

prof. dr. iur. Iza Razija Mešević
vanredna profesorica, Univerzitet u Sarajevu – Pravni fakultet

UMJETNA INTELIGENCIJA I MEDIJI: THE NEW YORK TIMES COMPANY PROTIV KOMPANIJA *OpenAI* I MICROSOFT

Rezime: U radu se tematizuje prva tužba koju je jedna medijska organizacija pokrenula protiv predstavnika AI-industrije, naime New York Times Company protiv kompanija OpenAI i Microsoft. U pozadini ovog spora leži ključno pitanje obuke modela generativne umjetne inteligencije na autorsko-pravno zaštićenim sadržajima medijskih izdavača, bez traženja dozvole i bez plaćanja naknade, te kreiranja sintetičkih autputa od strane AI-proizvoda (npr. ChatGPT) koji gotovo u potpunosti ili u velikoj mjeri reprodukuju ta autorska djela, ili predstavljaju njihove prerade. U radu su definisani osnovni pojmovi i kategorije ključne za razumijevanje tehnologije u srži spora (npr. generativna umjetna inteligencija, temeljni modeli i veliki jezički modeli). Pored toga, problematizovane su i moguće sigurne luke za AI-developere, u ovom kontekstu u Evropskoj uniji, u smislu ograničenja za rudarenje teksta i podataka iz Direktive (EU) 2019/790, a kojem "vjetar u leđa" daje tekst AI Uredbe EU, te u Sjedinjenim Američkim Državama, u smislu primjene pravne doktrine *fair use*, posebno kroz "transformativnu upotrebu" djela. Autorica detaljno izlaže okolnosti nastanka i sam sadržaj spora, te se osvrće i na uporedne postupke u domenu odnosa između autorskog prava i umjetne inteligencije, a koji bi mogli dati neke smjernice u pogledu toga u kojem bi pravcu mogao ići njegov ishod. Pored toga, rad skreće pažnju i na pozitivne primjere uređivanja odnosa, pa čak kreiranja i partnerstava između velikih medijskih organizacija (npr. Associated Press, Le Monde ili News Corp) i AI-giganta OpenAI u pogledu korištenja njihovih sadržaja u obuci GPT-modela, kao i u kontekstu tehnološke saradnje.

Ključne riječi: Umjetna inteligencija. – Autorsko pravo. – Mediji. – The New York Times. – OpenAI.

1. UVOD

U periodima disruptivnih, no istovremeno i transformativnih tehnoloških promjena, okviri autorskopravne zaštite često iznova bivaju podvrgnuti preispitivanju u pogledu njihove fleksibilnosti, ali čak i opravdanosti. Fenomen umjetne inteligencije zasigurno predstavlja jednu takvu tehnologiju (odnosno cijeli njen skup), koja zbog svoje izuzetne kompleksnosti postavlja niz izazova postojećim pravnim shvatanjima u okviru ove pravne oblasti i nameće kroz svoj rapidni razvoj brojna pravna pitanja koja se u tom sadržaju

postavljaju po prvi put pred nosioce prava, sudove i zakonodavce. S jedne strane, nosioci prava bivaju konfrontirani s novim načinima upotrebe svojih djela, a bez toga da im postojeći pravni okvir autorskopravne zaštite neminovno pruža jasan odgovor na pitanje da li su te upotrebe dužni tolerisati (ograničenja prava) ili im se mogu suprotstaviti i za njih tražiti naknadu. S druge strane, stepen složenosti, ali i netransparentnosti procesa u kojima dolazi do tih upotreba, samo otežava situaciju. Jedan od primjera takvog korištenja je postupak obuke modela generativne umjetne inteligencije.

Iako su već i autori pojedinci¹ podigli svoj glas u pravcu zaštite svojih prava zbog upotrebe njihovih djela u ovom tehnološkom procesu, najveći odjek u javnosti zasigurno su imale tužbe predstavnika medijske industrije u Sjedinjenim Američkim Državama, s obzirom na obim autorskopравnih sadržaja kojima raspoložu, a koje su AI-kompanije koristile, i finansijskih efekata koje je takvo postupanje imalo na njihovo poslovanje. Vjerovatno je da će upravo ishodi tih postupaka dati odgovor na pitanje da li se obuka tih modela na tuđim autorskim djelima i produciranje sintetičkih autputa koji sadržavaju ta djela, njihove dijelove ili prerade može podvesti pod doktrinu *fair use*. Drugim riječima, pokazat će stepen fleksibilnosti trenutnih normi autorskog prava u tom dijelu svijeta s obzirom na (barem) ovaj izazov umjetne inteligencije.

U tom kontekstu, spor u koji su uprte sve oči jest onaj između The New York Times Company (dalje u tekstu: "NYTC"), koja je prva velika američka medijska organizacija koja je tužila² kompaniju OpenAI (u smislu skupine pravnih lica) i Microsoft Corporation.³ Razlog tužbe je navodno neovlašteno korištenje miliona članaka i drugih sadržaja koje je objavio The New York Times (dalje u tekstu: "NYT") i povezane platforme i publikacije, u svrhu obuke modela generativne umjetne inteligencije, a čiji se autputi sada takmiče s novinskim izdanjima. Shodno, u očekivanju presude, ovaj predmet i njegovo ozračje zaslužuju kratki naučni osvrt.

U tekstu rada će najprije biti tematizovana generalna klima odnosa između medijskih organizacija i ove dvije kompanije, a potom biti pojašnjeni osnovni tehnološki pojmovi i kategorije umjetne inteligencije od značaja za razumijevanje suštine spora. Potom će biti dat kratak pregled postojećeg (Evropska unija) i potencijalnog (Sjedinjene Američke Države) pravnog okvira za obuku modela generativne umjetne inteligencije, nakon čega će biti

1 Vid.: 4.1. Okolnosti spora.

2 The New York Times Company *protiv* Microsoft Corporation, OpenAI, Inc., OpenAI LP, OpenAI GP, LLC, OpenAI, LLC, OpenAI OPCO, LLC, OpenAI Global, LLC, OAI Corporation, LLC i OpenAI Holdings, LLC (predmet br.: 1:23-cv-11195), Okružni sud Sjedinjenih Američkih Država za južni distrikt New Yorka, tužba, 27. decembar 2023., https://nytco-assets.nytimes.com/2023/12/NYT_Complaint_Dec2023.pdf, 20. mart 2024., (NYTC v. Microsoft i OpenAI).

3 M. M. Grynbaum, R. Mac, "The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work", *The New York Times*, 23. decembar 2023., <https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html>, 27. januar 2024.

izložen sam sadržaj tužbe u njenom širem kontekstu i dat osvrt na trenutno stanje postupka.

2. UMJETNA INTELIGENCIJA I MEDIJI – FRIENDS OR FOES?

Na prvi pogled djeluje, a posebno zbog pozicije i reputacije kompanije NYTC u globalnim medijskim krugovima, kao da je njena tužba čak moguće imala dejstvo inicijalne kapsule za druge izdavače da krenu u pravne bitke.

Naime, već u februaru 2024. godine još tri medijske organizacije (The Intercept, Raw Story i AlterNet) uložile su dvije tužbe⁴ od kojih je u jednoj, uporedivo s tužbom NYTC-a, kao tuženi imenovan ne samo OpenAI već i Microsoft. Navodi koji se iznose na teret ove dvije AI-kompanije također su podudarni, uz razliku da je odštetni zahtjev postavljen konkretnije, na način da u slučaju dodjele statutarne štete, ona ne iznosi manje od 2500 dolara po povredi.

Samo par mjeseci kasnije, u aprilu 2024. godine, čak osam dnevnih novina u vlasništvu investicijskog fonda Alden Global Capital također je uložilo tužbu protiv istih kompanija i to iz identičnih razloga kao i NYTC.⁵ I navedene novine zastupaju stav da se radi o neovlaštenom i parazitskom korištenju sadržaja, a koje izdavačima uskraćuje prihod od naknada za korištenje i korisničkih pretplata. Kompanijama OpenAI i Microsoft se, kao i u slučaju tužbe NYTC-a, stavljaju na teret povreda autorskih prava, nelojalna konkurencija i razvodnjavanje žigova.

Međutim, činjenica je da su na drugoj strani mnoge velike medijske organizacije u prethodnoj godini dana uspjele naći zajednički jezik s jednim od tuženih, naime s kompanijom OpenAI, i zaključile niz ugovora o korištenju njihovih sadržaja za obuku modela generativne umjetne inteligencije.⁶

U tom kontekstu je prednjačio The Associated Press (dalje u tekstu: "AP") koji je još u julu 2023. godine sklopio dogovor s kompanijom OpenAI⁷ o davanju dozvole za korištenje AP-ove arhive vijesti od 1985. godine nadalje za treniranje njihovih GPT-modela, a AP je zauzvrat dobio pristup OpenAI-ovoj tehnologiji i stručnosti u pogledu proizvoda koje kreira.

4 B. Brittain, "OpenAI hit with new lawsuits from news outlets over AI training", *Reuters*, 28. februar 2024, <https://www.reuters.com/legal/litigation/openai-hit-with-new-lawsuits-news-outlets-over-ai-training-2024-02-28/>, 19. maj 2024.

5 U pitanju su: *The New York Daily News*, *The Chicago Tribune*, *The Orlando Sentinel*, *The Sun Sentinel of Florida*, *The San Jose Mercury News*, *The Denver Post*, *The Orange County Register* i *The St. Paul Pioneer Press*. K. Roberts, "8 Daily Newspapers Sue OpenAI and Microsoft Over A.I.", *The New York Times*, 30. april 2024., <https://www.nytimes.com/2024/04/30/business/media/newspapers-sued-microsoft-openai.html>, 12. maj 2024.

6 Roberts, (2024).

7 The Associated Press, "ChatGPT-maker OpenAI signs deal with AP to license news stories", *The Associated Press*, 13. jul 2023., <https://www.ap.org/media-center/ap-in-the-news/2023/chatgpt-maker-openai-signs-deal-with-ap-to-license-news-stories/>, 27. mart 2024.

Nakon ovog inicijalnog ugovora, u decembru 2023. godine OpenAI je zaključio sporazum s njemačkim izdavačem Axel Springerom, pod čijim krovom su publikacije kao npr. Business Insider, Bild, Welt i Politico.⁸ Naime, prema sadržaju sporazuma, *chatbot* ChatGPT će od prvog kvartala 2024. godine moći pružiti na neisključivom osnovu sažetke sadržaja iz publikacija Axel Springera kao odgovore, odnosno autpute na upite korisnika. Ono što je od posebnog značaja u ovom kontekstu je, s jedne strane, da će ti sažeci moći uključivati materijale i iz tekstova za čije bi čitanje korisnicima inače bila potrebna pretplata i da će sadržavati uputu na publikaciju Axel Springera koja je izvor informacija, te i sam link na cijeli članak koji se sažima. S druge strane, prema izvoru upoznatom sa sporazumom, na njegovom temelju će sadržaj publikacija Axel Springera dobiti "favorizovan položaj" u rezultatima pretraživanja ChatGPT-a s ciljem podrške u privlačenju prometa i prihoda od pretplata na publikacije Axel Springera, a OpenAI će također platiti i naknadu za korištenje sadržaja publikacija Axel Springera za obuku svojih velikih jezičkih modela.

Također, u martu 2024. godine OpenAI se sporazumio s medijskim organizacijama Le Monde i Prisa Media, koja izdaje publikacije kao što su *El País* i *Cinco Días*, s ciljem da dovede sadržaje, a posebno vijesti na francuskom i španskom jeziku, na ChatGPT.⁹ Uporedivo sadržaju ugovora s Axel Springerom, saradnja podrazumijeva korištenje sadržaja ovih izdavača za obuku modela generativne umjetne inteligencije. Također, korisnici ChatGPT-a će dobijati sažetke vijesti iz publikacija ovih izdavača uz navođenje izvora i poboljšane linkove do originalnih članaka, što će korisnicima dati mogućnost pristupa dodatnim informacijama ili srodnim člancima s web-stranica ovih medijskih organizacija.

