekonomicznie model to konstrukcja skomplikowana. Nie, nie powiemy „ta książka będzie potrzebna z pięcioma równoległymi dostępami”, bo dziś wystarczy jeden, a w przyszłym tygodniu dziesięć to będzie za mało. Musimy zatem doprowadzić nasze sylabusy (a raczej ich nowe, zaawansowane wydania) do takiej postaci użytkowej, by obsługiwały rytm pracy każdej grupy studenckiej w ciągu semestru i umożliwiały (sztucznej inteligencji?) prognozowanie zapotrzebowania na każdy tekst. I to – uwaga – zapewne nie na jednostkę tekstu typu książka czy artykuł, ale na jednostkę tekstu typu rozdział, fragment rozdziału, a może dane i informacje do jakiegoś tematu, zbierane z całości narracji (przyjmując, być może błędnie, że jednostka tekstu w zbiorze będzie nadal wyszukiwalna poprzez taki czy inny identyfikator obiektu cyfrowego lub opis bibliograficzny). I ta machina obsługująca grupowe czytelnictwo nie może nie uwzględniać zmiennych rytmów pracy czy indywidualnych sposobów radzenia sobie z tekstem, nie może blokować czegokolwiek czytelnikom typu *early bird* czy karać spóźnialskich. Innymi słowy, nasz przyszły clearinghouse będzie się uczył na wszystkich odmiennościach i zmianach, wprowadzanych w toku nauczania i samouctwa, których biblioteka jako instytucja dotychczas nie obserwowała, nie widząc takiej potrzeby albo nie będąc w stanie tego się podjąć. I to wszystko będzie się działo z troską o racjonalizację wydatków i samo się będzie regulowało, tak by nie „leżały” bezużytecznie nieskonsumowane, lecz opłacone dostępy.

Wszystkie takie zabiegi są od dawna potrzebne również książkom drukowanym. Piśmiennictwo bibliotekoznawcze traktuje to w ten sposób, jakby jedyną decyzją do podjęcia była obecność lub nieobecność w zbiorze wielu egzemplarzy tego samego tytułu w tym samym wydaniu. O księgozbiorach wieloegzemplarzowych nie wspomina najważniejszy polski podręcznik (Tokarska, 2013). Wieloegzemplarzowość zbiorów dydaktycznych w bibliotekach uczelnianych jest oczywistością, ale nie widać opracowań, które pomagałyby określać zapotrzebowanie za pomocą parametrów. Na myśl przychodzi tu zróżnicowanie liczby egzemplarzy ze względu na liczbę studentów, ale także cenę egzemplarza, prawdopodobieństwo trwałości merytorycznej i użytkowej zakwalifikowanych wydań czy najtrudniejsze do określenia: stopień niezbędności zawartych w książce tekstów oraz termin i czas potrzebny do wykorzystania książki. (Przez kogo – przez przeciętnego użytkownika? Jak to liczyć?). Poszerzanie do-

stępu jest związane z wydatkami oraz wysiłkiem logistycznym, zatem bibliotekarz działa tu zgodnie ze znaną zasadą *just in time*. Półki pełne podręczników, które stoją sobie przez większość roku w długich rzędach, po czym nagle znikają – i wtedy okazuje się, że ich było za mało – to chyba nie jest rzadki widok. Jasne, że dobór parametrów stanowi tu zadanie dla inteligencji naturalnej i niekoniecznie wybitnej, ale ich nastroszenie wymaga, jak w każdym rachunku ekonomicznym, równoczesnego panowania nad tyloma zmiennymi, że sztuczna inteligencja miałaby tu zapewne coś do zrobienia. Oczywiście tylko w wielkiej uczelni. Co oznacza zmierzch uczelni niewielkich czy przynajmniej niezrzeszonych w konsorcjach, głęboko penetrujących ich procedury i struktury organizacyjne.

## Od tekstu do treści

No i wreszcie sam bibliotekarski matecznik: opracowanie rzeczowe. Coraz mniej potrzebne, skoro możemy wnikać w głąb pełnych tekstów i penetrować ich zawartość. Coraz mniej możliwe, skoro piśmiennictwo wciąż się różnicuje, a wraz z przyrostem wiedzy terminologia wydaje się wymykać wszelkiej standaryzacji. Zauważmy interdyscyplinarny charakter wszystkich znaczących prac naukowych – od artykułu medycznego, pisanego przez zespoły o nieograniczonej niemal liczebności, składające się z lekarzy wielu specjalności i nie tylko lekarzy, a kończąc na humanistach, których coraz trudniej przypisać do konkretnej dyscypliny, choć administracje i politycy po staremu spierają się, jakie są dyscypliny i jak je oddzielić od siebie.