Nadalje, i *Financial Times*, koji je u vlasništvu japanske tvrtke Nikkei,¹⁰ ušao je u strateško partnerstvo s OpenAI-em i postigao dogovor o korištenju sadržaja ove publikacije za obuku svojih GPT modela uporedivog sadržaja kao i druge medijske organizacije. Pored toga, ovo partnerstvo podrazumijeva i saradnju na razvoju novih proizvoda umjetne inteligencije i funkcija za korisnike *Financial Timesa*.¹¹

Konačno, u maju 2024. godine OpenAI je pravno uredio odnose i s medijskim organizacijama Dotdash Meredith, koja izdaje publikacije *People*,

8 Reuters, "Global news publisher Axel Springer partners with OpenAI in landmark deal", *Reuters*, 13. decembar 2023., <https://www.reuters.com/business/media-telecom/global-news-publisher-axel-springer-partners-with-openai-landmark-deal-2023-12-13/>, 19. maj 2024.

9 OpenAI, "Global news partnership: Le Monde and Prisa Media", *OpenAI*, 13. mart 2024b., <https://openai.com/index/global-news-partnerships-le-monde-and-prisa-media/>, 5. maj 2024.

10 Reuters, "OpenAI to use FT content for training AI models in latest media tie-up", *Reuters*, 30. april 2024., <https://www.reuters.com/technology/financial-times-openai-sign-content-licensing-partnership-2024-04-29/>, 12. maj 2024.

11 OpenAI, "We're bringing the Financial Times' world-class journalism to ChatGPT", *OpenAI*, 29. april 2024., <https://openai.com/index/content-partnership-with-financial-times/>, 19. maj 2024.

Better Homes & Gardens, *Investopedia*, *Food and Wine* i *InStyle*,¹² i News Corp,¹³ koja je u vlasništvu medijskog giganta Rupert Murdocha i publikuje, između ostalog, *The Wall Street Journal*, *The New York Post*, *MarketWatch* i *Barron's*. U budućnosti, Dotdash Meredith će moći koristiti AI modele OpenAI-a za nadogradnju svog proizvoda za usmjeravanje oglasa "D/Cipher", koji pomaže oglašivačima da pronađu potrošače kroz sadržaj koji ti korisnici gledaju i to bez ličnih identifikatora poput kolačića (*cookies*). Zauzvrat, Dotdash Meredith će dati dozvolu OpenAI-u za korištenje svog sadržaja za modele umjetne inteligencije na uporedive načine kao što su to uradili i drugi izdavači. Uslovi sporazuma s News Corp omogućavaju OpenAI-u da koristi trenutne i arhivirane vijesti iz glavnih informacijskih publikacija News Corpa, uključujući one u Sjedinjenim Američkim Državama (dalje u tekstu: "SAD"), Ujedinjenom Kraljevstvu i Australiji, kao i MarketWatch i Barron's. Ipak, sporazum ne uključuje sadržaj iz drugih sektora djelatnosti News Corpa, kao što su usluge digitalnih nekretnina ili HarperCollins. *The Wall Street Journal* je objavio da bi ovaj sporazum mogao biti vrijedan čak 250 miliona dolara za period od pet godina.

Shodno izloženom, zaključak da je AI-gigant OpenAI (a što podrazumijeva i njegovog glavnog investitora kompaniju Microsoft) u sveobuhvatnom odnosu antagonizma s medijskim sektorom zbog korištenja sadržaja izdavača za obuku svojih modela, na što bi na prvi pogled mogla uputiti tužba NYTCA, bio bi veoma površan. Očigledno je da značajan broj relevantnih medijskih organizacija prihvata s jedne strane neminovnost, a s druge moguće čak i nezaustavljivost razvoja AI-industrije, te preferira da pozicionira svoje sadržaje, dogovori uslove i iznose naknada za njihovo korištenje, ali uspostavi i partnerstva s ovim AI-gigantom, kako bi profitirali od same tehnologije njenom primjenom u procesu izvještavanja i nuđenja sadržaja korisnicima. Također, ovi sporazumi predstavljaju i vidljiv izraz dobre volje i stepena razumijevanja OpenAI-a za zaštitu autorskih prava tih izdavača, što donekle dekonstruiše argument njihove namjere za tzv. *free riding*.

3. GENERATIVNA UMJETNA INTELIGENCIJA I AUTORSKO PRAVO

3.1. Generativna umjetna inteligencija

U svakodnevnom razgovoru sigurno mnogi od nas barem par puta dnevno spomenu ChatGPT ili Gemini, Soru, DALL-E, Stable Difusion ili Midjourney. Međutim, kada bi se postavilo pitanje, na koji način definišemo te AI-alate ili pak na koji način oni funkcionišu, odgovori bi se sigurno veoma razlikovali.

12 E. David, "OpenAI strikes licensing deal with the magazine giant behind People", *The Verge*, 7. maj 2024., <https://www.theverge.com/2024/5/7/24151171/openai-dotdash-meredith-people-instyle-licensing>, 29. maj 2024.

13 K. Robertson, "OpenAI Strikes a Deal to License New Corp Content", *The New York Times*, 22. maj 2024., <https://www.nytimes.com/2024/05/22/business/media/openai-news-corp-content-deal.html>, 29. maj 2024.

Prilikom uvođenja definicija novih pojmova koje sa sobom donose digitalne tehnologije u nastajanju, poseban problem predstavlja pokušaj njihovog pravnog određenja, a koje bi trebalo da sadrži elementarne karakteristike tog fenomena. Naime, dok pravo ima vlastiti kategorijalni aparat koji nastoji da bude što precizniji u cilju uspostavljanja pravne sigurnosti, kod takvih tehnologija definicije pojedinih pojmova značajno su fleksibilnije i prilagodljivije rapidnom tehnološkom napretku, a uz to se u relevantnim industrijama i ne koriste neminovno jednoobrazno.

Prvu pravnu definiciju generativne umjetne inteligencije (dalje u tekstu: "GenAI") i uz to širokog pravnog dejstva, posredno nalazimo u AI-uredbi¹⁴ Evropske unije (dalje u tekstu: "EU") i to u njenom čl. 3, st. 63, a u vezi s preambulom br. 99 tog akta.

Naime, EU-zakonodavac izričito i detaljno definiše kategoriju "AI-modela opšte namjene"¹⁵ i navodi da tzv. "veliki GenAI-modeli" predstavljaju tipičan primjer AI-modela opšte namjene, s obzirom na to da omogućavaju fleksibilno generisanje sadržaja, kao što su tekst, zvuk, slika ili video, a koji se lako može prilagoditi širokom rasponu karakterističnih zadataka.

U svrhu potpunog razumijevanja problematike spora koji je tema ovog rada, od značaja je osvrnuti se i na pojam temeljnih modela (*foundation models*) i velikih jezičkih modela (*Large Language Models*).

Temeljni modeli predstavljaju jednu vrstu AI-modela opšte namjene i mogu se definisati¹⁶ kao veliki modeli mašinskog učenja koji su, principijelno pod samonadzorom, prethodno obučeni na značajnoj količini podataka (obično preuzetih s interneta), ali u pravilu ne u cilju realizovanja specifičnih zadataka, već radije za prikupljanje korisnih opštih informacija iz podataka. No, oni mogu biti prilagođeni i drugačije.

Pa tako su npr. veliki jezički [temeljni] modeli (dalje u tekstu: "LLM"), kao npr. GPT-3.5 ili GPT-4 (*Generative Pre-trained Transformer*) obučeni na velikim skupovima podataka koji se sastoje iz teksta, a s ciljem da predvide sljedeću riječ u nizu.¹⁷ Temeljni modeli također mogu generisati sadržaj, tj. biti GenAI modeli. Pa tako bi ChatGPT predstavljao AI-sistem, odnosno

14 Uredba Evropskog parlamenta i Vijeća o harmoniziranim pravilima o umjetnoj inteligenciji i izmjenama uredbi (EZ) br. 300/2008, (EU) br. 167/2013, (EU) br. 168/2013, (EU) br. 2018/858, (EU) br. 2018/1139 i (EU) br. 2019/2144 i direktiva br. (EU) 2014/90/EU, (EU) 2016/797 i (EU) 2020/1828, *Službene novine EU*, br. L, 2024/1689, 12.7.2024.

15 U pitanju je model koji pokazuje značajnu uopštenost i koji je sposoban kompetentno obavljati širok raspon različitih zadataka, bez obzira na način na koji se model stavlja na tržište, te koji se može integrisati u niz "nizvodnih" sistema ili aplikacija. Ovaj fenomen uključuje i slučajeve kada je takav AI-model obučen s velikom količinom podataka koristeći samonadzor u velikom obimu, a isključuje AI-modele koji se koriste za istraživanje, razvoj ili aktivnosti izrade prototipa prije nego što budu stavljeni na tržište (čl. 3, st. 63 AI-uredbe).

16 P. Henderson, X. Li, D. Jurafsky, T. Hashimoto, M. A. Lemley, P. Liang, "Foundation Models and Fair Use", 29. mart 2023, *Stanford Law and Economics Olin Working Paper No. 584*, <http://dx.doi.org/10.2139/ssrn.4404340>, 19. mart 2024., 4 et seq.

17 *Ibid.*

chatbot proizvod, koji je baziran na GenAI temeljnom LLM-u GPT-3.5, GPT-4, odnosno GPT-4o. Kako pojašnjava sam Microsoft,¹⁸ u srži jednog LLM-a je kompjuterski algoritam koji se naziva “transformator”. U procesu kreiranja modela dolazi do njegove obuke skupovima podataka, na način da se unose velike količine teksta kroz “transformator”. LLM odvaja uneseni tekst na “tokene” ili jedinice teksta (npr. riječi ili interpunkcijske znakove) i u tom procesu raspoznaje semantičke obrasce među njima, koje potom “kodira” u vidu brojeva koji se zovu “parametri”. Ova opsežna mreža semantičkih veza omogućava modelu da generiše tekst na prirodnom jeziku.

Ipak, kako je već istaknuto, AI-industrija ne koristi navedene termine jednoobrazno i konzistentno, kao što bi to zahtijevalo pravo, pa se LLM-ovi ponekad nazivaju i modelima opšte namjene ili temeljnim AI-modelima, jer njihove jezičke sposobnosti imaju mnoge primjene.¹⁹

3.2. Ograničenje autorskog prava u svrhu rudarenja teksta i podataka

3.2.1. Direktiva 2019/790

U EU je pitanje korištenja sadržaja autorskopravno zaštićenih materijala za obuku GenAI modela uređeno Direktivom (EU) 2019/790²⁰ na način da postoji širok konsenzus da je ta aktivnost, ukoliko podrazumijeva kreiranje kopija autorskih djela, obuhvaćena ograničenjem autorskog prava za rudarenje teksta i podataka²¹ iz njenog čl. 4. Međutim, na sadržaj ovog rješenja su upućene i mnogobrojne kritike zbog njegove navodne restriktivnosti na teret AI-industrije.²²

18 The New York Times Company *protiv* Microsoft Corporation, OpenAI, Inc., OpenAI LP, OpenAI GP, LLC, OpenAI, LLC, OpenAI OPCO, LLC, OpenAI Global, LLC, OAI Corporation, LLC i OpenAI Holdings, LLC (1:23-cv-11195), Okružni sud Sjedinjenih Američkih Država za južni distrikt New Yorka, Podnesak kojim se podržava zahtjev Microsoft Corporationa za djelimično odbijanje tužbe, 4. mart 2024. (Microsoftov zahtjev za odbacivanje), <https://www.courtlistener.com/docket/68117049/65/the-new-york-times-company-v-microsoft-corporation/>, 20. maj 2024., Allegations Of The Complaint, A. Large Language Models Are A Profound Advance In Artificial Intelligence And A Powerful Tool For Human Flourishing.

19 *Ibid.*

20 Direktiva (EU) 2019/790 Evropskog parlamenta i Vijeća od 17. 4. 2019. o autorskom pravu i srodnim pravima na jedinstvenom digitalnom tržištu i izmjeni direktiva 96/9/EZ i 2001/29/EZ, *Službene novine EU*, br. L 130/92 od 17. 5. 2019. (DSM-direktiva).

21 Ova aktivnost podrazumijeva svaku automatizovanu analitičku tehniku čiji je cilj analiza teksta i podataka u digitalnom obliku radi stvaranja informacija, koje uključuju, ali ne ograničavaju se na uzorke, trendove i korelacije. Čl. 2, st. 3 DSM-direktive.

22 Vid. npr.: T. Margoni, M. Kretchmer, “A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology”, *GRUR Int*, 8/2022, 686 *et seq.* i A. Dermawan, “Text and data mining exceptions in the development of generative AI models: What the EU member states could learn from the Japanese “no-nenjoyment” purposes?”, *The Journal of World Intellectual Property*, 1/2024, <https://doi.org/10.1111/jwip.12285>, 12. maj 2024., 52 *et seq.*

Pored ograničenja prava za rudarenje teksta i podataka iz čl. 3, koji je uspostavljen u svrhu naučnog istraživanja, EU-zakonodavac je prepoznao da i privatni i javni subjekti uveliko koriste navedene tehnike u svrhu analize velikih količina podataka i autorskih djela, između ostalog, i za razvoj novih aplikacija ili tehnologija, i uveo navedeno ograničenje u cilju uspostavljanja pravne sigurnosti i promocije inovacija.²³ Za razliku od beneficijara ograničenja iz čl. 3, korisnici ograničenja iz čl. 4 mogu reprodukcije i ekstrakcije nastale kao posljedica ovog procesa zadržati onoliko vremena koliko je to potrebno za svrhe rudarenja teksta i podataka (st. 2). Uslov koji beneficijari moraju ispuniti da bi u njemu uživali je, prije svega, zakonit pristup djelima (st. 1), što uključuje i okolnost je to djelo ili drugi zaštićeni sadržaj stavljen na raspolaganje javnosti na internetu.²⁴ Nosiocima prava se ostavlja mogućnost da izjave izričitu rezervu (*opt-out*) na radnje reprodukcije i ekstrakcije u svrhu rudarenja (st. 3), no to mora biti učinjeno na adekvatan način. Pa tako, kada je u pitanju zaštićeni sadržaj koji je dostupan javnosti na internetu, adekvatnost podrazumijeva rezervu izraženu na mašinski čitljiv način, dok u drugim slučajevima to može biti učinjeno i jednostranom izjavom volje ili u okviru ugovornog aranžmana (*contract out*), a nosioci prava trebali bi biti u mogućnosti da primijene mjere kako bi osigurali da se njihova rezerva poštuje.²⁵ Ukoliko je rezerva izražena, potrebna je dozvola nosilaca prava za upotrebu autorskopravno zaštićenog sadržaja za obuku. Upravo ovo rješenje je jedno od onih koje je najviše bilo na meti kritičara.²⁶

Konsekventno, iako EU-zakonodavac nije uveo navedeno ograničenje izričito u svrhu izuzeća obuke GenAI-modela na autorskopravno zaštićenim materijalima od povrede prava, smatra se da ono principijelno predstavlja “sigurnu luku” za takve aktivnosti AI-kompanija. To potvrđuju i jasne reference sadržane u AI-uredbi na relevantno ograničenje iz čl. 4 DSM direktive. S druge strane, predmetni akt nalaže njegovim beneficijarima dodatne obaveze u kontekstu njegove primjene.

3.2.2. AI-uredba

Također, i dugo očekivana AI-uredba EU pozabavila se pitanjem efikasne realizacije navedenih rezervi u korist nosilaca prava.

Pa tako, prije svega preambule 104–107 tematizuju obavezu transparentnosti za provajdere AI-modela opšte namjene (posebno velikih GenAI-modela) u pogledu obaveze izrade i stavljanja dostupnim javnosti dovoljno detaljnog sažetka o sadržaju koji se koristi za obuku tog modela, te obaveze uspostavljanja politike usklađenosti s autorskim pravom EU, a posebno u pogledu identifikovanja i pridržavanja iskazanih rezervi nosilaca prava u skladu sa čl. 4, st. 3 DSM-direktive. Pri tome, AI-uredba (preambula 106) ne pravi

23 DSM-direktiva, preambula br. 18.

24 *Ibid.*

25 *Ibid.*

26 A. Dermawan, 53.

razliku između domaćih i stranih provajdera i navodi da svaki pružalac usluga koji stavlja AI-model opšte namjene na tržište EU treba da se pridržava ove obaveze, i to bez obzira na jurisdikciju u kojoj se odvijaju akti relevantni za autorska prava koji podupiru obuku tih modela.²⁷

Predmetni sažetak ne bi trebao biti tehnički detaljan, već generalno sveobuhvatan, npr. na način da navodi glavne zbirke ili skupove podataka koji su ušli u obuku modela, kao što su velike privatne ili javne baze podataka ili arhive podataka, te da pruži narativna objašnjenja o drugim korištenim izvorima podataka. U ovom procesu provajderi bi trebali dobiti podršku novooosnovanog AI-ureda EU, tako što će ovo tijelo obezbijediti obrazac za sažetak, koji bi trebao biti jednostavan i efikasan, ali i pratiti da li je provajder ispunio obavezu izvještavanja.

3.3. *Fair use*

Dok su u državama koje pripadaju kontinentalno-evropskom konceptu autorskopravne zaštite okolnosti zbog kojih i svrhe u koje može doći do ograničenja i izuzetaka od primjene autorskog prava izričito zakonski propisane i sadržane u iscrpnim listama (npr. citiranje, informisanje javnosti, nastavne potrebe, umnožavanje djela u privatne i druge vlastite svrhe, upotreba djela trajno izloženih na javnim mjestima i sl.), u autorskom pravu SAD-a²⁸ se pored takvih definisanih ograničenja (npr. reprodukcija od strane biblioteka i arhiva ili upotreba od strane slijepih lica ili drugih lica s invaliditetom) primjenjuje i doktrina *fair use* ili “pravične upotrebe” (§ 107 ZAP SAD-a). Za razliku od prava EU, autorsko pravo SAD-a ne sadrži izričito propisano ograničenje za rudarenje teksta i podataka. Shodno tome, na ovu aktivnost, a konkretno, na korištenje autorskopravno zaštićenog materijala u svrhu obuke GenAI modela i kreiranje sintetičkih autputa na upite korisnika, mogao bi se potencijalno primijeniti pravni standard “pravične upotrebe”, ukoliko su prema ocjeni suda ispunjeni odgovarajući kriterijumi.²⁹

Pri utvrđivanju da li je korištenje djela u bilo kojem konkretnom slučaju pravično, u obzir se uzimaju sljedeći faktori (§ 107): (1) svrha i prirode upotrebe, uključujući i to da li je takva upotreba komercijalne prirode ili je u neprofitne obrazovne svrhe; (2) priroda djela zaštićenog autorskim pravima;

27 Domet navedene obaveze je u skladu sa stipuliranim obimom primjene AI-uredbe (čl. 2). Naime, ona se primjenjuje na pružatelje usluga koji stavljaju na tržište ili koriste AI-sisteme ili pak stavljaju na tržište AI-modele opšte namjene u EU i to bez obzira na to jesu li ti pružatelji osnovani ili locirani unutar EU ili pak u trećoj državi.

28 Zakon o autorskom pravu Sjedinjenih Američkih Država i srodni zakoni sadržani u naslovu 17 Zakonika Sjedinjenih Država, decembar 2022., <https://copyright.gov/title17/title17.pdf>, 28. april 2024. (ZAP SAD).

29 Za detaljan pregled principa koje sudovi primjenjuju prilikom primjene doktrine *fair use*, te za bazu sudskih mišljenja u ovom kontekstu po kategoriji i vrsti upotrebe (npr. muzika, internet/digitalizacija, parodija) vidjeti: U.S. Copyright Office, “U.S. Copyright Office Fair Use Index”, U.S. Copyright Office, novembar 2023., <https://www.copyright.gov/fair-use/index.html>, 6. maj 2024. (U.S. Copyright Office Fair Use Index).

(3) obim i supstancijalnost dijela upotrijebljenog djela³⁰ u odnosu na djelo kao cjelinu; i (4) efekat korištenja na potencijalno tržište ili vrijednost zaštićenog djela. U kontekstu svrhe i prirode upotrebe, vjerovatnije je da će se pravičnom smatrati tzv. "transformativna" upotreba, u smislu one upotrebe koja upravo dodaje nešto novo (drugu svrhu i prirodu), a ne zamjenjuje izvorni način upotrebe djela.³¹

Da li se ova aktivnost tuženih u predmetu *NYTC v. OpenAI i Microsoft* zaista može podvesti pod *fair use*, nadležni sud još treba da odluči. AI-kompanije, naravno, zastupaju mišljenje da može i pri tome se pozivaju na istovrsne stavove bibliotekskih udruženja, predstavnika akademske zajednice, grupa civilnog društva, ali i autora.³² Međutim, drugi autori argumentuju da kod GenAI modela tranformativnost kod autputa nije neminovno zagaranтована.³³

4. NEW YORK TIMES COMPANY PROTIV KOMPANIJA *OpenAI* I *MICROSOFT*

4.1. Okolnosti spora

Kao što je iz prethodnog izlaganja evidentno, tuženi OpenAI veoma proaktivno i uz to i uspješno nastoji da uredi svoje odnose s medijskim organizacijama, a što je navodno bio slučaj i s NYTC-om. Naime, u tekstu objavljenom na svojoj web-stranici, OpenAI navodi³⁴ da je kompanija s NYTC-om bila u pregovorima koji su ostavljali dojam da konstruktivno napreduju kroz posljednju komunikaciju 19. decembra 2023, te da su za samu tužbu saznali iz članka u NYT-u. U fokusu tih pregovora bilo je kreiranje partnerstva u pogledu prikazivanja sadržaja NYTC u realnom vremenu uz navođenje izvora, što je, prema navodima OpenAI-a, trebalo NYTC-u dati novi način da se poveže sa svojim postojećim i novim čitaocima, a korisnicima GPT-modela dati pristup izvještavanju NYTC-a.

Pored toga, Open AI tvrdi da je pojasnio NYTC-u da, kao i svaki pojedinačni izvor, njegov sadržaj nije značajno doprinio obuci postojećih GenAI-modela, a da također neće imati dovoljno utjecaja i na buduću obuku. Konačno, OpenAI ističe da je NYTC u razgovorima spomenuo tzv. „regurgitaciju“ svog sadržaja u autputima predmetnog AI-sistema, ali bez navođenja konkretnih primjera, što je ipak shvaćeno ozbiljno i rezultiralo u privremenom uklanjanju ChatGPT funkcije odmah po saznanju da može reprodukovati sadržaj u stvarnom vremenu na nenamjeran način.

30 Vid. detaljnije u predmetu Paul Tremblay, et. al. *protiv* OpenAI, INC., et al. (23-cv-03416-AMO), 3.1. Okolnosti spora.

31 U.S. Copyright Office Fair Use Index.

32 OpenAI, "OpenAI and journalism", *OpenAI*, 8. januar 2024a, <https://openai.com/index/openai-and-journalism/>, 30. april 2024., (OpenAI, 2024a).

33 P. Henderson *et al.*, 6 *et seq.*

34 OpenAI (2024a).

Međutim, ova nagla promjena stava NYTC-a u odnosu prema jednom od tuženih ne iznenađuje, uzimajući u obzir okolnost da je tužilac u ovom sporu u cilju zaštite svog poslovanja očigledno spreman poduzeti sve neophodne korake, pa čak doživjeti i “promjenu svijesti”.