Tradycja opracowania rzeczowego każe mieć kontrolowane słowniki, tak by do jednego zagadnienia prowadził (umownie) jeden komunikat językowy i by użytkownik mógł sprawności informacyjne ukształtowane w jednej bibliotece przenieść na drugą. Takie zadania da się całkowicie lub częściowo skomputeryzować; tak też stało się z tezaurem medycznym i rolniczym w Stanach Zjednoczonych (Golub, 2021, s. 709). Możemy zobaczyć zadanie dla sztucznej inteligencji w odejściu od standardowych, wielkich kartotek. Niechże ona raczej profiluje nam użytkowników i dla nich dobiera słowa kluczowe jako punkty dostępu do poszczególnych tekstów. Skromne doświadczenia niżej podpisanego, na tyle świeże, by wciąż go zdumiewać, zdają się podpowiadać, że każdy użytkownik potrzebuje książki do innych celów, ale nie bardzo wie, do jakich; trzeba mu je podsunąć. „Choroby ludzi” jako temat książki zatytułowanej *Patologia*<sup>7</sup> to rozwiązanie nikomu do niczego nieprzydatne. Żadne „deskryptory” nic tu nie pomogą. Statystyczny użytkownik, któremu ta książka jest potrzebna „do nauki”, i tak ją rozpozna i zamówi. Natomiast równie statystyczna użytkowniczka, np. pielęgniarka, która będzie szukała w niej materiału do pracy magisterskiej, będzie potrzebowała jednego słowa kluczowego – ale spomiędzy dziesiątków odnoszących się do tej samej pracy. Skorzystają z nich być może inne koleżanki czy koledzy.

Rzecz jasna stosunkowo proste algorytmy od dawna potrafią obsłużyć statystykę tekstu. Wiele serwisów ją prowadzi. Jedną z kilku przyjemności Google Books jest statystyka słów w książkach udostępnianych w postaci pełnotekstowej. Wybrawszy takie np. *Lectures on general pathology* z 1889 r.<sup>8</sup> możemy zajrzeć do listy *Kluczowych wyrazów i wyrażeń*, ułożonej w ten sposób, że te częściej wymieniane w tekście stają się większe, to znaczy są wyrażone pismem o większym stopniu (w kilku przedziałach). Dowiadujemy się w ten sposób, że w tekście najczęściej pojawia się *inflammation* i *thrombus*, nieco rzadziej *anaemia*, *capillaries*, *haemorrhage* i *heart*, i tak dalej. Faktycznie osobliwa to hierarchia

---

<sup>7</sup> Np. GUPTA, M.E.P. (2002). *Patologia*. Wrocław: Edra Urban & Partner.

<sup>8</sup> CONHEIM, J. (1889). *Lectures on general pathology, a handbook for practitioners and students*. London: New Sydenham Society.



w porównaniu z dzisiejszymi podręcznikami z tej dziedziny. Z kolei w zupełnie innej książce, współczesnej i humanistycznej, poświęconej historii edukacji seksualnej w Polsce<sup>9</sup>, „największe litery” to *homosexuality, Lew-Starowicz, Poland, sexologist*, na drugim miejscu *hormonal contraception, girls, masturbation, planned parenthood, pornography, sex education*. Przy takich ćwiczeniach terminy, które może i wydawałyby się nam równie ważne w takim tekście, są nieco rzadziej notowane (np. *boys, Catholic, condoms, marriage, Wisłocka*), ponieważ występują wielokrotnie, ale w różnych przypadkach, i nie „liczą się” jako jeden. Dotyczy to zwłaszcza fleksyjnego języka polskiego. Zapewne dałoby się łatwo wysubtelnić algorytm umożliwiający takie odławianie słów kluczowych z kontekstu, z założonym skalowaniem ich ważności i ilościową granicą przyjętą dla każdego prezentowanego zbioru wyrażeń.