Naime, pojedini autori³⁵ su po obznani tužbe i njenog sadržaja optužili NYTC za hipokriziju. Pri tome podsjećaju na slučaj *New York Times Co. v. Tasini*,³⁶ kad se ova medijska organizacija našla u ulozi tužene i pri tome nedvosmisleno dala prioritet svojim finansijskim interesima i napretku tehnologije u odnosu na interese svojih *freelance* autora,³⁷ dok se u tužbi protiv kompanija OpenAI i Microsoft oslanja na koncept tzv. “romantičnog autorstva”. Naime, NYTC već u prvim paragrafima tužbe donekle dramatično ilustruje posljedice navodnog uključenja velikog broja autorskih djela koja pripadaju NYTC-u u skupove podataka za obuku GenAI modela tuženih bez traženja dozvole i plaćanja naknade, apelujući na okolnost da je nezavisno novinarstvo od vitalnog značaja za demokratiju u SAD-u. Također potcrtava da NYTC upravo nudi uslugu pružanja pouzdanih informacija, analiza vijesti i komentara i poziva na to da ZAP SAD-a štiti originalno novinarstvo NYTC-a.³⁸ Jedan od tužbenih zahtjeva koji postavlja je čak uništavanje svih GPT ili drugih LLM modela i skupova podataka za obuku koji uključuju djela koja pripadaju NYTC-u.³⁹

Međutim, u predmetu *New York Times Co. v. Tasini*, NYTC je nastupao potpuno drugačije i odbijao mogućnost da bi presuda mogla zahtijevati brisanje članaka iz baza podataka u velikom obimu, jer bi uticaj tog postupka na elektronske arhive kao korisno sredstvo za istraživanje bio neprocjenjiv, a nije htio ni pristati na predložene šeme kompenzacije za autore.⁴⁰ Sve ovo,

35 A. Pope, “Blog Essay: NYT v. OpenAI: The Times’s About-Face”, *Harvard Law Review*, 10. april 2024., <https://harvardlawreview.org/blog/2024/04/nyt-v-openai-the-times-about-face/>, 11. april 2024.

36 The New York Times Company, Inc., et al. protiv Jonathan Tasini et al., Vrhovni sud Sjedinjenih Američkih Država (533 U.S. 483), 25. 6. 2001., presuda, <https://www.law.cornell.edu/supct/pdf/00-201PZO>, 26. maj 2024.

37 Postupak je pokrenulo šest *freelance* autora, a odnosi se na njihove članke koji su uključeni u tri printane publikacije. U okviru ugovora s izdavačima, ali bez saglasnosti *freelance* autora, dvije kompanije za kompjuterske baze podataka (i.a. LexisNexis) su kopije članaka tih autora unijele u tri baze podataka. U tužbi autora navodi se da su njihova autorska prava povrijeđena uključivanjem njihovih članaka u baze podataka. Izdavači su se u odgovoru pozvali na privilegiju reprodukcije i distribucije koju im daje §201(c) ZAP SAD-a. Konkretno, NYTC je tvrdio da je kao nosilac prava na kolektivnom djelu, tj. originalnoj štampanoj publikaciji (NYT), samo iskoristio “privilegiju” koju mu daje §201(c) da “reprodukuje i distribuiraju” autorski doprinos pojedinih autora. Međutim, Vrhovni sud je zauzeo stav da se ta odredba ZAP-a ne može primijeniti, budući da baze podataka reprodukuju i distribuiraju članke kao samostalna djela, a ne u širem kontekstu (kao dio tog određenog kolektivnog djela), te da su stoga kako štampani tako i elektronski izdavači prekršili autorska prava *freelance* autora.

38 NYTC v. Microsoft i OpenAI, Nature Of The Action (1–3).

39 NYTC v. Microsoft i OpenAI, Prayer For Relief.

40 A. Pope (2024).

dok u postupku protiv OpenAI-a i Microsofta višestruko ukazuje na njihovo besplatno korištenje (*free-riding*) značajnih napora NYTC-a i ulaganja ljudskog kapitala u prikupljanje informacija na kojima su bazirana djela koja želi zaštititi.

Nadalje je u kontekstu spora između OpenAI-a i NYTC-a od značaja istaći i djelomični uspjeh ovog tuženog u drugim, uporedivim postupcima koje su pokrenuli autori (i.a. Sarah Silvermann, Paul Tremblay i Michael Chabon) čije knjige su prema njihovim navodima korištene za treniranje GenAI modela tuženog. Naime, autputi tih modela (konkretno, *chatbota* ChatGPT) bi na temelju uputa korisnika producirali sažetke tih knjiga, što su tužioci smatrali djelima prerade i povredom svojih autorskih prava. No, u februaru 2024. godine, kalifornijski sud⁴¹ je u velikoj mjeri usvojio zahtjeve tužilaca za odbacivanje tužbe. Između ostalog je istaknuto da, s obzirom na to da tužioci nisu naveli da autputi ChatGPT-a sadrže izravne kopije knjiga zaštićenih autorskim pravima, moraju pokazati "značajnu sličnost" između tih autputa i samih knjiga. Također, sud je zauzeo stav da nije dovoljno navesti da je "svaki autput OpenAI jezičkih modela djelo prerade koje krši autorska prava", te da se taj navod odbacuje budući da tuženi nisu uspjeli objasniti šta autputi sadrže ili navode, da bi bilo koji specifični autput bio bitno sličan njihovim knjigama.⁴²

Shodno tome, u kontekstu spora između NYTC-a i OpenAI/Microsofta, ova naredba potencijalno postavlja presedan prema kojem, čak i ako su AI-modeli obučeni na zaštićenim autorskim djelima, njihovi autputi moraju dokazano uključivati aspekte tih djela da bi se mogla ustanoviti povreda, te ukazuje na to da su u postupcima potrebni veoma konkretni dokazi o tome kako autputi neposredno posuđuju iz zaštićenih djela.⁴³ NYTC je dokaze u svojoj tužbi formulisao upravo na takav način.

4.2. Sadržaj spora

4.2.1 Stranke i njihove uloge u sporu

Stranke u predmetnom postupku su NYTC, kompanija sa sjedištem u New Yorku, a koja izdaje digitalne i printane proizvode, od kojih je ključni publikacija *The New York Times* (NYT). Pored toga što je NYT dostupan kao printana novina, korisnicima je na raspolaganju i na mobilnoj aplikaciji, web-

41 Paul Tremblay, et. al. *protiv* OpenAI, INC., et al. (23-cv-03416-AMO), Okružni sud Sjedinjenih Američkih Država za sjeverni distrikt Kalifornije, Naredba kojom se djelimično prihvata i djelimično odbija zahtjev za odbacivanje, 12. februar 2024. (Naredba Tremblay v. OpenAI), <https://fingfx.thomsonreuters.com/gfx/legaldocs/gdvzdmxayvw/OPENAI%20COPYRIGHT%20LAWSUIT%20mtdruling.pdf>, 13. maj 2024.

42 Naredba Tremblay v. OpenAI, III. Analysis, A. Vicarious Copyright Infringement (Count II).

43 R. Yerramseti, "Case study of AI in publishing industry – New York Times vs. OpenAI and Microsoft case", *Medium*, 12. januar 2024., <https://medium.com/@ryerrams2003/case-study-of-ai-in-publishing-industry-new-york-times-vs-openai-and-microsoft-case-b14bfc85fbd3>, 13. maj 2024.

stranici (NYTimes.com), te putem povezanog sadržaja, kao što su podcasti. Prema navodima iz tužbe,⁴⁴ NYTC je nosilac prava na preko 3 miliona registrovanih autorskih djela.

Na strani tuženih, Microsoft Corporation (dalje u tekstu: Microsoft) je kompanija sa sjedištem u Washingtonu (Redmond), a koja je uložila minimalno 13 milijardi dolara u OpenAI Global LLC. Na osnovu navedenog ulaganja, do povrata te investicije, Microsoftu pripada 75% profita te kompanije, a nakon toga, 49% udjela u njoj. Kako je u tužbi istaknuto,⁴⁵ partnerstvo između Microsofta i ostalih tuženih podrazumijeva i značajnu tehničku saradnju na nastanku GenAI modela i omogućavanje preferencijalnog pristupa Microsoftu takvim najnovijim modelima, a koji su trenirani neovlaštenim primjercima djela na kojima prava ima NYTC i koji služe kreiranju sadržaja koji krši ta prava.

Ostali tuženi (dalje u tekstu: "OpenAI") predstavljaju skupinu povezanih lica sa sjedištem u Delawareu, pri čemu OpenAI Inc. neprofitno pravno lice osnovano 2015. godine, a koje posredno ima u vlasništvu ili na neki način kontroliše sva druga OpenAI pravna lica.⁴⁶ Sam OpenAI⁴⁷ definiše se kao kompanija za istraživanje i implementaciju čija je misija da osigura da umjetna generalna inteligencija koristi cijelom čovječanstvu.

U činjeničnim navodima tužbe,⁴⁸ NYTC, između ostalog, ističe tradiciju NYT-a i kroz skoro dva vijeka iskazanu novinarsku originalnost, ekspertizu, nepristrasnost, kvalitet i integritet od svog osnivanja 1851. godine. Temeljne oblasti djelovanja NYT-a obuhvataju: istraživačko novinarstvo, izvještavanje o udarnim vijestima, specijalizovano izvještavanje, preglede i analize, te komentare i mišljenja. Potcrtava se broj od preko 10 miliona digitalnih pretplatnika diljem svijeta od 2011. godine (tzv. *paywall* na webstranici), ca. 5800 zaposlenika, djelatnost novinara na terenu u preko 160 država, kreiranje u prosjeku 250 originalnih članaka dnevno, te nagrađivost (npr. 135 Pulitzerovih nagrada). Pored toga, skreće se pažnja na djelatnost stotina urednika koji provjeravaju tačnost, nezavisnost i pravičnost novinarskog rada, te na velika finansijska ulaganja u redakciju, samu publikaciju i tehnologiju.

U pogledu plaćenih sadržaja po osnovu digitalne pretplate,⁴⁹ ističe se da NYT daje registrovanim korisnicima slobodan pristup ograničenom broju članaka/sadržaja, tek nakon čega traži pretplatu za pristup punom/dodatnom sadržaju. Pored toga, izvršena je i sveobuhvatna digitalizacija članaka od osnivanja publikacije, te kreirane i digitalne replike svih dosadašnjih izdanja od 1851. do 2002. godine, što je također zahtijevalo značajno ulaganje.

44 NYTC v. Microsoft i OpenAI, III. Stranke, 14.

45 *Ibid.*, 15 *et seq.*

46 *Ibid.*, 18–25.

47 OpenAI, <https://openai.com/about/>, 19. april 2024.

48 NYTC v. Microsoft i OpenAI, III. Činjenični navodi tužbe, 26 *et seq.*

49 *Ibid.*, 45 *et seq.*

U pogledu GenAI proizvoda u ovom kontekstu,⁵⁰ u tužbi se navodi da je zaštita intelektualnog vlasništva NYTC-a ključna za njegovu kontinuiranu sposobnost da finansira novinarstvo svjetske klase u javnom interesu i da će u suprotnosti biti ugrožena sposobnost NYT-a da unovči novinarski sadržaj. Pri tome se argumentira da NYTC zavisi od svojih isključivih autorskih prava na, između ostalog, reprodukciju i preradu kako bi se suprotstavio ovim tendencijama, te da je registrovao autorska prava na svom štampanom izdanju svakog dana više od 100 godina, da održava *paywall* i da je implementirao opšte uslove poslovanja koji sadrže ograničenja u kopiranju i korištenju sadržaja.

Pa tako, korištenje autorskog sadržaja i žigova NYTC-a u komercijalne svrhe podrazumijeva prethodnu komunikaciju zainteresovane strane u pogledu uslova za davanje dozvole za takvo korištenje, što NYTC-u daje kontrolu nad samim korištenjem, ali i njegovim okolnostima i rokovima, te obezbjeđuje pravičnu naknadu. Skreće se pažnja, da NYTC ima zaključene ugovore o korištenju sadržaja s velikim digitalnim platformama, te da se dozvola za nekomercijalno i pojedine vidove komercijalnog korištenja može steći i putem "Copyright Clearance Center (CCC)". U tom kontekstu tužilac potcrtava da je mjesecima pokušavao da pregovorima⁵¹ postigne dogovor i da je tokom njih nastojao da osigura pravičnu kompenzaciju za korištenje svog sadržaja, ali i facilitira zdrav ekosistem djelatnosti informisanja i pomogne u razvoju GenAI na odgovoran način koji koristi društvu i podržava dobro informisanu javnost. Ipak, pregovori su bili neuspješni i tuženi su u javnosti insistirali na tome da je njihovo ponašanje zakonito zbog primjene doktrine *fair use*, jer njihovo korištenje sadržaja zaštićenog autorskim pravima, bez prethodno stečene dozvole, a za obuku modela umjetne inteligencije, služi novoj transformativnoj svrsi što NYTC kategorički odbija.

Tužilac ne troši vrijeme na to da u tužbi iznosi kontraargumente u pogledu ovog stava tuženih i osvrće se na kategorije "transformativna upotreba" i *fair use* svega nekoliko puta i to na vrlo površan i generalan način, dok fokus stavlja na argumente kako i zašto je došlo do povrede njegovih prava.

U cilju održavanja i širenja kruga digitalnih pretplatnika, te digitalnog marketinga, NYT dozvoljava internetskim pretraživačima pristup i indeksiranje svog sadržaja, što je neophodno kako bi se omogućilo korisnicima da pronađu publikaciju koristeći ove tražilice.⁵² Međutim, izričito se naglašava da NYTC nikada nije dao dozvolu bilo kojem subjektu, uključujući tužene, da koristi sadržaj publikacije u svrhe kreiranja proizvoda GenAI. S druge strane, NYTC se, prema navodima iz tužbe, obratio Microsoftu i OpenAI-u u aprilu 2023. godine kako bi otvorio mogućnost kreiranja sporazumnog rješenja za korištenje sadržaja od strane tuženih, uz određene komercijalne uslove i tehnološka ograničenja, no do tog rješenja nije došlo.

Tužilac ističe, govoreći o modelima umjetne inteligencije koje tuženi razvija,⁵³ da je inicijalno OpenAI djelovao na neprofitnoj osnovi i da su

50 *Ibid.*, 48 *et seq.*

51 NYTC v. Microsoft i OpenAI, Nature of the Action (7. i 8).

52 NYTC v. Microsoft i OpenAI, III. Činjenični navodi tužbe, 52 *et seq.*

53 NYTC v. Microsoft i OpenAI, B. Defendants' GenAI Products, 55 *et seq.*

LLM-ovi „GPT-1“ i „GPT-2“, a koji su proizašli iz OpenAI-a, bili publicirani na principu „otvorenog izvora“ (*open source*), te da ih je pratio visok stepen transparentnosti u pogledu stavljanja na raspolaganje javnosti kako seta podataka kojim su trenirani tako i dizajna te hardvera. Međutim, nakon što je OpenAI krenuo profitnim putem, za nove verzije spomenutih modela i to posebno „GPT-3.5“ i „GPT-4“, te informacije nisu bile učinjene dostupnim trećim licima iz razloga zaštite od konkurenata. Tuženi se, kako se u tužbi izlaže,⁵⁴ pozicionirao na tržištu posebno plasiranjem *chatbota* „Chat GPT“ u novembru 2022. godine, a koji je već u prvom mjesecu od objave dosegao broj od milion korisnika, a preko 100 miliona u prva tri mjeseca. Pored toga, tuženi s velikim uspjehom nudi pakete usluga zasnovanih na svojim LLM-ovima (besplatno i naplatno), a koje su usmjerene kako ka potrošačima tako i ka poslovnim subjektima. Pa tako, npr. preko 80% kompanija s liste „Fortune 500“ koristi ChatGPT, a OpenAI ostvaruje prihode od 80 miliona dolara mjesečno i na putu je da premaši preko milijardu dolara u narednih 12 mjeseci.

Tužilac argumentuje da je ovaj komercijalni uspjeh OpenAI-a velikim dijelom izgrađen na kršenju autorskih prava NYTC-a. Navedeno na način da je OpenAI napravio brojne reprodukcije djela zaštićenih autorskim pravima NYTC-a u toku „obuke“ predmetnih LLM-ova, pri čemu su sva povezana lica pod krovom OpenAI-a bila ili direktno uključena u tu povredu ili su njom upravljala, imala je pod svojom kontrolom ili profitirala od nje.⁵⁵

Što se tiče Microsofta,⁵⁶ prema tvrdnjama iz tužbe, ova kompanija je od 2019. bila blisko uključena u obuku, razvoj i komercijalizaciju GPT-proizvoda OpenAI-a i to barem na dva načina. Naime, tako što su Microsoft i OpenAI sarađivali na dizajniranju superkompjuterskih sistema koje pokreće Microsoftova računarska platforma u „oblaku“ pod nazivom „Azure“, a koji su korišteni za obuku svih GenAI modela OpenAI-a nakon „GPT-1“. Pored toga, Microsoft i OpenAI sarađuju i na komercijalizaciji tehnologije bazirane na GPT-modelu, npr. kroz kreiranje i plasiranje tzv. „Bing Chata“ u februaru 2023. godine, a koji predstavlja generativni AI *chatbot* na tražilici „Microsoft Bing“ koji pokreće GPT-4. Nadalje, u maju iste godine, Microsoft i OpenAI predstavili su „Browse with Bing“ dodatak (*plug-in*) za ChatGPT koji mu je omogućio pristup najnovijem sadržaju na internetu putem tražilice „Microsoft Bing“. Zajednički, „Bing Chat“ i „Browse with Bing“ dodatak imaju sposobnost generisanja sintetičkih rezultata pretraživanja, a što se u kontekstu korištenih sadržaja NYT-a manifestuje na način da oni predstavljaju opsežne parafraze i izravne citate tekstova NYT-a i iz tih tekstova izvedene nove tekstove i to često bez hiperlinka koji upućuje na izvor, te da se time gubi potreba korisnika da posjećuju web-stranicu NYT-a.⁵⁷

54 *Ibid.*, 61 *et seq.*

55 *Ibid.*, 64 *et seq.*

56 *Ibid.*, 66 *et seq.*

57 NYTC v. Microsoft i OpenAI, C. Defendants' Unauthorized Use and Copying of Times Content, 108. *et seq.*

4.2.2. Upotreba sadržaja NYTC-a

U pogledu navodne neovlaštene upotrebe sadržaja koji pripadaju NYTC-u, a za koju se tvrdi da je bila svjesna i namjerna i da je NYTC na nju višestruko skretao pažnju,⁵⁸ u tužbi se ističe⁵⁹ da su Microsoft i OpenAI izvršili njihovu reprodukciju, te distribuciju takvih reprodukcija na nekoliko neovisnih načina u procesu obuke svojih LLM-ova i rada s proizvodima koji ih uključuju. Ti načini obuhvataju sljedeće radnje i to: a) neovlaštenu reprodukciju autorskih djela NYTC-a tokom obuke GPT-modela; b) inkorporisanje neovlaštenih reprodukcija i djela prerade autorskih djela NYTC-a u GPT modele c) neovlašteno javno prikazivanje autorskih djela NYTC-a u GPT-autputima proizvoda i d) neovlašteno preuzimanje i širenje aktuelnih vijesti. U tekstu rada ćemo se osvrnuti posebno samo na neovlaštenu reprodukciju i donekle na inkorporisanje takvih reprodukcija u GPT-modele tuženih.

Pa tako, naprimjer, u kontekstu stavljanja na teret neovlaštene reprodukcije autorskih djela NYTC-a tokom obuke GPT-modela, u tužbi se obrazlaže, da spomenuti modeli tuženih predstavljaju “porodicu LLM-ova” (od GPT-1 do GPT-4). LLM-ovi u *chat* stilu (GPT-3.5 i GPT-4) razvijani su u dvije faze, i to kroz prethodni trening na veoma velikoj količini podataka, zatim kroz “fino podešavanje” modela s manjim i nadziranim skupom podataka koji bi trebao pomoći modelu u rješavanju specifičnih zadataka. Prethodna obuka uključuje i prikupljanje i pohranjivanje tekstualnog sadržaja za stvaranje skupova podataka (*datasets*) za obuku i obradu tog sadržaja putem GPT-modela. OpenAI je objavio opšte informacije o svom procesu prethodne obuke za GPT-modele, pa tako iz podataka o obučenoj verziji GPT-2 proizlazi da u korištenom tzv. “WebText” skupu podataka domena NYTimes.com čini petu po redu top domenu u tom skupu s preko 300.000 upisa.

Za razliku od GPT-2 koji je uključivao 1.5 milijardi parametara, GPT-3 je već uključivao 175 milijardi parametara,⁶⁰ a pretpostavlja se da GPT-4 obuhvata čak 1.8 biliona parametara.⁶¹

58 *Ibid.*, 124 *et seq.*

59 *Ibid.*, 82 *et seq.*

60 Parametar je vrijednost koja određuje postupanje modela mašinskog učenja, a koji nije ništa drugo do matematički algoritam koji je dizajniran da prepoznaje uzorke u skupu podataka kojim je obučavan i da daje predviđanja na temelju tih podataka. Svaki parametar djeluje kao varijabla koja određuje kako će model obraditi te podatke i generisati output (predviđanje), kao npr. u formi kriterijuma za stil i/ili sadržaj generisanog teksta na upit (*prompt*) korisnika. Međutim, parametri su varijable koje uče vlastite vrijednosti iz skupa podataka kojim se obučava model, a u kojem procesu se ti parametri ažuriraju od strane algoritma sve dok ne dobiju svoje optimalne vrijednosti. Principijelno, što više parametara jedan model ima, to je bolja izvedba modela u pravcu kvaliteta i/ili tačnosti outputa, iako previše parametara može imati i obratan efekat. Keary, A. “What is the Role of Parameters in AI”, *Techopedia*, 28. novembar 2023, <https://www.techopedia.com/experts/what-is-the-role-of-parameters-in-ai>, 16. april 2024.

61 NYTC v. Microsoft i OpenAI, C. Defendants’ Unauthorized Use and Copying of Times Content, 85, 86. i 91.

Pa tako, jedan od skupova podataka koji su korišteni za trening GPT-3,⁶² tzv. "WebText2", jest skup podataka koji obuhvata sadržaj visokog kvaliteta i koji nosi 22% udjela u procesu treninga, te uključuje značajan broj izvora NYTC-a. Također i "Common Crawl" skup podataka koji nosi čak 60% udjela u procesu treninga i titulira se kao "kopija interneta", uključuje, prema navodima iz tužbe i otvorenim podacima iz 2019. godine, najmanje 16 miliona jedinstvenih zapisa sadržaja od NYTC-a, obuhvatajući izdanja kao *News*, *Cooking*, *Wirecutter* i *The Athletic*, te više od 66 miliona ukupnih zapisa sadržaja iz NYT-a.

Tužilac zaključuje⁶³ da su tuženi više puta kopirali i opsežno komercijalno koristili ovu masu sadržaja na kojima autorska prava ima NYTC, a sve u cilju razvoja svojih GenAI- modela i to bez ikakve dozvole ili plaćanja naknade tužiocu. Naime, kao dio obuke GPT-modela, Microsoft i OpenAI sarađivali su na razvoju složenog, prilagođenog superračunarskog sistema za smještaj i reprodukciju kopija skupova podataka za obuku, uključujući kopije sadržaja koji pripada NYTC-u, a kako bi ti modeli oponašali sadržaj i autore NYT-a. Pored toga, Microsoft i OpenAI su, navodno, sarađivali i u dizajniranju GPT-modela, selekciji skupova podataka za obuku i nadgledanje procesa obuke, te, između ostalog, nastavljaju da kreiraju neovlaštene kopije u formi rezultata pretrage koji proizlaze iz upotrebe "Bing Chata" i "Browse with Bing" dodatka.

Konačno, na temelju informacija i uvjerenja, NYTC tvrdi da tuženi već sada kreiraju, ili će uskoro početi da kreiraju, dodatne kopije autorskih djela ove kompanije u cilju obuke ili "finog podešavanja" naredne generacije iz ove LLM-porodice, odnosno GPT-5.

U pogledu inkorporisanja neovlaštenih kreiranih reprodukcija autorskih djela⁶⁴ i djela prerade autorskih djela NYTC-a u GPT-modele, tuženi prilaže Dokaz J⁶⁵ u svrhu potvrđivanja argumenta da su i sami LLM-ovi tuženih "memorisali"⁶⁶ kopije mnogih autorskih djela koja pripadaju tužiocu, a koja

62 *Ibid.*, 86.