Oczywiście jesteśmy tu jeszcze bardzo daleko od sztucznej inteligencji. Niemniej jednak widać, iż dostępny jest wgląd w tekst w celu rozpoznania jego tematyki (*aboutness*) z pominięciem indywidualnych, subiektywnych oględzin dokonywanych przez bibliotekarza uzbrojonego w kartotekę haseł wzorcowych. Tekst, który się sam indeksuje, stosując statystykę użytych w nim wyrażeń, mógłby mieć następcę w postaci tekstu, który zestawia charakteryzujące go terminy spoza kontekstu, dobierając adekwatne słownictwo, w tym wyrażenia abstrakcyjne i teoretyczne, zgromadzone dzięki zastosowaniu wielkiego modelu językowego (LLM) w fazie przygotowywania programu do pracy. Należy mieć nadzieję, że będzie to rozwiązanie gorączkowo poszukiwane dla technik automatycznego opracowania rzeczowego. Będą one bezbronne wobec każdego tekstu, którego treść nie ogranicza się do sumy użytych w nich wyrażeń. Zwłaszcza prac nawiązujących do szerszych pojęć, takich jak style, poglądy, ideologie, zjawiska historyczne, które nie muszą być nazwane w tekście, by rzutować na jego przekaz. Zjawisko to traktujemy tu intuicyjnie, ponieważ jest oczywiste. Nie nawiązujemy do opublikowanych badań, ale jest ono znane twórcom wyszukiwarek i ma swoje piśmiennictwo (zwłaszcza z zakresu SEO, czyli *search engine optimization*).

Maszynowe opracowanie rzeczowe stosowane przez Niemiecką Bibliotekę Narodową (Deutsche Nationalbibliothek, DNB) nie wchodzi chyba jeszcze w tę fazę. Zrobiono jednak wiele i odnotowano to w nowym niemieckim podręczniku (Sühl-Strohmer i Tappenbeck, 2024, s. 17). Od kilku lat w katalogu DNB widzimy *schlagwörter* (słowa kluczowe, hasła przedmiotowe) opatrzone gwiazdką, którym w odróżnieniu od pozostałych przysługuje cecha *maschinell ermittelt*. Maszynowo wytwarzane są także symbole Klasyfikacji Dziesiątej Deweya. Widać, że biblioteka stosuje system kwalifikowania opisów do automatycznego wytworzenia, nie wszystkie zatem jednostki bibliograficzne mu podlegają. Pomijane – tzn. opisywane tradycyjnie, przez ludzi – są prace z zakresu humanistyki. Teksty historyczne, polityczne czy prawnicze okazują się za trudne dla programu komputerowego, od SI jesteśmy tu wciąż dość daleko. Często triumf człowieka nad maszyną okazuje się tu mało imponujący, jak np. wówczas, gdy prawnicza cegła *Europäisches Steuerrecht* (2022) otrzymuje hasła: *Europäische Union; Steuerrecht*. Przypomina to przypadek filozofa koreańsko-niemieckiego nazwiskiem Byung-Chul Han, którego książce *Hyperkulturalität: Kultur und Globalisierung* (2005) dostały się boleśnie oczywiste hasła *Kultur; Globalisierung*. Wiele książek, np. opublikowanych za granicą lub opublikowanych w ramach serii, haseł przedmiotowych nie otrzymuje wcale. Podobnie beletrystyka (nie do wiary!) i prace popularne.

---

<sup>9</sup> KOŚCIAŃSKA, A. (2021). *To see a moose: the history of Polish sex education*. New York: Bergham Books.

Automatyczne tematowanie w DNB bywa jednak stosowane także do prac z pogranicza nauk przyrodniczych czy medycznych i społecznych; nie unika się tu nowych nurtów piśmiennictwa. Tekst elektroniczny o tytule *Funktionelle Stimmtransition von Trans\*Frauen* (2024) otrzymał *schlagwörter*: *Transsexualität\**; *Transgender\**; *Frau\**; *Logopädie\**; *KulturGut Holzhausen VOICES Internationales Stimmenfestival (Nieheim)\**; *Diskriminierung\**. Nawet dla książki z zakresu (*Sachgruppe*) teologii chrześcijańskiej oraz nauk społecznych i antropologii pt. *Biblische Schöpfungsordnung in der Spannung von Genderideologie und sexueller Vielfalt* (2023) hasła wyszukał automat: *Bibel*; *Transgender\**; *Sexualethik\**; *Transsexualität\**; *Sexualverhalten\**; *Geschlechterrolle\**; *Exegese\**. Podobnie jak dla typowej dysertacji biologicznej *Aufklärung der tRNAHis-Prozessierung in Eukaryoten und Charakterisierung von anzestralen tRNA-Nukleotidyltransferasen* (2021): *Prozessierung\**; *Eukaryoten\**; *Sequenz\**; *Enzym\**; *In vitro\**.