63 *Ibid.*, 92 *et seq.*

64 *Ibid.*, 98. *et seq.*

65 The New York Times Company *protiv* Microsoft Corporation, OpenAI, Inc., OpenAI LP, OpenAI GP, LLC, OpenAI, LLC, OpenAI OPCO, LLC, OpenAI Global, LLC, OAI Corporation, LLC i OpenAI Holdings, LLC (1:23-cv-11195), Okružni sud Sjedinjenih Američkih Država za južni distrikt New Yorka, Dokaz J, dokument 1-68, Stotinu primjera kako GPT-4 memoriše sadržaj iz New York Timesa, <https://nytco-assets.nytimes.com/2023/12/Lawsuit-Docket-dkt-1-68-Ex-J.pdf>, 20. mart 2024., (Dokaz J).

66 Iako se ne radi o temeljnom modelu generativne umjetne inteligencije, kao što je to GPT, već o difuzijskom modelu, u ovom kontekstu vrijedi obratiti pažnju na argumentaciju iz odgovora na tužbu u sporu (Visoki sud u Londonu) između tuženog, kompanije Stability AI koja stoji iza sistema "Stable Diffusion", a koji automatski generiše slike na temelju tekstualnih ili slikovnih upita korisnika i kompanije u domenu vizuelnih medija, i tužio-ca Getty Images. Prema navodima tužene, "Stable Diffusion" ne sadrži kopiju bilo kojeg sadržaja na kojem je model obučavan i kod kreiranja outputa pravi razliku da li je uputa tekstualna, kada nastaje output bez korištenja bilo kojeg autorskog djela, ili slikovna. Ukoliko nakon slikovne upute model generira output koji sadrži elemente koji odgovara-

su kroz proces obuke kodirana u njihovim parametrima. Pa tako, navedeni dokaz sadrži stotinu primjera koji pokazuju kako je GPT-4 memorisao sadržaj NYT-a i da može uz odgovarajuću uputu (*prompt*) pružiti autpute koji predstavljaju gotovo doslovne kopije značajnih dijelova autorskih djela NYTC-a.⁶⁷ Navedeno dovodi do zaključka da su GenAI modeli tuženih obučeni na tim konkretnim i bezbrojnim drugim autorskim djelima NYTC-a, jer u suprotnom ne bi mogli generisati tako doslovne tekstove i izvode iz originalnih tekstova NYT-a.

Tuženi na teret tužiocima također stavlja i “neovlašteno prisvajanje komercijalnih preporuka”⁶⁸ zbog kreiranja sintetičkih rezultata pretrage i prikazivanja isječaka i parafraza putem “Bing Chata” i dodatka “Browse with Bing” iz sadržaja *online* publikacije *Wirecutter*, čime se često potpuno reprodukuju preporuke proizvoda s te platforme i umanjuje *online* promet na izvornu stranicu.

Naime, stranica *Wirecutter*⁶⁹ nudi usluge preporuke proizvoda, na način da njeni novinari neovisno testiraju i pregledaju razne kategorije proizvoda. Finansira se putem pretplata i raznih partnerskih marketinških programa (*affiliate marketing*), što podrazumijeva da dobija proviziju za proizvode kupljene preko njenih linkova ka proizvodima na web-stranicama prodavača/proizvođača. Shodno tome, *Wirecutter* neće dobiti proviziju od preporuka ako korisnik kupi proizvod koji *Wirecutter* preporučuje putem linka na platformama tuženih. Prema tome, smanjeni *online* promet na *Wirecutterove* članke zbog sintetičkih rezultata “Bing Chata” i dodatka “Browse with Bing” podrazumijeva i smanjeni promet na pridružene linkove ka preporučenim proizvodima, a što dovodi do gubitka prihoda za *Wirecutter*. Na ovaj način se GenAI proizvodi tuženih direktno i nesavjesno takmiče sa sadržajem NYTC-a i uzurpiraju komercijalne prilike tužioca.⁷⁰

ju elementima fotografije korištene za obuku, taj rezultat samo reflektuje korisnički upit koji je formulisan tako da generiše fotografiju koja sadrži te elemente. Odnosno, takav autput ne proizlazi iz sadržaja na kojem je model treniran, već iz upita koji je dizajnirao korisnik. Također, Stability AI se poziva i na ograničenje upotrebe autorskih djela za stvaranje pastiša, a koje je u pravo Ujedinjenog Kraljevstva ušlo implementacijom čl. 17 DSM direktive. Naime, tuženi tvrde da autput modela predstavlja pastiš, u smislu da je to kreacija u stilu koji može oponašati stil drugog djela, umjetnika, perioda ili mnoštva elemenata iz velikog broja različitih izvora materijala za obuku. Vid. detaljnije kod: C. Wyn Davis, G. Dennis, Getty Images v Stability AI: the implications for UK copyright law and licensing, *Pinsent Masons*, 29. april 2024., <https://www.pinsentmasons.com/out-law/analysis/getty-images-v-stability-ai-implications-copyright-law-licensing>, 27. maj 2024.

67 Npr.: Autput iz GPT-4: “In January 2017, outgoing Obama administration officials ran an extensive exercise on responding to a pandemic for incoming senior officials of the Trump administration. The full story of *why the Trump administration failed to heed these warnings is still emerging, and will be the subject of a commission, if Congress gets its way.*” Stvarni tekst iz NYT: “In January 2017, outgoing Obama administration officials ran an extensive exercise on responding to a pandemic for incoming senior officials of the Trump administration. The full story of *the Trump administration’s response to the coronavirus.*”, Dokaz J, str. 60.

68 NYTC v. Microsoft i OpenAI, D. Misappropriation of Commercial Referrals, 127. *et seq.*

69 Wirecutter, <https://www.nytimes.com/wirecutter/about/>, 28. april 2024.

70 NYTC v. Microsoft i OpenAI, D. Misappropriation of Commercial Referrals, 129.

Pored toga, tužilac je utvrdio da “Bing Chat” i dodatak “Browse with Bing” ponekad lažno pripisuju *Wirecutteru* preporuke za proizvode koje ova stranica uopšte nije dala (“halucinacija”), čime se ugrožava ugled *Wirecuttera*.⁷¹

Konačno, ove tzv. “halucinacije”, kako se nazivaju u žargonu svijeta umjetne inteligencije, koje ne predstavljaju ništa drugo do dezinformaciju,⁷² nisu vezane samo za *Wirecutter*, već i za druge sadržaje koji se pripisuju tužiocu. Naime, nakon što dobiju uputu, bez obzira na to da li raspolažu s tačnim informacijama na samu temu upita, GPT-proizvodi tuženih (npr. *chatbot* “Chat GPT”) će pružiti autput s informacijama koje u najboljem slučaju nisu posve tačne, a u najgorem su potpuno lažne i izmišljene. Problem leži u okolnosti što su korisnici tih modela ljudi, kojima je teško razlikovati tačno od fabriciranog (tzv. “halucinacije”) i što se na takav način dovode u zabludu u pogledu izvora informacija koje dobijaju, pogrešno vjerujući da je predmetne informacije provjerio i objavio NYT.

4.2.3. Povrede prava

Tužilac u tužbenom zahtjevu ističe tri vida povrede autorskih prava i to: neposrednu povredu od strane svih tuženih,⁷³ posrednu povredu od strane određenog broja tuženih⁷⁴ i tzv. “doprinosnu” povredu⁷⁵ od strane Microsofta, ali i svih tuženih zajedno.

U pogledu neposredne povrede autorskih prava na autorskim djelima registrovanim u korist tužioca, u tužbi je istaknuto da je došlo do nje na način da su tuženi u skupove podataka koje su kreirali za obuku uključili djela NYTC-a tako što su ih direktno “povukli” sa stranica NYTC-a ili reprodukovali iz skupova podataka trećih lica. Pored toga, povreda je izvršena, prema navodima tužioca, i pohranjivanjem, obradom i reprodukcijom na Microsoftovoj superračunarskoj platformi skupova podataka za obuku GPT-modela koji sadrže milione primjeraka djela NYTC-a. Pored toga, tužilac konstatuje povredu i kroz diseminaciju generativnog autputa koji sadrži kopije djela NYTC-a, ali i djela prerade tih kopija putem ChatGPT-a i Bing Chata.

U kontekstu posredne povrede, tužilac naglašava da su jedni tuženi kontrolisali i usmjeravali povredu, te profitirali od te povrede, a da su tu povredu neposredno počinili drugi tuženi (npr. Microsoft kontroliše i usmjerava superračunarsku platformu koja se koristi za pohranu, obradu i reprodukciju skupova podataka za obuku).

U pogledu “doprinosne povrede”, koja predstavlja jedan vid sekundarne odgovornosti, tužilac Microsoftu stavlja na teret da je materijalno doprinio i

71 *Ibid.*, 134.

72 NYTC v. Microsoft i OpenAI, E. “Hallucinations” Falsely Attributed to The Times, 136 *et seq.*

73 NYTC v. Microsoft i OpenAI, COUNT I: Copyright Infringement (17 U.S.C. § 501), 158 *et seq.*

74 NYTC v. Microsoft i OpenAI, COUNT II: Vicarious Copyright Infringement, 169 *et seq.*

75 NYTC v. Microsoft i OpenAI, COUNT III: Contributory Copyright Infringement, 174 *et seq.* i COUNT IV: Contributory Copyright Infringement, 178 *et seq.*

direktno pomogao u neposrednom kršenju prava koje se može pripisati grupi OpenAI tuženih. Navedeno na način da je znao ili je imao razloga da zna za direktnu povredu koju je počinila grupa OpenAI tuženih, jer se njihovo partnerstvo proteže na razvoj, komercijalizaciju i monetizaciju GPT-proizvoda. Navod “doprinosne povrede” svih tuženih zajedno tužilac ograničava na mjeru u kojoj krajnji korisnik može biti odgovoran za neposrednu povredu na osnovu autputa GPT-proizvoda, budući da su tuženi u tom slučaju materijalno doprinijeli i direktno pomogli u neposrednom kršenju koje su počinili ti korisnici.

Na teret svih tuženih tužilac stavlja i tvrdnju povrede *Digital Millennium Copyright Acta* zbog uklanjanja informacija o upravljanju autorskim pravima.⁷⁶ Naime, NYT inkorporiše takve informacije u svako od djela NYTC-a, a koje obuhvataju, naprimjer, obavijest o autorskim pravima, naslov i druge identifikacijske informacije, odredbe i uslove korištenja, te identifikacijske brojeve ili simbole koji se odnose na informacije o upravljanju pravima. Prema navodima iz tužbe, tuženi su prilikom neovlaštenog kopiranja djela NYTC-a u svrhu obuke svojih modela generativne umjetne inteligencije uklonili informacije o upravljanju pravima NYTC-a, a one nisu prisutne ni u autputima baziranim na tim djelima. Shodno tome, tužilac potcrtava da je uklanjanje ili izmjena podataka o upravljanju pravima učinjena svjesno i s namjerom da se potakne, omogućiti, olakša ili prikrije kršenje autorskih prava, te da tužiocu pripadaju statutarne šteta, stvarna šteta, povrat dobiti te puni troškovi postupka i advokatske naknade.

Nadalje⁷⁷ tužilac stavlja na teret tuženim nelojalnu konkurenciju po pravilima precedentnog prava zbog neovlaštenog prisvajanja. Naime, prema navodima tužbe, tuženi se, na način da nude sadržaje koje je producirao GenAI, pri čemu je taj sadržaj isti ili sličan sadržaju koji objavljuje NYT, direktno takmiče sa sadržajem NYT-a i uzurpiraju i prisvajaju specifične komercijalne prilike NYTC-a, kao što je npr. prihod generisan preporukama *Wirecuttera*. Pored toga, upotreba sadržaja NYT-a od strane tuženih za obuku GenAI modela koji proizvode istovrsni informativni tekst podrazumijeva da se taj sadržaj takmiči sa sadržajem NYT-a za promet i ujedno predstavlja i besplatno korištenje NYTC-evih značajnih napora i ulaganja ljudskog kapitala u prikupljanje ovih informacija. Time NYTC trpi, prema navodima iz tužbe, stvarnu štetu zbog uskraćivanja koristi od svog rada, kao što je npr. gubitak prihoda od oglašavanja i partnerskih preporuka.

Konačno, tužitelj ističe i povredu svojih čuvenih i poznatih žigova (i. a. “The New York Times”, “nytimes” i “nytimes.com”) kroz njihovo razvodnjavanje.⁷⁸

76 NYTC v. Microsoft i OpenAI, COUNT V: Digital Millennium Copyright Act – Removal of Copyright Management Information (17 U.S.C. § 1202), 181 *et seq.*

77 NYTC v. Microsoft i OpenAI, COUNT VI: Common Law Unfair Competition By Misappropriation, 192 *et seq.*

78 NYTC v. Microsoft i OpenAI, COUNT VII: Trademark Dilution (15 U.S.C. § 1125(c)), 198 *et seq.*

Na temelju izloženog, tužilac je postavio niz tužbenih zahtjeva,⁷⁹ između ostalog, dodjelu statutarne i kompenzacijske štete, i restituciju, trajnu zabranu tuženim da vrše nezakonite, nepravedne i kršeće radnje navedene u tužbi i čak uništavanje svih GPT ili drugih LLM-modela i skupova podataka za obuku koji uključuju djela koja pripadaju NYTC-u. Interesantno je da u tužbi nije naveden tačan iznos potraživanja naknade materijalne štete.

4.3. Trenutni epilog spora

OpenAI u prilogu objavljenom na svojoj web-stranici⁸⁰ opetovano ističe kako smatra da je tužba NYTC-a neosnovana i da korištenje javno dostupnih materijala na internetu za obuku GenAI-modela, a što uključuje materijale koji su autorskopravno zaštićeni, predstavlja *fair use*. No, nezavisno od toga također skreće pažnju na to da je OpenAI vodeća kompanija u AI-industriji koja pruža jednostavan *opt-out* proces⁸¹ za izdavače kako bi oni uskratili korištenje svojih sadržaja,⁸² te da je i sam NYTC iskoristio tu opciju u avgustu 2023. godine. Pored toga, navodi da memorisanje sadržaja od strane modela predstavlja rijetku grešku (*a rare bug*) na čijem se eliminisanju radi.

U februaru i martu 2024. godine, obje tužene kompanije podnijele su zahtjeve za odbijanje tužbe.

U svom podnesku⁸³ OpenAI zahtijeva odbacivanje tužbe NYTC-a po četiri osnova i to: za neposrednu povredu autorskih prava (I) u onoj mjeri u kojoj se ona temelji na aktima reprodukcije koji su se dogodili prije više od tri godine od tužbe; za "doprinosnu povredu" autorskih prava (IV) u cijelosti, jer je u tužbi propušteno da se učini vjerovatnim da je kompanija imala stvarno znanje o konkretnim radnjama navodne neposredne povrede; za navodnu povredu podataka o upravljanju pravima (V) u cijelosti, jer je propušteno da se identifikuju sporni podaci, da se učini vjerovatnim da je OpenAI uklonio te podatke iz bilo kojeg skupa podataka ili autputa, da se učini vjerovatnom "distribucija", te da se navedu činjenice koje sugerišu da je OpenAI djelovao s namjerom; i za navodnu nelojalnu konkurenciju neovlaštenim prisvajanjem (VI) prema pravu države New York u cijelosti, jer se na okolnost spora prvenstveno primjenjuje ZAP SAD-a.

79 NYTC v. Microsoft i OpenAI, Prayer For Relief.

80 OpenAI, (2024a).

81 Na OpenAI platformi nosioci prava mogu odabrati opciju potpune ili djelimične zabrane pristupa svojim stranicama za oba "korisnička agenta" koji vrše indeksiranje weba i korisničku pretragu. OpenAI, "GPTbot", <https://platform.openai.com/docs/gptbot>, 14. maj 2024.

82 Vidjeti *opt-out* opciju (čl. 4, st. 3) prema DSM direktivi i obim primjene AI uredbe (čl. 2).

83 The New York Times Company *protiv* Microsoft Corporation, OpenAI, Inc., OpenAI LP, OpenAI GP, LLC, OpenAI, LLC, OpenAI OPCO, LLC, OpenAI Global, LLC, OAI Corporation, LLC i OpenAI Holdings, LLC (1:23-cv-11195), Okružni sud Sjedinjenih Američkih Država za južni distrikt New Yorka, Podnesak kojim se podržava zahtjev OpenAI-a za odbacivanje tužbe, 26. februar 2024., IV. Argument (OpenAI-ev zahtjev za odbacivanje), <https://static01.nyt.com/newsgraphics/documenttools/82a013b9ba852548/9d4b1790-full.pdf>, 19. maj 2024.

U zahtjevu za odbacivanje, zastupnici OpenAI-a optužili su NYTC da je platio nekome da “hakira” njihov *chatbot* i da im je trebalo na desetine hiljada pokušaja da bi generisali zavaravajuće i anomalne rezultate koje su prezentovali u Dokazu J, a što je jedino bilo izvodivo ciljanjem i iskorištavanjem buga [u „*chatbotu*“] korištenjem obmanjujućih upita koji krše OpenAI-ove uslove korištenja.⁸⁴ Tužilac, NYTC, negirao je te tvrdnje.⁸⁵

Podnesak također sadrži tvrdnje da je zakonito koristiti materijal zaštićen autorskim pravima u svrhu obuke GenAI modela jer to predstavlja tzv. “nekonzumentsku upotrebu” koja je vid upotrebe kroz tehnološki proces.⁸⁶ Pri tome se OpenAI pozvao na korpus precedentnog prava (npr. *Authors Guild v. HathiTrust*, *Authors Guild v. Google, Inc.* [GoogleBooks] i *Google LLC v. OracleAm., Inc.*) koji potvrđuje da je korištenje zaštićenog sadržaja zakonito kada je ono dio tehnološkog procesa koji rezultira stvaranjem novih, drugačijih i inovativnih proizvoda. OpenAI ističe da autorsko pravo ne predstavlja pravo veta na transformativne tehnologije koje interno maksimiziraju iskorištavanje postojećih djela, a čime se unapređuje osnovna svrha ovog prava, bez da se pri tome potkopavaju mogućnosti autora da utrži svoja djela.

U svom podnesku sudu⁸⁷ Microsoft zahtijeva djelimično odbacivanje tužbe NYTC-a u pogledu tri tačke i to, identično kao i OpenAI, za “doprinosnu odgovornost” za povredu autorskih prava od strane krajnjih korisnika alata zasnovanih na GPT-u, za povredu *Digital Millennium Copyright Acta* u kontekstu informacija o upravljanju pravima, te za neovlašteno prisvajanje po osnovu zakona države New York, pri čemu ZAP SAD-a ima prevagu u primjeni.

U kontekstu upotrebe autorskopravno zaštićenog materijala NYTC-a, Microsoft tvrdi da ta zaštita ne predstavlja ništa veću prepreku za obuku

84 OpenAI zahtjev za odbacivanje, I. Introduction. Vid. i: J. Henley, “OpenAI claims New York Times ‘hacked’ ChatGPT to build copyright lawsuit”, *The Guardian*, 27. februar 2024., <https://www.theguardian.com/technology/2024/feb/27/new-york-times-hacked-chatgpt-openai-lawsuit>, 28. april 2024. i B. Brittain, “OpenAI says New York Times ‘hacked’ ChatGPT to build copyright lawsuit”, *Reuters*, 27. februar 2024., <https://www.reuters.com/technology/cybersecurity/openai-says-new-york-times-hacked-chatgpt-build-copyright-lawsuit-2024-02-27/>, 20. april 2024.

85 B. Brittain, “New York Times denies OpenAI’s ‘hacking’ claim in copyright fight”, *Reuters*, 12. mart 2024., <https://www.reuters.com/legal/litigation/new-york-times-denies-openais-hacking-claim-copyright-fight-2024-03-12/>, 20. april 2024.

86 OpenAI zahtjev za odbacivanje, I. Introduction i C. Reliance on Longstanding Fair Use Principles. Vid. također: C. Metz, K. Roberts, “OpenAI Seeks to Dismiss Parts of The New York Times’s Lawsuit”, *The New York Times*, 27. februar 2024., <https://www.nytimes.com/2024/02/27/technology/openai-new-york-times-lawsuit.html>, 30. april 2024.

87 The New York Times Company protiv Microsoft Corporation, OpenAI, Inc., OpenAI LP, OpenAI GP, LLC, OpenAI, LLC, OpenAI OPCO, LLC, OpenAI Global, LLC, OAI Corporation, LLC i OpenAI Holdings, LLC (1:23-cv-11195), Okružni sud Sjedinjenih Američkih Država za južni distrikt New Yorka, Podnesak kojim se podržava zahtjev Microsoft Corporationa za djelimično odbacivanje tužbe, 4. mart 2024. (Microsoftov zahtjev za odbacivanje), <https://www.courtlistener.com/docket/68117049/65/the-new-york-times-company-v-microsoft-corporation/>, 20. maj 2024.

LLM-a nego što je to onomad bila za videorekordere (ili klavir, fotokopirni aparat, kompjuter, internet ili pretraživač). Na ovom mjestu Microsoft aludira na presudu *Sony Corp. of America protiv Universal City Studios*⁸⁸ koja je bila epilog tužbe ovog filmskog studija zbog Betamax videorekordera, a u kojoj se tvrdilo da će ova nova tehnologija ljudima omogućiti ilegalno kopiranje filmova i televizijskih emisija. Međutim, Vrhovni sud SAD-a je zauzeo stav da je izrada tih kopija za lično/privatno gledanje *fair use* prema zakonu.⁸⁹ Na tom tragu Microsoft tvrdi da zaštićeni sadržaj koji se koristi za obuku LLM-ova ne potiskuje tržište za autorska djela, već samo podučava GenAI o jeziku, a da je to bit onoga što ZAP SAD-a smatra transformativnom i pravičnom upotrebom (*fair use*).⁹⁰

5. ZAKLJUČNA RAZMATRANJA

Kad posmatramo NTYC-ovu tužbu u širem kontekstu trenutnog medijskog ekosistema i njegovog stava prema umjetnoj inteligenciji generalno i industriji koja stoji iza ove tehnologije posebno, teško je ne steći dojam da NYTC moguće vodi jednu donekle bespotrebnu, pomalo kapricioznu, no pravno krajnje zanimljivu i relevantnu bitku.

Naime, očigledno je kako iz same tužbe tako i iz navoda OpenAI-a, da su stranke u ovom postupku neposredno prije toga bile stranke u pregovorima, vrlo moguće uporedivim s onima koje je ovaj AI-gigant paralelno vodio i s drugim medijskim organizacijama (npr. Axel Springer, Financial Times ili News Corp). No, u tim drugim slučajevima, pregovori su polučili zaključivanjem ugovora kojim se naplatno dozvoljava korištenje sadržaja tih organizacija u procesu obuke GenAI modela, te njihov preferencijalni status u ChatGPT autputima, a uz reference i linkove na originalne izvore kako bi se povećao digitalni promet na njihove web- stranice. Uz to, OpenAI je kreirao i tehnološka partnerstva s nekim od tih medijskih kompanija, između ostalog, u cilju korištenja alata umjetne inteligencije u obavljanju njihove djelatnosti (npr. za nadogradnju proizvoda za usmjeravanje oglasa "D/Cipher" kompanije Dotdash Meredith). To apsolutno ne iznenađuje, jer medijska industrija ima, uz brojne zadržke i bojazni, ipak generalno pozitivan stav prema techno-

88 Sony Corporation of America *protiv* Universal City Studios, Inc., Vrhovni sud Sjedinjenih Američkih Država (464 U.S. 417), 17. januar 1984., presuda, <https://supreme.justia.com/cases/federal/us/464/417/>, 21. maj 2024.

89 C. Metz, K. Roberts, "Microsoft Seeks to Dismiss Parts of Suit Filed by The New York Times", *The New York Times*, 4. mart 2024., <https://www.nytimes.com/2024/03/04/technology/microsoft-ai-copyright-lawsuit.html>, 30. april 2024. Također vidjeti: E. David, "Microsoft invokes VCRs in motion to dismiss The New York Times' AI lawsuit", *The Verge*, 6. mart 2024., <https://www.theverge.com/2024/3/5/24091719/microsoft-new-york-times-openai-motion-dismiss-copyright-lawsuit>, 15. 5. 2024. i M. Sigalos, H. Field, "Microsoft compares The New York Times' claims against OpenAI to Hollywood's early fight against VCR", *CNBC*, 5. mart 2024., <https://www.cnn.com/2024/03/05/microsoft-seeks-dismissal-parts-of-new-york-times-suit-against-openai.html>, 29. april 2024.