W tym miejscu ponownie oddaliśmy się od przyjętego zadania, jakim miało być ukazywanie obszarów praktyki bibliotecznej, podatnych na przejmowanie przez narzędzia sztucznej inteligencji. Filozofowie i historycy, politologowie i reporterzy będą jednak dalej zanurzeni w fali słów, które wywołuje każda ich wypowiedź, i zrezygnują w pewnym momencie z dociekania, czy obieg ten obsługuje w wystarczającym stopniu sieci neuronowe naturalne, czy też trzeba przystosować do współpracy także sztuczne. Nowe systemy charakterystyki rzeczowej, oparte na automatycznej analizie tekstu, wygenerują wiele szumu wynikającego z niezrozumienia analizowanych i transmitowanych wyrażań i kontekstów, nie będą jednak w tym względzie nic mniej sprawne niż ludzie i narzędzia, którymi posługują się teraz. Przyniosą także nowe niebezpieczeństwa. Pokusa profilowania będzie wzbudzona i ciągle pobudzana przez inteligencję generatywną; nie wiedząc gdzie i jak się zatrzymać, nasi następcy zaczną do każdej książki dołączać inną charakterystykę rzeczową – zależną od tego, kto jej będzie szukał. Może zresztą się okazać, że tego właśnie było nam trzeba. Zapewne czeka nas zmiana o wiele głębsza niż tzw. automatyzacja bibliotek, będąca na początku – i w dużym stopniu pozostająca do dziś – komputeryzacją jej katalogów.

## Bibliografia:

1. BARON, N. S. (2023). *Who wrote this? How AI and the lure of efficiency threaten human writing*. Stanford: Stanford University Press.
2. BOROWSKA, K. (2023). *Obliczenia granularne w przetwarzaniu danych niebalansowanych* [Rozprawa doktorska, Politechnika Białostocka] [online]. Białystok. [Dostęp 18.03.2024]. Dostępny w: <https://bazawiedzy.pb.edu.pl/info/phd/BUTda295628a9c9455ebcaa68335ab5c713>.
3. CHOMSKY, N. (2023). The false promise of Chat GPT. *The New York Times* [online]. 08.03. [Dostęp 18.03.2024]. Dostępny w: <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>.
4. CRONIN, J. (2024). Noam Chomsky On Artificial Intelligence. A ChatGPT Rebuttal. W: *LinkedIn* [online]. 24.01. [Dostęp 18.03.2024]. Dostępny w: <https://www.linkedin.com/pulse/noam-chomsky-artificial-intelligence-chaptgpt-rebuttal-jon-cronin-7rze/>.
5. DUKAJ, J. (2023). Się pisze. *Książki : Magazyn do Czytania*, nr 4, s. 12–19.
6. ENIS, M. (2021). OverDrive refines readtelligence, a new AI-driven collection management tool. *Library Journal* [online]. Oct. 07. [Dostęp 30.01.2024]. Dostępny w: <https://www.libraryjournal.com/story/overdrive-refines-readtelligence-a-new-ai-driven-collection-management-tool>.
7. FRANKE, J. (2024). OMNIS 2 i pół. Karawana jedzie dalej. Materiał niepublikowany.
8. GOLUB, K. (2021). Automated subject indexing: an overview. *Cataloging and Classification Quarterly* [online]. Nr 8, s. 702–719. [Dostęp 13.03.2024]. Dostępny w: <https://doi.org/10.1080/01639374.2021.2012311>.
9. HOLLENDER, H. (2021). Cyfrowe wzmoczenie w zbiorach specjalnych. W: Puchalski, J. (red.). *Książki mają swoją historię: studia ofiarowane Profesor Barbarze Bieńkowskiej*. Warszawa: Wydaw. SBP, s. 410–446.