90 Microsoftov zahtjev za odbacivanje, Preliminary Statement.

loškom fenomenu umjetne inteligencije i to ne odnedavno. Naime, da navedemo samo jedan primjer, najranije forme tzv. „robotizovanog novinarstva“ ili kompjuterski asistiranog novinarstva prisutne su u ovoj branši još od 50-ih godina prošlog vijeka, a u današnje vrijeme ovaj način izvještavanja koristi se u velikoj mjeri u domenu sporta, vremenske prognoze ili finansijskih vijesti.⁹¹ Prema tome, mediji žele biti dio ovog fenomena, a ne na njegovim marginama. Ta okolnost neminovno otvara pitanje, da li se NYTC s trenutnim stavom i postupkom koji je inicirao upravo dovodi u tu graničnu zonu?

Naime, sadržaji navedenih sporazuma predstavljaju odraz u ogledalu navoda koje NYTC stavlja na teret kompanijama OpenAI i Microsoft. Odista je teško povjerovati da iste, ako ne i preferencijalne uslove, OpenAI nije ponudio jednom takvom medijskom dinosaurusu kao što je NYTC.

S druge strane, sigurno je da je ova NYTC-ova tužba otvorila novo poglavlje u odnosu između tehnologije i medija na temu dozvola za korištenje zaštićenih sadržaja i distribucije prihoda od njihovog korištenja. Granice *fair use* doktrine u ovom predmetu će biti strogo testirane, do čega je u nekom trenutku u kontekstu obuke GenAI modela i sadržaja autputa GPT-proizvoda moralo doći. Vjerovatno će ovaj postupak polučiti jednim izuzetno značajnim precedentsom koji će moguće, u zavisnosti od njegovog sadržaja, uticati i na sam razvoj tehnoloških procesa. Trenutni postupci inicirani od autora pojedinaca upućuju na pozitivan stav sudova u pogledu korištenja zaštićenih materijala od strane AI-industrije, međutim, nedostaci na koje je ukazano u pogledu sadržaja dokaza (npr. supstancijalnost) nisu prisutni u temeljito potkrijepljenim navodima NYTC-a koji ukazuju na tzv. „regurgitaciju“ tekstova NYT-a u značajnom obimu.

Ipak, kako neki autori⁹² potpuno ispravno primjećuju, ključan element američkog prava, a koji posebno dolazi do izražaja u njegovoj primjeni, jest pragmatizam. To podrazumijeva da će sud uzeti u obzir ne samo prethodno uspostavljena načela, već i društveni, kulturni i ekonomski kontekst konkretnog slučaja. Shodno tome, moguće je da u ovom predmetu uvaži argumente promocije inovacija i razvoja tehnologije, te navodnog transformativnog dejstva umjetne inteligencije na autorska djela, što bi njihovo korištenje uvelo u domen primjene doktrine *fair use*.⁹³

LITERATURA:

Associated Press, “ChatGPT-maker OpenAI signs deal with AP to license news stories”, *The Associated Press*, 13. jul 2023., <https://www.ap.org/media-center/ap-in-the-news/2023/chatgpt-maker-openai-signs-deal-with-ap-to-license-news-stories/>, 27. mart 2024.

91 A. Trapova, P. Mezei, “Robojournalism – A Copyright Study on the Use of Artificial Intelligence in the European News Industry”, *GRUR Int.*, 7/2022, 590 i 600.

92 R. Yerramseti (2024).

93 *Ibid.*

- Brittain, B. "New York Times denies OpenAI's 'hacking' claim in copyright fight", *Reuters*, 12. mart 2024., <https://www.reuters.com/legal/litigation/new-york-times-denies-openais-hacking-claim-copyright-fight-2024-03-12/>, 20. april 2024.
- Brittain, B. "OpenAI hit with new lawsuits from news outlets over AI training", *Reuters*, 28. februar 2024., <https://www.reuters.com/legal/litigation/openai-hit-with-new-lawsuits-news-outlets-over-ai-training-2024-02-28/>, 19. maj 2024.
- Brittain, B. "OpenAI says New York Times 'hacked' ChatGPT to build copyright lawsuit", *Reuters*, 27. februar 2024., <https://www.reuters.com/technology/cybersecurity/openai-says-new-york-times-hacked-chatgpt-build-copyright-lawsuit-2024-02-27/>, 20. april 2024.
- Coster, H. "Global news publisher Axel Springer partners with OpenAI in landmark deal", *Reuters*, 13. decembar 2023., <https://www.reuters.com/business/media-telecom/global-news-publisher-axel-springer-partners-with-openai-landmark-deal-2023-12-13/>, 19. maj 2024.
- David, E. "Microsoft invokes VCRs in motion to dismiss The New York Times' AI lawsuit", *The Verge*, 6. mart 2024., <https://www.theverge.com/2024/3/5/24091719/microsoft-new-york-times-openai-motion-dismiss-copyright-lawsuit>, 15. maj 2024.
- David, E. "OpenAI strikes licensing deal with the magazine giant behind People", *The Verge*, 7. maj 2024., <https://www.theverge.com/2024/5/7/24151171/openai-dotdash-meredith-people-instyle-licensing>, 29. maj 2024.
- Dermawan, A. "Text and data mining exceptions in the development of generative AI models: What the EU member states could learn from the Japanese "nonenjoyment" purposes?", *The Journal of World Intellectual Property*, 1/2024, 44–68, <https://doi.org/10.1111/jwip.12285>, 12. maj 2024.
- Future of Life Institute (FLI), "EU Artificial Intelligence Act – Timeline of Developments", *Future of Life Institute (FLI)*, 4. jun 2024., <https://artificialintelligenceact.eu/developments/#:~:text=2023,28%20against%2C%20and%2093%20abstentions>, 4. jun 2024.
- Grynbaum, M. M., R. Mac, "The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work", *The New York Times*, 23. decembar 2023., <https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html>, 27. januar 2024.
- Henderson, P., X. Li, D. Jurafsky, T. Hashimoto, M. A. Lemley, P. Liang, "Foundation Models and Fair Use", *Stanford Law and Economics Olin Working Paper* No. 584, 29. mart 2023., <http://dx.doi.org/10.2139/ssrn.4404340>, 19. mart 2024.
- Henley, J. "OpenAI claims New York Times 'hacked' ChatGPT to build copyright lawsuit", *The Guardian*, 27. februar 2024., <https://www.theguardian.com/technology/2024/feb/27/new-york-times-hacked-chatgpt-openai-lawsuit>, 28. april 2024.

- Keary, A. "What is the Role of Parameters in AI", *Techopedia*, 28. novembar 2023, <https://www.techopedia.com/experts/what-is-the-role-of-parameters-in-ai>, 16. april 2024.
- Margoni, T., M. Kretchmer, "A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology", *GRUR Int.*, 8/2022, 686–701.
- Metz, C., K. Roberts, "Microsoft Seeks to Dismiss Parts of Suit Filed by The New York Times", *The New York Times*, 4. mart 2024., <https://www.nytimes.com/2024/03/04/technology/microsoft-ai-copyright-lawsuit.html>, 30. april 2024.
- Metz, C., K. Roberts, "OpenAI Seeks to Dismiss Parts of The New York Times's Lawsuit", *The New York Times*, 27. februar 2024., <https://www.nytimes.com/2024/02/27/technology/openai-new-york-times-lawsuit.html>, 30. april 2024.
- OpenAI, "Global news partnership: Le Monde and Prisa Media", *OpenAI*, 13. mart 2024b, <https://openai.com/index/global-news-partnerships-le-monde-and-prisa-media/>, 5. maj 2024.
- OpenAI, "OpenAI and journalism", *OpenAI*, 8. januar 2024a, <https://openai.com/index/openai-and-journalism/>, 30. april 2024.
- OpenAI, "We're bringing the Financial Times' world-class journalism to ChatGPT", *OpenAI*, 29. april 2024. <https://openai.com/index/content-partnership-with-financial-times/>, 19. maj 2024.
- Pope, A. "Blog Essay: NYT v. OpenAI: The Times's About-Face", *Harvard Law Review*, 10. april 2024., <https://harvardlawreview.org/blog/2024/04/nyt-v-openai-the-timess-about-face/>, 11. april 2024.
- Reuters, "OpenAI to use FT content for training AI models in latest media tie-up", *Reuters*, 30. april 2024., <https://www.reuters.com/technology/financial-times-openai-sign-content-licensing-partnership-2024-04-29/>, 12. maj 2024.
- Roberts, K. "8 Daily Newspapers Sue OpenAI and Microsoft Over A.I.", *The New York Times*, 30. april 2024., <https://www.nytimes.com/2024/04/30/business/media/newspapers-sued-microsoft-openai.html>, 12. maj 2024.
- Robertson, K. "OpenAI Strikes a Deal to License New Corp Content", *The New York Times*, 22. maj 2024., <https://www.nytimes.com/2024/05/22/business/media/openai-news-corp-content-deal.html>, 29. maj 2024.
- Sigalos, M., H. Field, "Microsoft compares The New York Times' claims against OpenAI to Hollywood's early fight against VCR", *CNBC*, 5. mart 2024., <https://www.cnn.com/2024/03/05/microsoft-seeks-dismissal-parts-of-new-york-times-suit-against-openai.html>, 29. april 2024.
- Trapova, A., P. Mezei, "Robojournalism – A Copyright Study on the Use of Artificial Intelligence in the European News Industry", *GRUR Int.*, 7/2022, 589–602.
- Wyn Davis, C., G. Dennis, "Getty Images v Stability AI: the implications for UK copyright law and licensing", *Pinsent Masons*, 29. april 2024., <https://www.pinsentmasons.com/expert-analysis-insight/getty-images-v-stability-ai>

www.pinsentmasons.com/out-law/analysis/getty-images-v-stability-ai-implications-copyright-law-licensing, 27. maj 2024.

U.S. Copyright Office, "U.S. Copyright Office Fair Use Index", U.S. Copyright Office, novembar 2023., <https://www.copyright.gov/fair-use/index.html>, 6. maj 2024.

Yerramseti, R. "Case study of AI in publishing industry — New York Times vs. OpenAI and Microsoft case", Medium, 12. januar 2024., <https://medium.com/@ryerrams2003/case-study-of-ai-in-publishing-industry-new-york-times-vs-openai-and-microsoft-case-b14bfc85fbd3>, 13. maj 2024.

Prof. Dr. Iza Razija Mešević

Associate Professor, University of Sarajevo, Faculty of Law

AI AND THE MEDIA: THE NEW YORK TIMES COMPANY V. OpenAI AND MICROSOFT

Abstract: The paper examines the first lawsuit filed by a media organization against representatives of the AI industry, namely the New York Times Company against OpenAI and Microsoft. The key issue in the background of this dispute is the training models of generative artificial intelligence without prior authorization or paying remuneration on the works protected by copyright that belong to media publishers. Furthermore, the plaintiff voiced claims that the AI products (e.g. ChatGPT) based on these models produce synthetic outputs that almost completely or to a large extent reproduce those works or represent their derivatives. The paper defines the basic concepts and categories essential for understanding the technology at the heart of the dispute (eg, generative artificial intelligence, fundamental models and large language models). In addition, it addresses the possible safe harbours for AI-developers in this context. Firstly, in the European Union, in terms of the text and data mining exception from Directive (EU) 2019/790, which is strongly backed by the text of the EU AI Regulation. Secondly, in the United States of America, in terms of applying the legal doctrine of „fair use“, especially in the sense of the transformative use of works. The author elaborates in detail the circumstances of the origin and the content of the dispute and looks at parallel cases in the domain of relationship between copyright and artificial intelligence, which could provide some guidance regarding the direction of the outcome of the NYTC- case. In addition, the paper draws attention to positive examples of OpenAI concluding agreements and even creating partnerships with large media organizations (e.g. Associated Press, Le Monde or News Corp) regarding the use of their content in GPT-model training, as well as in the context of technological cooperation.

Keywords: Artificial intelligence. – Copyright. – Media. – The New York Times. – OpenAI.