10. JANICKA, E. (2017). O kontrkulturę – tu i teraz. Wystąpienie na Nadzwyczajnym Zjeździe Etnologów i Antropologów Polskich „Antropologia przeciw dyskryminacji”. *Lud* [online]. Nr 101, s. 93–95. [Dostęp 18.03.2024]. Dostępny w: <https://apcz.umk.pl/LUD/article/view/lud101.2017.01d>.
11. KRUEGER, A. (2024). Why the Russia sanctions are failing. *Project Syndicate* [online]. Jan. 18. [Dostęp 30.01.2024]. Dostępny w: <https://www.project-syndicate.org/commentary/how-russia-circumvented-western-sanctions-by-anne-o-krueger-2024-01>.
12. LEE, J. Y. (2023). Can an artificial intelligence chatbot be the author of a scholarly article? *Journal of Educational Evaluation for Health Professionals* [online]. Nr 6. [Dostęp 18.03.2024]. Dostępny w: <https://doi.org/10.3352/jeehp.2023.20.6>.
13. MYKOWIECKA, A. (2007). *Inżynieria lingwistyczna: komputerowe przetwarzanie tekstów w języku naturalnym*. Warszawa: Wydaw. Polsko-Japońskiej Wyższej Szkoły Technik Komputerowych. (Podręczniki akademickie, t. 27).
14. OYELUDE, A. (2021). AI and libraries: trends and projections. *Library Hi Tech News* [online]. R. 38, nr 10, s. 1-4. [Dostęp 30.01.2024]. Dostępny w: <https://doi.org/10.1108/LHTN-10-2021-0079>.
15. OŚWIADCZENIE (2023). *Oświadczenie CEATL w sprawie wykorzystania sztucznej inteligencji* [online]. Stowarzyszenie Tłumaczy Literatury. Aktualności. [Dostęp 13.03.2024.] Dostępny w: <https://stl.org.pl/oswiadczenie-ceatl-w-sprawie-wykorzystania-sztucznej-inteligencji/>.
16. ROTKIEWICZ, M. (2024). Więcej czatu! Jak sztuczna inteligencja wspomaga badaczy. *Polityka*, nr 17 (3461), 14.04-23.04, s. 56-58.
17. SAMPLE, I. (2023). Science journals ban listing of ChatGPT as co-author on papers. *The Guardian* [online]. [Dostęp 18.03.2024]. Dostępny w: <https://www.theguardian.com/science/2023/jan/26/science-journals-ban-listing-of-chatgpt-as-co-author-on-papers>.
18. SÜHL-STROHMENGER, W., TAPPENBECK, I. (red.) (2024). *Praxishandbuch Wissenschaftliche Bibliothekar: innen. Wandel von Handlungsfeldern, Rollen und Perspektiven im Kontext der digitalen Transformation* [online]. Berlin: De Gruyter, Saur. [Dostęp 18.03.2024]. Dostępny w: <https://doi.org/10.1515/9783110790375>.
19. SZOT, W. (2023). Literatura, czyli małż. *Magazyn Gazety Wyborczej*, 29-30 lipca, s. 12-13.
20. TANZI, N. (2023). How can AI be used in libraries? *The Digital Librarian* [online]. Jan. 23. [Dostęp 18.03.2024]. Dostępny w: <https://the-digital-librarian.com/2023/01/23/artificial-intelligence-in-the-library/>.
21. THORP, H. (2023). ChatGPT is fun, but not an author. Editorial. *Science* [online]. R. 379, nr 6630, s. 313. [Dostęp 13.03.2024]. Dostępny w: <https://www.science.org/doi/10.1126/science.adg7879>.
22. TOFFLER, A. (1974). *Szok przyszłości*. Przeł. W. Osiatyński, E. Grabczak-Ryszka, E. Woydyłło. Warszawa: Państwowy Instytut Wydawniczy.
23. TOKARSKA, A. (red.) (2013). *Bibliotekarstwo*. Warszawa: Wydaw. SBP.
24. WHEATLEY, A., HERVIEUX, S. (2019). Artificial intelligence in academic libraries: an environmental scan. *Information Services & Use* [online]. R. 39, nr 4, s. 347-356. [Dostęp 18.03.2023]. Dostępny w: <http://dx.doi.org/10.3233/ISU-190065>.

---

HOLLENDER, H. (2024). Sztuczna inteligencja i prawdziwe biblioteki: szok przyszłości i co dalej? *Biuletyn EBIB* [online]. Nr 2(213), Sztuczna inteligencja w bibliotekach – rewolucja czy ewolucja. [Dostęp 22.05.2024]. ISSN 1507-7187. Dostępny w: <https://ebibojs.pl/index.php/ebib/article/view/910>.